A Biocultural Examination of Student Learning Behaviors in Large Undergraduate Lectures

Daniel Z. Grunspan

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:
Steven M. Goodreau, Co-Chair
Daniel T.A. Eisenberg, Co-Chair
Darryl J. Holman

Program Authorized to Offer Degree:

Department of Anthropology

University of Washington

Abstract

“A Biocultural Examination of Student Learning Behaviors in Large Undergraduate Lectures”

Daniel Z. Grunspan

Co-Chairs of Supervisory Committee:

Professor Steven M. Goodreau

Professor Daniel T.A. Eisenberg

Department of Anthropology

Humans rely on culturally acquired information. Survival and reproductive success often hinge on whether or not individuals are culturally adapted to the environment they live in. These cultural variants can include important local social norms, foraging or hunting techniques, how to construct boats, among countless other examples. Given the importance of culture to human success, the field of Gene Culture Co-evolution suggests that humans have social learning strategies shaped by natural selection to preferentially adopt adaptive cultural variants while avoiding maladaptive ones. These predicted strategies can dictate *when*, *how*, and *from whom* to learn socially, with the fitness benefits of each defined by the state of the environment, the type of information being learned, and attributes of the individuals available to copy.

This dissertation empirically examines the social learning behaviors of college students longitudinally throughout a large lecture classroom. The first study examines who students identify as knowledgeable peers; identifying culturally competent individuals is a prerequisite ability to use a payoff-based bias, so students should be particularly attuned to who is doing well around them. Results reveal that students are more likely to be nominated as strong in the course material if they are performing well in the class, indicating a general student ability to know who knowledgeable peers are. However, it is also found that male students overrate their male peers after controlling for performance in the course and participation in class. Potential biological and cultural roots for this gender bias are discussed.

Using the perception networks from the first study, the second study examines whether students form learning relationships with peers they previously identified as strong in the material. This would indicate study group formation follows payoff-based biases. This study also tests whether student study groups follow homophily-based biases, a bias both predicted in social learning literature and supported in previous sociological research. These learning biases are supported; students are more likely to study with peers they previously identified as strong in the material, and students of the same ethnicity are also more likely than chance study together. As implied in its name, the use of payoff-based biases is expected to benefit the learner by giving them access to someone who is particularly skilled or knowledgeable. However, no evidence is found that student exam scores increase from having a study partner that was nominated as strong in the material.

The last study examines biocultural roots of learning preferences for social or individual learning. Here, social learning is described by passive learning through listening, while individual learning is more exploratory and involves more active work in having the learner self-discover information. Specifically, this study tests for an impact of a well-studied polymorphism of the gene

Dopamine Receptor D4 (*DRD4*). The 2R and 7R alleles of *DRD4* show positive selection in human migratory populations despite links to inattention, hyperactivity, different types of risk taking, and sensitivity to local norms and environments. This study lends evidence that these alleles may differentially impact learning preferences of students depending on the ethnic background of the student, suggesting a gene by environment interaction for the formation of learning styles, and a possible role of social learning dynamics to the evolution of *DRD4*.

This dissertation contributes to growing literature empirically testing human learning strategies. It also contributes to a small field of work applying evolutionary thought to improving educational practices, as well as work that examines social networks in classrooms. Thus, this work integrates several fields of research, and represents the value of interdisciplinary work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This dissertation is all about social learning, and how we come to know things. Specifically, I focus on the importance of the traits of “cultural models” – those from whom we can learn. Throughout the process of completing this dissertation, I have not only studied how humans navigate their network of potential cultural models, but have been lucky to find myself with access to some of the best cultural models possible. Excuse me as I effuse gratitude to these wonderful people.

I am unbelievably grateful to my two extremely committed, kind, and supportive co-chairs Steve Goodreau and Dan Eisenberg. I was never able to relate to PhD comics that poked fun at any of the annoying traits PhD advisors apparently have. I have their unbelievable mentorship to thank. I am also grateful to Darryl Holman for continued support throughout my graduate career, and for always entertaining the good, bad, and occasional ugly research ideas I would come up.

I can’t thank Emily Wojcik enough for being the incredible person, friend, and partner she has been. Her support, enthusiasm, sense of humor, patience, and karaoke performances are unmatched.

In 2011 I accepted a position as a teaching assistant in the Biology Department. Upon accepting this position, I received an e-mail from Ben Wiggins, in which he offered a rigorous TA position with the promise that I would develop my teaching, communication, and organizational skills. He was not lying. Ben is one of the best role models I have ever met, and none of this research would have happened had we never crossed paths. Unbelievably, I’ve never even seen him throw a frisbee.

I am grateful to have been a part of the faculty, staff, and students who make up the always growing Biology Education Research Group at UW. The content I have learned from Sara Brownell, Alison Crowe, Jennifer Doherty, Sarah Eddy, Scott Freeman, Jon Herron, Hannah Jordt, Joya Mukerji, Elli Theobald, and Mary Pat Wenderoth has always been at the top of Bloom’s taxonomy. I am especially thankful to have met Sarah Eddy, who not only exemplifies everything that makes a great academic, but also a great friend.

My first two years of graduate school were spent as an IGERT fellow in the IGERT Program for Evolutionary Modeling. In these years, I was treated to countless discussions with my advisor Eric Smith, and my cohort mates Kyle Bocinsky, Doug Barron, Brad Newbold, Bryan Roosien, and Daryl Trumbo. I am thankful for all these interactions, as they were foundational to the development of this dissertation.

This dissertation required a lot of drudging through messy raw data. I thank Dan Butler for both his programming expertise in handling this data. I also thank Katherine Cook, Sarah Davis, Arielle DeSure, and Carrie Sjogren for their hours of what can only be described as mind numbing data cleaning. I also thank Rebecca Ferrell, for introducing me to the amazing world of RMarkdown and showing me how first-hand how professionals perform statistical analyses.

I thank the wonderful friends and community I have had during my time in Seattle. I thank the Biocultural and Anthropology graduate student community for years of fascinating discussions and mediocre softball. I am grateful to Daniel Wershow and Elaine Katz for providing a family when my own were so far away. I owe much to my friend Hal Wershow, who epitomizes the benefits of always being optimistic, but having contagious amounts of passion for every endeavor.

Last, and certainly not least, I thank each and every student who participated in this research.

DEDICATION

This dissertation is dedicated to my parents, Avie and April Grunspan, for their endless love and support.

Chapter 1: Introduction

Overview

This dissertation studies human social learning behaviors through an evolutionary lens. It generates predictions about student learning behaviors in large lecture classrooms through Gene Culture Co-evolutionary theory (GCC). With this approach, each of the three research studies included has two main goals. First, this research empirically tests predictions from formal evolutionary models in GCC, a field which has had limited empirical testing. Thus, this work help fills the ongoing need for empirical tests of predictions arising from this theory. Second, by combining this evolutionary behavior approach with novel research methods, this research investigates social dynamics and learning diversity in college classrooms.

The first chapter of this dissertation focuses on who students perceive as strong with course materials. This chapter uncovers a strong gender discrepancy in how students perceive one another, and potential educational ramifications of this finding are discussed. An addendum to this chapter provides a post-hoc contextualization of these findings through a biocultural perspective.

The second chapter examines how study partnerships form in large classrooms, and explicitly tests whether patterns of study partnership formation and dissolution follow predictions from GCC. The perception networks in the first chapter, as well as other demographic factors, indeed impact who students study with. This chapter further examines whether studying with peers who are perceived as strong in the material has a positive impact on performance in the classroom.

The final chapter focuses on potential behavioral impacts of dopamine receptor alleles linked to Attention Deficit Hyperactivity Disorder (ADHD). I derive novel social learning based hypotheses for heightened frequencies of these alleles in populations with high rates of past migration. Based off of these hypotheses, I test the impact of these alleles on learning gains and engagement of students under two different forms of active learning compared to normal class days.

Together, these three chapters provide a broad scope of the social aspects of learning in large lecture classrooms. This work contributes an understanding of how learning behaviors and cultural transmission strategies are enacted within a classroom, and the impacts these behaviors have on students. It takes into account not only how different cultural and demographic backgrounds may impact behaviors, but how genetic factors may interact within this pathway.

Theoretical background

Humans exhibit complex culture and exceptional capabilities for learning. In every society, learned information is necessary for survival (Boyd, et al. 2011). In subsistence populations humans must know how to extract water, know the edible flora, and which fauna are beneficial to hunt. The importance of cultural learning does not disappear in Western cultures, where specialized expertise is often required to make a living and knowing local norms such as what side of the road to drive on is critical for survival.

Taking the importance of this extra-somatic knowledge to heart, GCC takes an evolutionary approach to how culture changes over time, and how this cultural evolution has an important dialogue with the evolution of human behaviors (Richerson and Boyd 2008). Culture is defined here, and for the rest of this dissertation, as any information that can be socially transmitted from one individual to another (Henrich and McElreath 2003; Richerson and Boyd 2008). The theoretical approach of GCC consists of three main arguments that drive research within the field: 1) Culture changes over time through a process that shares many characteristics with genetic evolution (Mesoudi 2011a; Richerson and Boyd 2008; Whiten, et al. 2011). 2) Human psychology has evolved learning rules that increases the adoption of locally adaptive culture. In turn, these learning rules further shape the evolution and frequency of cultural traits (Laland 2004; Rendell, et al. 2011), and 3) changes to the cultural landscape can lead to new genetic evolutionary pressures, further shaping human traits (Gerbault, et al. 2011; Laland, et al. 2001; Odling-Smee, et al. 2003). These three research paradigms are intricately related, as cultural, behavioral, and genetic evolution are all reliant upon one another. This dissertation focuses on human learning psychology and behavior, but the connected nature of the three main points in GCC means understanding one often requires understanding the others. Thus, to review the work on human learning psychology in GCC, one must first understand the first point: culture evolves.

Culture evolves:

Darwinian evolution rests on the existence of three pre-conditions. So long as variation, competition, and heritable traits exist, a population can be expected to evolve over time (Darwin 1872). Similar to replicating organisms, culture exhibits these three features, and likewise, changes in culture over time can be understood in a Darwinian manner. Important differences between genetic and cultural evolution do exist. For one, culture can change through intentional changes of humans, a process referred to as guided variation (Boyd and Richerson 1988; Richerson and Boyd 2008). Another key difference lies in transmission directions. While genetic inheritance is vertical from parent to offspring, culture transmits both vertically and horizontally

(Cavalli-Sforza and Feldman 1981a; Cavalli-Sforza, et al. 1982; Hewlett and Cavalli-Sforza 1986). For this reason, researchers have suggested using terms such as cultural “replication” or “transmission” as opposed to “inheritance” (Mesoudi, et al. 2004). A last important consideration is the non-particulate nature of culture. While it is conceptually helpful to model culture as if it existed in discrete units, as has been done with the concept of a “meme” (Dawkins 2006), this does not accurately represent the nature of culture.

Early attempts to study the evolution of cultures were unsophisticated, and in many ways prejudicial and damaging (Morgan 1877; Spencer 1857; Tylor 1871). Modern approaches to cultural evolution took place in the 1970s and 1980s in the form of formal mathematical models (Boyd and Richerson 1976; Boyd and Richerson 1988; Cavalli-Sforza and Feldman 1981b), which reject earlier assumptions that culture, as well as evolution, follows a linear trajectory towards sophistication . These cultural evolution models adopt methods from population genetics. These models simulate populations who possess a set of cultural traits, along with genetic rules dictating how individuals will learn these traits. Cultural trait frequencies change over time based on the transmission dynamics in the model, such as vertically from parents or horizontally between peers. By tracking how these frequencies change over time, these models formally predict how the relationship between transmission modes and cultural change should work, while integrating the role of genes and culture. This research has led to the birth of cultural phylogenies (Gray and Atkinson 2003; Gray and Jordan 2000; Mace and Holden 2005; Mace, et al. 1994), where empirical studies of culture change can help uncover likely past behaviors. Further, these models have been expanded to understand how some cultural traits spread more than others. Specifically, these models focus on how human behavior and psychology may have co-evolved along with culture to form learning biases which lead to the adoption of beneficial cultural traits (Laland 2004).

Social learning biases:

One way human learning psychology can lead to preferential adoption of useful culture is if we are predisposed to learn some kinds of content over others – a content bias (Richerson and Boyd 2008). It has been argued that human psychology has evolved to be more attracted to cultural units that serve a function towards reproductive fitness. For example, humans are more likely to share stories that include disgusting imagery, including rotten food (Heath, et al. 2001). An innate preference for learning and spreading information pertinent to survival instead of more esoteric information can help the fitness of an individual and their kin.

Another type of theorized learning bias is frequency dependence. Here, the prevalence of a cultural trait may make it more or less attractive to learners. An example of frequency dependent learning is conformist transmission (Boyd and Richerson 1988), which is a well-known and studied phenomenon in humans (Asch 1951; Sherif 1936). Natural selection often works to keep locally optimal traits; if a similar process acts on cultural variants, then the most common or “mean” cultural variant in a population can be expected to be locally optimal (or at least not maladaptive). In this instance, common cultural variants can be expected to be worth copying. Conformist transmission can result in runaway cultural selection (Boyd and Richerson 1988), where a commonly held trait quickly becomes the only cultural unit, which cannot be unseated.

A last class of learning biases is model-based, where the likelihood of a trait being transmitted is dependent on aspects of the individual who is transmitting. It may be easier or cheaper to learn from some individuals than others. At the same time, individuals vary in the quality and relevance of their cultural information. Thus, learning from individuals who are particularly successful, prestigious, older, or similar to the learner can represent useful strategies to obtain cultural variants that will make them more successful (Henrich and McElreath 2003; Laland 2004; Richerson and Boyd 2008).

Historically, most research in GCC on learning biases has been in the form of theoretical mathematical models. While it has taken time for empirical work to catch up to this theoretical work, it is now becoming more common. Empirical work to date has examined learning biases using both experimental lab studies and observational studies in both Western and subsistence populations (Atkisson, et al. 2012; Chudek, et al. 2012; Derex, et al. 2013; Efferson, et al. 2008; Efferson, et al. 2007; Henrich and Broesch 2011; Henrich and Henrich 2010; McElreath, et al. 2005; Mesoudi, et al. 2015b). An emergent finding as more empirical work has been done is that social learning strategies are not consistent across individuals or populations (Mesoudi 2016). Thus, an emerging question is why are there cross cultural and individual differences in learning strategies?

Cultural and Genetic differences in learning style

A recent review of formal evolutionary models of social learning behaviors points out that all of the most influential models make the assumption that learning strategies are genetically determined (Aoki and Feldman 2014). This is partially out of convenience – after all, all models are wrong (Box 1979), as they require making simplifying assumptions to gain tractability. In this case, these models utilize the phenotypic gambit (Heyes 2016). This means they avoid the concern of linking learning phenotypes to genotypes (Grafen 1984; Smith 1978). So long as learning styles are inherited, genetically, culturally, or otherwise, the main findings of these models are consistent. However, as evidence mounts that social learning behaviors differ between individuals, and between populations (Mesoudi, et al. 2015a), investigating the actual causes of variation in social information use (i.e. culturally learned or genetically inherited) becomes important (Mesoudi, et al. 2016).

Considerable work has looked at individual differences in learning strategies across many different species (reviewed by Mesoudi (2016)). Within humans, the use of social information has been linked to variation in personality of the learner (Mesoudi 2011b), past experiences with the person being learned from (Corriveau and Harris 2009), as well as cultural backgrounds (Berl and Hewlett 2015; Correa-Chávez and Rogoff 2009; López, et al. 2010; Mesoudi, et al. 2015a). The variation between cultures is of particular interest in this dissertation. To date, a main focus on cultural differences in social learning has between Eastern and Western cultures (Chang, et al. 2011; Mesoudi, et al. 2016; Mesoudi, et al. 2015a).

Cultures have long been categorized as being culturally collectivist or individualistic (Hofstede and Bond 1988; Markus and Kitayama 1991; Nisbett 2010). Collectivist cultural norms, common in countries such as China, Korea, and Japan, tend to emphasize individuals as being embedded within a larger group – a focus that inclines individuals to focus on others. Individualistic cultural norms are common in Western countries such as the United States, and focus instead on individual accomplishments and standing out among a group – a focus that inclines individuals to focus on themselves. It has been hypothesized that these collectivist and individualistic norms culturally evolved due to links to social and individual learning (Chang, et al. 2011). Chang, et al. (2011) argue that the history of Eastern populations include more stable environments, and thus have been favorable to increased social learning.

Even if these cultural differences in how the individual is construed did not emerge for their roles in social learning, the expression of these norms seems to be correlated with how often individuals utilize social information, with Eastern cultures showing a greater use of social information. New products have been shown to diffuse faster in collectivist cultures compared to individualistic cultures (Yaveroglu and Donthu 2002). Students in mainland China used social information more frequently in a computer based task compared to students in the United Kingdom and the more culturally Westernized Hong Kong (Mesoudi, et al. 2015a). These norms also seem to be reflected in educational systems, where personal discovery is a focus in Western educational systems, and respect for instructors and rote memorization a focus in Eastern educational systems (Chang, et al. 2011; Tweed and Lehman 2002). This last point is salient, as the learning and adoption of cultural norms often occurs in educational settings, which may make educational systems particularly important for enculturation. Research on cultural boundaries has shown that indigenous Matsigenka children who attend Mestizo schools are considerably less normative to Matsigenka culture than those who don't (John Bunce, personal communication). Thus, we may expect social learning strategies to differ according to cultural upbringing – we learn how to learn. However, to fully untangle the phenotypic gambit, the importance of genetic variation to social learning must be considered.

Dopamine Receptor D4:

While it is clear that culture plays an important role in human behavior, investigating potential genetic underpinnings of human learning psychology allots a more complex understanding of human phenotypes. Thus, approaches that account for genetic and cultural variation which underlie learning styles can move us towards an understanding of why we see this diversity. A potentially interesting gene with regards to human learning phenotypes is Dopamine Receptor D4 (*DRD4*).

DRD4 exhibits a 48 base pair variable number tandem repeat (VNTR) polymorphism (Van Tol, et al. 1992b). The most common variant across human populations is the 4R variant (four repeats of the 48 bp VNTR in the third exon) (Chang, et al. 1996; Wang, et al. 2004). Of particular interest are the second and third most common variants: 2R and 7R, which have associations with various endophenotypes. Both of these alleles are associated with attention deficit hyperactivity disorder (ADHD) (DiMaio, et al. 2003; Elia, et al. 2012; Gizer, et al. 2009; Grady, et al. 2003; Leung, et al. 2005b; Zhang, et al. 2012). 7R has been further linked to risk-taking (Dreber, et al. 2009; Dreber, et al. 2011), and with mixed results to novelty seeking (Munafò, et al. 2008).

While this work suggests a linear association between 2R and 7R alleles and differential behaviors, other work shows that associations may exist through differential susceptibility (Bakermans-Kranenburg, et al. 2008). Some of the evidence for this comes from research investigating levels of independence versus interdependence among American college students of either European American descent or Asian students born in Asia (Kitayama, et al. 2014). By surveying students, they found that students with the most interdependent self-construal were Asian born Asians who were carriers of either DRD4-2R or DRD4-7R, while the most independent students were carriers of European descent (Kitayama, et al. 2014). This has led Kitayama, et al. (2016) to suggest a norm-sensitivity hypothesis for *DRD4*. This hypothesis states that 2R and 7R alleles lead to individuals who are particularly norm congruent with the environment and culture they grow up in, potentially to aid in cooperative ventures where shared norms are important.

The behavioral links of the *DRD4* 48-bp VNTR are interesting in their own right, but are made much more interesting given the global distribution and timing of emergence of these alleles. Patterns of linkage disequilibrium in DRD4-7R suggest that it is derived from the 4R allele 40,000-50,000 years ago, while DRD4-2R appears to have emerged around 10,000 years ago (Matthews and Butler 2011), both critical times in human migratory history. Population frequencies of these two alleles correlate with that population's distance from East Africa, which suggests that migratory environments select for these alleles (Chen, et al. 1999; Matthews and Butler 2011).

Based on the range of work showing behavioral links between *DRD4* and behavior, as well as the temporal and geographic data associated with this VNTR, many researchers have suggested evolutionary benefits to either having a 2R or 7R allele, or having higher frequencies of individuals with these alleles in the population. Some of these include increased exploratory behaviors leading to better resource exploitation in migratory societies (Chen, et al. 1999), reduced reactivity to novel stressors in migratory settings (Matthews and Butler 2011), and increased norm-regulated cooperation in expanding populations (Kitayama, et al. 2016). Other similar variations of evolutionary accounts for *DRD4* exist. All of these accounts have connections to learning dynamics. Thus, GCC represents a powerful lens to help make sense of the high frequencies of 2R and 7R alleles.

A critical parameter in many cultural evolution models is the rate of environmental change. The frequency of social learning is expected to be different in stagnant environments compared to rapidly changing ones; when the environment changes quickly, socially available information may quickly become outdated, creating a need for individuals to explore the environment (Richerson and Boyd 2008). If the frontier environment encountered by migratory populations elicits a high rate of environmental change, then perhaps *DRD4* allele frequencies differ by environment due to a role of 2R and 7R in mediating social learning in these frontier settings. By considering the importance of learning dynamics in relation to differential impacts of *DRD4*, this dissertation combines GCC with *DRD4* is one of the first studies to explicitly link genetics to GCC, an important direction for the field to take (Richerson, et al. 2010).

GCC and the undergraduate classroom

In summary, GCC suggests that human psychology has evolved to navigate our complex cultural landscape. Across individuals and populations, the propensity for learning socially varies, and is likely dependent on both cultural and genetic inputs. Within this learning psychology, though, are possible innate learning biases that should attract us towards optimal and fitness enhancing information. Compared to the various environments in which humans have evolved in the past, the college classroom represents a very novel culturally constructed environment designed for the transmission of cultural information. With hypotheses made about how humans have evolved to learn, including a concentrated area on *who* to learn from, GCC offers a unique lens to study the social learning dynamics within college classrooms.

In 2014, 17.3 million students were enrolled in degree-granting institutions; a number that is only expected to increase (Kena, et al. 2015). College classrooms play a central role in the social experiences of undergraduates at these institutions. It has been established that social involvement in classrooms is critical to the success and persistence of students (Astin 1984; Astin 1993; Terenzini and Pascarella 1991; Tinto 1997). The President’s Council of Advisors on Science and Technology recently called for an increase in the use of empirically grounded pedagogical techniques in science classrooms, in hopes to increase the number of STEM graduates (Holdren and Lander 2012). Most evidence based pedagogical techniques largely increase the amount of peer interaction in classrooms compared to rote lecture. Thus, as more classrooms move towards active student centered pedagogy, understanding the social behaviors and psychology of students in a classroom increases in importance. This dissertation, beyond testing evolutionary theories of social learning, also contributes in this manner, by illuminating the social environment of undergraduate Biology students and the potential impacts these environments have on affect and learning.

Study Population

The data for all three studies that make up this dissertation come from the same design. Data collection was performed three times over three iterations of the same introductory biology classroom (Class A, Class B, and Class C). Each class had a split lecture design, where half of the enrolled students attended an early lecture, and the other half attended a late lecture. Class A was a 9 week course, while classes B and C were 10 weeks. Course lectures met for 50 minutes four times a week, and students met in smaller 24 person labs once a week for 2.5-3 hours. Each class consisted of four exams making up the majority of the class grade, with lab section grades, in class participation grades (assessed via clicker technology), online reading quizzes, and other small assignments in the course making up the rest of the grade.

Within class days, both lectures were taught as similarly as possible by the same instructor. The only exception to this occurred on days where active learning experiments were being run. Here, lectures were taught by the same instructor, but the two sections experienced different pedagogical techniques to teach the same material. These differences are discussed further in chapter three. Overall pedagogical style in each course focused on active learning, with heavy use of peer instruction (Mazur 1997). This teaching method makes peer interactions frequent, as students are discussing course material with each other throughout every class day.

The students in these courses are generally between 18 and 22 years old, and are for the most part are *WEIRD* (Western, Educated, Industrialized, Rich, and Democratic (Henrich, et al. 2010)). Thus, as with considerable research on human evolution (Arnett 2008), this research must be contextualized as representing contemporary University students. While this does limit extrapolations to human behavioral roots due to the environmental novelty, it is important to consider two strengths in this approach. First, this research is motivated in part to understand classroom interactions and improve college pedagogy. Second, as previously discussed, GCC recognizes the importance of both cultural and genetic factors in social learning behaviors. The cultural and genetic variation within large undergraduate classrooms allows certain hypotheses to be tested that could not be investigated in smaller subsistence settings.

Specific Aims:

This research aims to empirically test predictions from GCC in a novel environment. In doing so, this research contributes to our understanding of complex co-evolutionary processes and human social learning. This research also aims to contribute to our understanding of social environments of college classrooms, and how they may impact student outcomes.

> **Chapter 2** investigates who students perceive as strong in class material. This chapter focuses on gendered differences within the classroom, and their potential impacts on equity in science, technology, engineering, and mathematical fields. The addendum to this chapter contextualizes these results within literature on human evolution, and considers how this work contributes to biocultural Anthropology.

Chapter 3 investigates whether or not college student study partnership patterns follow context biases, and whether or not the use of these biases improves exam performance.

Chapter 4 investigates whether learning preferences and learning gains differ between instructional styles that are very active and those that are more passive for students who have different genetic and cultural backgrounds. This specifically tests whether *DRD4* impacts preference for more social or more individualistic pedagogical settings.

A final conclusion and discussion provides a synthesis of the body of research presented in this dissertation, and further discusses the use of GCC in thinking about higher education.

Chapter 2: Males under-estimate academic performance of their female peers in undergraduate biology classrooms

Daniel Z. Grunspan[1]

Sarah L. Eddy[2]

Sara E. Brownell[3]

Benjamin L. Wiggins[4]

Alison J. Crowe[4]

Steven M. Goodreau[1]

[1]Department of Anthropology, University of Washington, Seattle, WA 98185

[2]Texas Institute for Discovery Education in Science, College of Natural Sciences, University of Texas at Austin, Austin, TX 78712

[3]School of Life Sciences, Arizona State University, Tempe, AZ 85287

[4]Department of Biology, University of Washington, Seattle, WA 98195

Note: This chapter was published in PLoS ONE:

Grunspan DZ, Eddy SL, Brownell SE, Wiggins BL, Crowe AJ, Goodreau SM (2016) Males Under-Estimate Academic Performance of Their Female Peers in Undergraduate Biology Classrooms. PLoS ONE 11(2): e0148405. doi:10.1371/journal.pone.0148405

Abstract

Women who start college in one of the natural or physical sciences leave in greater proportions than their male peers. The reasons for this difference are complex, and one possible contributing factor is the social environment women experience in the classroom. Using social network analysis, we explore how gender influences the confidence that college-level biology students have in each other's mastery of biology. Results reveal that males are more likely than females to be named by peers as being knowledgeable about the course content. This effect increases as the term progresses, and persists even after controlling for class performance and outspokenness. The bias in nominations is specifically due to males over-nominating their male peers relative to their performance. The over-nomination of male peers is commensurate with an overestimation of male grades by 0.76 points on a 4 point grade scale, indicating a strong male bias among males when assessing their classmates. Females, in contrast, nominated equitably based on student performance rather than gender, suggesting they lacked gender biases in filling out these surveys. These trends persist across eleven surveys taken in three different iterations of the same Biology course. In every class, the most renowned students are always male. This favoring of males by peers could influence student self-confidence, and thus persistence in this STEM discipline.

Introduction

Male faculty members outnumber female faculty members in every science, technology, engineering, and math (STEM) discipline (National Science Foundation 2013). The attrition of female STEM students from their disciplines can be seen in early stages of the progression to STEM careers including the transition into college and graduate school (Gillen and Tanenbaum 2014). The experiences of women in STEM that may lead to this attrition can be subtle. It is generally no longer considered a matter of outright discrimination, but rather the accumulation of smaller experiences that determine whether a female student identifies with and persists in a scientific field (Eddy and Brownell 2015; Isbell, et al. 2012; Settles, et al. 2006; Swim, et al. 2001). One factor linked to persistence in STEM fields is self-confidence. This factor is heavily influenced by social interactions, particularly for women in historically male-dominated fields (Robnett and Leaper 2013; Usher and Pajares 2008; Zeldin and Pajares 2000). For example, a student whose abilities are endorsed by an influential person may experience increased performance and confidence; conversely, a student not receiving this affirmation experiences a decrease in performance and confidence (Cohen, et al. 1999; Rosenthal, et al. 2011). While formative experiences like these can occur throughout one's life, certain periods are more influential than others. Individuals are particularly attuned to cues confirming or discrediting their ability to succeed in a field during major transition periods (Cech, et al. 2011; Usher and Pajares 2008). Student experiences at one of those key transition points, introductory STEM courses, thus seem likely to influence current and future decisions to persist in STEM disciplines.

STEM faculty members provide some of the first professional feedback and interactions that students receive in their disciplines. Unfortunately, both male and female faculty members behave in ways that subtly favor males in STEM disciplines: (a) they are more likely to spend time mentoring males (Moss-Racusin, et al. 2012), (b) they are more likely to respond to emails from males (Milkman, et al. 2014), and (c) they are more likely to call on males in class (Eddy, et al. 2014). These subtle yet consistent biases appear to cause at least some female STEM students to experience a lower sense of belonging or confidence in their discipline, resulting in an increased tendency to leave science (Seymour, et al. 1997).

In addition to interactions with faculty members, interactions with other students could impact a student's sense of belonging and confidence in her discipline. In contrast to the work on gender biases among faculty, only limited research has been performed on the disposition of current college-age students (the "millennial" generation) towards women in STEM and how this disposition may impact their female peers (but see (Eddy, et al. 2015; O'Brien, et al. 2015)). Such research would not only help us to measure one force that may be acting to decrease undergraduate females' sense of belonging in STEM fields; it would also help us predict whether we can expect these implicit biases to persist in future STEM faculty.

In this paper we focus on the formative experience of nascent STEM professionals during an introductory college science course, a key transition period for the development of a STEM identity (Olson and Riordan 2012). We explore this question in a biology classroom. We chose this field because females and males enroll equally in this discipline at the undergraduate level (Wiens, et al. 2003) and thus should represent a conservative case of the biases women in introductory STEM courses experience.

We explore the impact of gender on how students perceive their peers, as well as how students are perceived by their peers. It is important to note that the gender data used in this study come from the school registrar, and are thus defined by information given during student enrollment. The registrar constrains choice for gender identification to 'male' or 'female' choices. Given these complications, we choose to refer to student genders, but recognize that in some cases the data may not accurately reflect the true gender identity of each student.

To investigate how gender impacts peer perception, undergraduate students were asked to anonymously list class peers who they felt were "strong in their understanding of classroom material" at multiple time points throughout three iterations of a large introductory biology class. We employ longitudinal social network analyses of these data to (1) describe the distribution of nominations received between males and females, and (2) identify the factors that predict who a student will nominate as having mastered the content in their field. Finally, (3) we examine the characteristics of students receiving the most nominations in each class (to whom we refer to as "celebrities"). We focus on these students given our assumption that their ability to draw widespread acknowledgement of their excellence makes them among the most likely in the class to continue in the field beyond the undergraduate level.

Materials and methods

Ethics Statement:

We obtained human subjects approval from the University of Washington Institutional Review Board (#44438). Because students were not asked to do anything outside of the normal class curriculum, an altered consent process was approved for use in this study. Subjects were informed that a research study was taking place and that their data would be analyzed as part of this study. Students were informed that they could opt out of the study at any time by filling out a form in a centralized office.

Classroom and Classroom Data:

Data come from three different iterations of the same large introductory undergraduate biology class (*n=196, 759, and 760, corresponding to class A, B, and C, respectively*) at a large American university that engages in very high research activity (an R1 university). The class of interest is the second in a series of three introductory Biology classes, where the first class in the series served as a pre-requisite. Because this was the second class, many students enter the class already knowing many of their peers. Student demographic information, including gender, was collected from the Office of the University Registrar and course grades from the Department of Biology.

All three iterations of this course included a lab section with a maximum capacity of 24 students that met once a week for several hours. Classes A, B, and C contained 9, 33, and 33 lab sections, respectively. The gender distribution within lab sections is approximately normal and mirrors that of the overall class (Mean = 57.4% female, SD = 0.11). The lecture portion of the course met for 50 minutes a day four days out of the week, and employed active learning techniques in all three iterations of the course. In all three cases, lectures were split into two sections with approximately 100 students in each for class A, and approximately 375 in each for Classes B and C; the instructor stayed consistent between lectures each class day to assure minimal differences between the two sections. Classes A and B were both taught by a male instructor, while Class C had three total instructors: two male instructors teaching 75% of class days and one female instructor teaching 25% of class days. All three iterations of the class included three exams spaced throughout the quarter, and a non-cumulative final exam that took place one week after the end of the quarter. Grades were not publically posted in any of the three classes.

A measurement of student outspokenness was collected by polling the course instructor of record immediately after the end of each course, and thus represents active participation as perceived by the instructor who was blind to the hypotheses being tested. Thus, a student who frequently offers an incorrect answer in class is considered equally outspoken as students who frequently offer the correct answers. Because measurements come from instructors, the list may be subject to each instructor's own implicit biases.

All three classes consisted primarily of white and Asian students (40.5% and 29.9% of entire population across the three classes, respectively). Student ethnicity is not included in these analyses for two reasons. First, the diversity in each classroom is such that statistical power to understand the perception of minority students is lacking. Second, this issue is substantial enough to warrant its own separate analysis.

Student networks:

All network surveys were administered via a confidential online survey. For Class A, students were given a class roster after the first and second exams and were asked to mark students they felt were particularly strong with class material. In Class B students were asked at the beginning of the class to list students by name who they felt would do particularly well in the course. After the first, second, and third exams, they were asked to list students they felt were particularly strong with class material. The same collection method was performed in Class C as Class B, but in addition students were surveyed again before the final exam of the course. Surveys in Class C distinguished between students who responded and didn't know anyone they felt were knowledgeable and students who didn't list anybody due to a non-response to the survey. Thus, Class C offers the most accurate means to calculate response rates. An average of 81.4% (SD = 0.02) of students responded across the five surveys in this class, with 82.8% of female students responding (SD = 0.02) and 79.9% of males responding (SD = 0.01). We have no reason to believe that Classes A or B differed in response rates, or that response rates were skewed by gender in any manner.

Analysis of nominations

To assess the hypotheses about nomination structure, we used exponential-family random graph models (ERGMs). This approach can be thought of as a kind of generalization of logistic regression to social networks – with the log-odds of a tie (here, a nomination) between two actors being dependent on a set of predictors of interest (Robins, et al. 2007). Those predictors may include characteristics of either or both nodes (e.g. their gender, class performance or outspokenness). However, it can also include structural factors involving the other ties in the network – e.g. the tendency for ties in a directed network to be mutual, or to form a triangle. When such structural terms are present, ties become conditionally dependent and estimation becomes more difficult, with Markov chain Monte Carlo-based methods the current state of the art for estimation (Hunter, et al. 2008). Nevertheless, the coefficients may still be interpreted in terms of their contribution to the conditional log-odds of a tie, given all of the other ties in the network.

We specify two models, both of the general form:

$$\text{logit}\left(Y_{ij} = 1 \left| y_{ij}^{c}\right.\right) = \sum_{k=1}^{n} \theta_{\,k}\,\delta_{\,k}$$

Where Y_{ij} represents the value of the tie from i to j, which equals 1 if i nominates j and 0 if they did not (we discuss missing data for these values in the SI). The quantity y_{ij}^{c} represents the complement of y_{ij}, i.e. the state of all of the ties in the network other than y_{ij}. The δ vector represents the amount by which the model statistics change when y_{ij} is toggled from 0 to 1, and the θ vector represents the coefficients on these statistics.

The first model contains seven model statistics (δ_1 through δ_7) and the second model contains nine (δ_1 through δ_9):

δ_1 = 1 for all dyads [the main effect or intercept];

δ_2 = 1 if j nominates i, and 0 otherwise [mutuality];

δ_3 = 1 if i is female and 0 otherwise [female nominator];

δ_4 = 1 if both i and j are female and 0 otherwise [female-female bias];

δ_5 = 1 if both i and j are male and 0 otherwise [male-male bias];

δ_6 = 1 if i and j are in the same lab section [lab homophily];

δ_7 = -1 if j has no nominations other than that from i [0-indegree]

δ_8 = j's final grade in the class [grade of nominee];

δ_9 = 1 if j is outspoken, and 0 otherwise [outspokenness of nominee];

We use the R package *network* to process and store the data, and the R package *ergm* to estimate the θ coefficients for our two models for each survey wave (Butts 2008; Hunter, et al. 2008). The terms involving gender, grade, or outspokenness represent our core theoretical measures. We include mutuality since it is a basic phenomenon in directed networks (those where the relationship from *i* to *j* does not necessarily equal that from *j* to *i*), and we include lab homophily given that labs are a major structural element of the course. We include a unique propensity for individuals to have no nominations (called 0-indegree in network terminology) since this dramatically improved the fit of the model to the observed in-degree distribution, which is a condition for the statistical inference we later conduct. Moreover, it is reasonable to expect that measures of renown such as that here would have more variation than expected by chance – that is, with more students who have either no or many nominations than otherwise expected. The δ on this term is negative given the unique condition that adding a tie reduces the statistic of interest (nodes without ties).

Results

Classroom data and student outspokenness

A summary of student data stratified by gender can be found in **Error! Reference source not found.** All three classes consisted of a numerical female majority; classes A, B, and C were 56%, 55.4%, and 58.4% female, respectively. Males averaged higher course grades than females in all three classes; the differences in grade in class C, but not A and B, is significant (p = 0.0171; unpaired t-test).

Proportionately, more males than females were listed as outspoken (p = 0.0258; Mantel-Haenszel test). While instructor bias causing this gender difference in outspoken status is something we cannot check, it is worth noting that any male bias in the assignment of outspoken status would make our estimates of male bias in peer perception more conservative than they actually are.

Males are over-nominated by peers as mastering biology.

Across the 11 peer perception surveys, students received an average of 1.20 nominations with a standard deviation of 1.85; males averaged 1.31 nominations with a standard deviation of 2.23, while females averaged 1.12 nominations with a standard deviation of 1.51. Males consistently received more nominations than females in every survey, with the first survey in Class C as the only exception.

In all three classes, the number of nominations given to males increased throughout the course. This pattern was particularly strong in Classes B and C, where data were collected across a longer time span. No consistent longitudinal trend for females is visible in any of the three classes. Combined, these patterns result in a growing gender gap in the number of nominations received between males and females when comparing data from surveys early in the class to those taken later in the class (**Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.**)

To determine the significance of these results, we use Exponential Random Graph Models (ERGM). Our base model does not include grade or outspokenness in order to give an absolute sense of the gender differences in receiving nominations (**Error! Reference source not found.**). In all 11 surveys, males show a significant bias toward nominating other males, with an absolute value greater than seen among females. In the last survey from Class A, females also show a bias towards nominating males, but show no significant bias either way in the remaining 10 surveys.

Class performance and outspokenness predict classroom wide recognition, but males still nominate more males after controlling for these.

Over-representation of males in received nominations could be explained either by the higher frequency of outspokenness in males, or the higher average grades achieved by males compared to females, as both of these measures may indicate that, on average, males indeed know the material better or at least make their knowledge more visible to their peers. To test these explanations, we expanded our ERGM to include the class grade and outspokenness of the nominee as mediating factors (**Error! Reference source not found.**).

Performance is a strong and significant predictor of receiving a nomination in every survey, indicating that students have an accurate sense of other students' performance, despite not having any public way to view their peers' grades. In addition, outspokenness has a significant effect in all but one case, indicating that students also nominate based on this trait. Being in the same lab section is also universally predictive of a nomination from one student to another. There is a significant tendency for there to be more students with no nominations than expected by chance given the overall nomination rates and the other terms in the model. The female nominator coefficient indicates that females make more nominations overall than males do, without considering the gender of those they nominate.

With performance and outspokenness in the model, females no longer show a bias toward nominating males in any of the 11 surveys; their nominations do not diverge from gender expectations in either direction in any survey. Males, on the other hand, continue to show a significant bias towards males in all 11 surveys; in each case the magnitude of the effect declined, but remains significant. **Error! Reference source not found.** shows the consequences of this inequity by simulating the nominations that would occur according to this model in a hypothetical classroom with a 1:1 gender ratio and equal mean class performance and outspokenness by gender.

Another way to understand the magnitude of the gender bias is to compare its coefficient to that for class grade point average (GPA), our best proxy for actual mastery of course material scored on a 4 point scale. Averaged across the 11 surveys, females give a boost to fellow females relative to males that is equivalent to an increase in GPA of 0.040; i.e. they would be equally likely to nominate an outspoken female with a 3.00 and an outspoken male with a 3.04. On the other hand, males give a boost to fellow males that is equivalent to a GPA increase of 0.765; for an outspoken female to be nominated by males at the same level as an outspoken male her performance would need to be over three-quarters of a GPA point higher than the male's. On this scale, the male nominators' gender bias is 19 times the size of the female nominators'.

Being Male is a Prerequisite for Celebrity Status

The three-to-four most nominated students in all classes examined were male. In each class, most students received very few nominations, while several students emerged over the course of the class as exceptionally well known; we refer to these students as "celebrities". Several patterns are evident in the distribution of nominations in these classes (**Error! Reference source not found.**). First, celebrity students tend to have high grades and speak up frequently in class. Second, with no exceptions, the biggest celebrity students in each network are male. While some females rank towards the top, the most well-known females are tied for 4th in two classes, and are 5th most well-known in the other. Third, male students at the top of the distribution tend to be considerably more well-known than any other student in the course. This is especially pronounced in Class B, where the most renowned male (52 nominations) received 5.78 times the nominations as the most renowned female (9 nominations). The most renowned male in Class A (16 nominations) has twice as many nominations than the most renowned female (8 nominations), while in Class C the most renowned male (13 nominations) has 1.63 times as many nominations than the most renowned female (8 nominations). These high nomination counts are notable, given the low average number of nominations seen across all 11 surveys (1.20). While the number of nominations achieved by celebrities in each class varies, the male biased pattern among the most frequently nominated peers holds.

The male majority among classroom celebrities could be explained if males were the only students who both achieved high grades and spoke up frequently in class. However, this is not the case. While male students on average scored slightly higher than female students and were more likely to be outspoken in every class, outspoken females with grades as high as these most renowned male students exist in every class (**Error! Reference source not found.**, **Error! Reference source not found.**, and **Error! Reference source not found.**). However, females achieving high grades and outspoken status never gain the same celebrity status as their male counterparts. It appears that being male is a prerequisite for students to achieve celebrity status within these classrooms.

Discussion

The underrepresentation of women in STEM is a complex and daunting problem. Increasing gender equity requires tackling both inequalities in students' initial interest in STEM and the retention of women who have expressed that interest. While there is strong evidence that precollege factors influence a student's initial decision to major in a STEM field (Ceci and Williams 2011), the causes of attrition after students initially declare a STEM major are less commonly explored. Studies on attrition of STEM-oriented women have found sense of belonging (Seymour, et al. 1997), decisions to start families (Ceci and Williams 2011) , and confidence that one can succeed in one's chosen profession (11) all influence a woman's decision to leave STEM. In particular, professional confidence is lower for women in STEM than for men (Cech, et al. 2011). This confidence is influenced by many inputs, but one of the major ones, especially for women in STEM fields, is receiving verbal messages and encouragement from individuals with influence, such as teachers and peers (Zeldin and Pajares 2000). Unfortunately, current faculty hold a gender bias that impacts the experiences of women in STEM (Moss-Racusin, et al. 2012) and could result in less support from faculty. Here we demonstrate that the peers of female students in introductory biology classes can also exhibit gender biases, adding to the list of subtle experiences that can lead to the attrition of females from STEM careers.

In three iterations of an undergraduate biology class, we found that even after controlling for actual course performance and outspokenness, male peers still disproportionately nominate males as being knowledgeable about biology while females nominate males and females equally. This indicates that males hold a bias against their female peers' competence in biology. Our finding of peers as a second source of differential treatment by gender, beyond known biases of faculty, contributes to a more complete picture of the experiences of undergraduate women in STEM fields. The coalescence of subtle messages about their STEM abilities from both faculty and peers may undermine the self-confidence females have to persist in STEM fields beyond their undergraduate education (Becker 1984).

The finding that a gender bias impacts the perception of millennial students may at first seem surprising, but is supported by work on implicit biases. Implicit biases are unconscious associations that people hold related to certain groups. Across many cultures, STEM is associated with males and not females (Nosek, et al. 2002a). Interestingly, male STEM majors in the US hold the strongest associations between maleness and science, while female STEM majors show some of the weakest implicit biases between gender and science (Nosek and Smyth 2011). These differences in the gender-STEM stereotypes held may explain why male undergraduate STEM majors nominate more males, but females do not demonstrate this bias. It also helps explain why male faculty demonstrate biases in hiring and mentoring, but many female STEM faculty do not (Sheltzer and Smith 2014).

One potential analytical concern for the current study is multiple comparisons. This occurs when statistical analyses involves multiple outcome measures, testing for an effect of multiple independent variables on a single outcome measure, or when the research design is repeated across several populations. In each case, the chance of finding a false positive is increased by adding another test. Because we repeated our study design three times and include multiple independent variables in our models, we are performing multiple tests, and thus have increased chances of a false positive. However, the repeated significance of our main result (that males over-nominate their male peers) across every survey gives us no reason to suspect that they are spurious due to multiple comparisons. It appears that males consistently hold a bias against their female peers' competence in biology.

The classroom environment can influence student perceptions of their peers

Our work suggests that processes in the classroom may either be reinforcing pre-existing implicit biases over the quarter, or at least facilitating behaviors based on these biases. The end of every class term shows a stronger male bias than the beginning. This pattern is mediated by two class-related factors: 1) whether or not a student is outspoken in class and 2) level of achievement in the class. These factors, which seem to influence the opinions of both male and female peers, have previously been found to differ by gender in biology: males are more likely to be heard speaking in class and males slightly, but systematically, outperform women (Eddy, et al. 2014). Instructors may be able to interrupt this process by equalizing the rate at which students of all genders speak up in class, closing the achievement gaps in their classrooms, or using more student-centered instruction in ways that do not rely primarily on whole class discussions (e.g. small group-work only).

We propose that the specific classroom environment can influence the effect size of the male bias, with some support for this hypothesis from Class C. In this term males did not behave differently than in previous years. Females, however, developed a stronger bias towards nominating other females than in the other two classes. Though this bias was not significant, it effectively lessens the overall magnitude of bias towards male students. Although we cannot specifically pinpoint why this was the case, this class differed from the other two in two critical ways. First, one of the three instructors in this course was female, whereas all instructors were male in the other two classes. Female instructors, when they are considered role models, have been shown to reduce the science-gender biases of female students, and this may have impacted the latter's nomination patterns (Young, et al. 2013). Second, during classroom discussions, all three instructors in Class C employed 'random call,' in which the instructor selects students to speak based on a randomized class list rather than by choice, more extensively than in the two other classes. Random call has been shown to eliminate the gender gap in class volunteers, leading to more females speaking up in class (Eddy, et al. 2014) and limiting the opportunity for one student to dominate classroom conversations and the instructors attention (Martin, et al. 2006). These differences in Class C seem to indicate that other factors in the classroom environment could mitigate the extent to which gender and renown are correlated. It is important to keep in mind that this mitigation seems to come from a larger female-female bias. This counteracting gender bias is likely undesirable compared to eliminating the male-male bias and achieving complete gender equity. Further research is needed to understand how to best achieve this equity in peer perception.

Biology is a conservative case; patterns may be more extreme across STEM

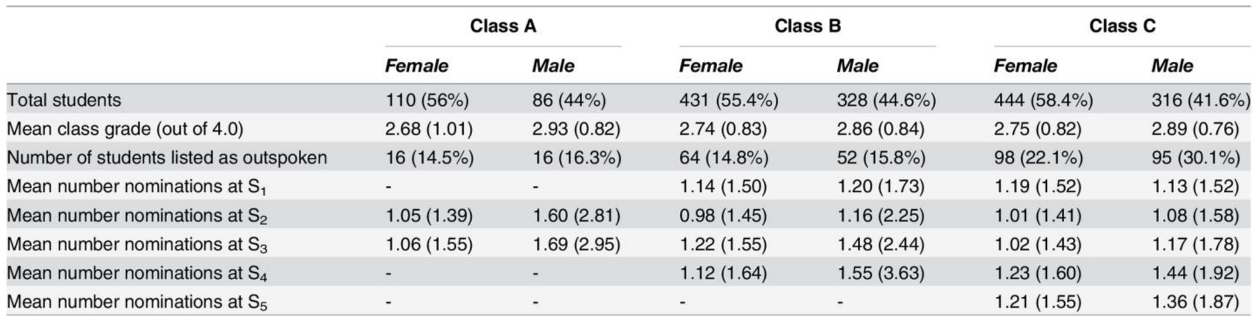
The context of this research on peer perceptions was an introductory biology classroom. We can only speculate on the peer biases present in other STEM fields, but we predict that the male bias observed in this study may be conservative relative to other STEM fields for three reasons. First, biology is thought to be the STEM field with the most gender equity: undergraduate enrollment is nearly equal in terms of males and females (Board 2014) and slightly more women than men earn degrees in the biological sciences. Thus females in biology do not have to contend with the biases associated with being the sole representative of their gender in a STEM classroom (Thompson and Sekaquaptewa 2002). Second, there is also a perception that biology lacks a strong math basis, and does not invoke the math-gender stereotype as strongly (Schmader, et al. 2004). Thus, stereotypes about women's math ability may not be undermining how their peers perceive them to the same extent it might in more explicitly math-based fields like physics or computer science. Finally, biology is a field that people believe does not require "brilliance", unlike other STEM fields (Leslie, et al. 2015). This perception means that stereotypes that males are more intelligent may not impact peer perceptions as strongly as it does in fields that are considered to require brilliance, like physics and math. For these reasons, we argue that the gender inequities in peer perception in the classrooms presented in this paper are likely conservative compared to classrooms in other STEM fields. Further, this dynamic may exist beyond STEM fields. However, explicit tests are required to confirm these hypotheses.

Conclusion

Our findings have strong implications regarding the effectiveness of existing strategies to increase women in STEM fields. Without addressing social dynamics that perpetuate gender biases in the college classroom, simply increasing the number of young women entering STEM majors may not be enough. The patterns of uneven peer perceptions by gender shown in our student population suggest that future populations of academics may perpetuate the same gender stereotypes that have been illuminated among current faculty. This may not only be the case because the male students receiving high celebrity are reaffirmed in their abilities and are better able to advance through the STEM pipeline than women who do not receive this affirmation, but also because the existence of "celebrity" males and other individuals with distinction can impact and reaffirm the stereotypes held by others (Rothbart, et al. 1978). This gender biased pattern in celebrity was experienced by over 1,500 students in our analyses. This number is striking, but less worrisome than the millions of students who attend college STEM classes that may perpetuate the same biases described here. In addition to current impacts on the peers in their classes, the students in these classes are potential future faculty members. Although we cannot directly compare the magnitude of gender bias between current faculty and millennial students, our work implies that the chilly environment for women may not be going away any time soon.

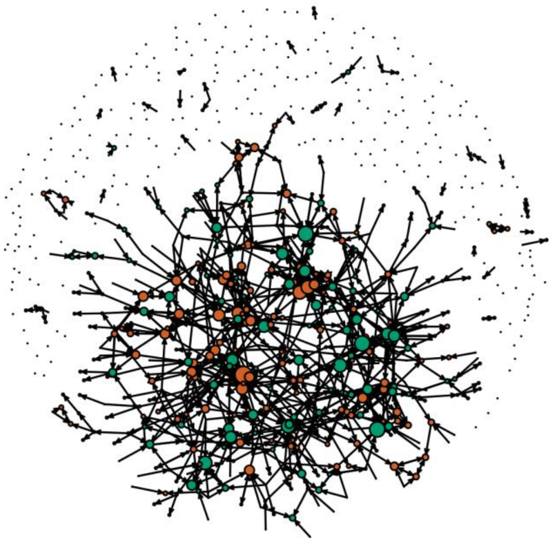
Table 2.1: Student demographics from all three classes.

Classes are majority female in all three cases. Males performed slightly better than females in each class, and also tended to be more outspoken. Numerical counts are accompanied by total percentage in the class in parentheses. Means are accompanied by standard deviations in parentheses.



	Class A		**Class B**		**Class C**	
	Female	***Male***	***Female***	***Male***	***Female***	***Male***
Total students	110 (56%)	86 (44%)	431 (55.4%)	328 (44.6%)	444 (58.4%)	316 (41.6%)
Mean class grade (out of 4.0)	2.68 (1.01)	2.93 (0.82)	2.74 (0.83)	2.86 (0.84)	2.75 (0.82)	2.89 (0.76)
Number of students listed as outspoken	16 (14.5%)	16 (16.3%)	64 (14.8%)	52 (15.8%)	98 (22.1%)	95 (30.1%)
Mean number nominations at S_1	-	-	1.14 (1.50)	1.20 (1.73)	1.19 (1.52)	1.13 (1.52)
Mean number nominations at S_2	1.05 (1.39)	1.60 (2.81)	0.98 (1.45)	1.16 (2.25)	1.01 (1.41)	1.08 (1.58)
Mean number nominations at S_3	1.06 (1.55)	1.69 (2.95)	1.22 (1.55)	1.48 (2.44)	1.02 (1.43)	1.17 (1.78)
Mean number nominations at S_4	-	-	1.12 (1.64)	1.55 (3.63)	1.23 (1.60)	1.44 (1.92)
Mean number nominations at S_5	-	-	-	-	1.21 (1.55)	1.36 (1.87)

Class B, Week 1 (S1)



Class B, Week 7 (S4)

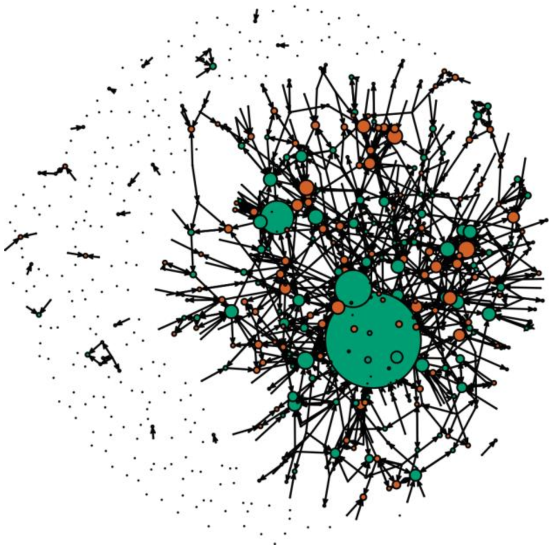
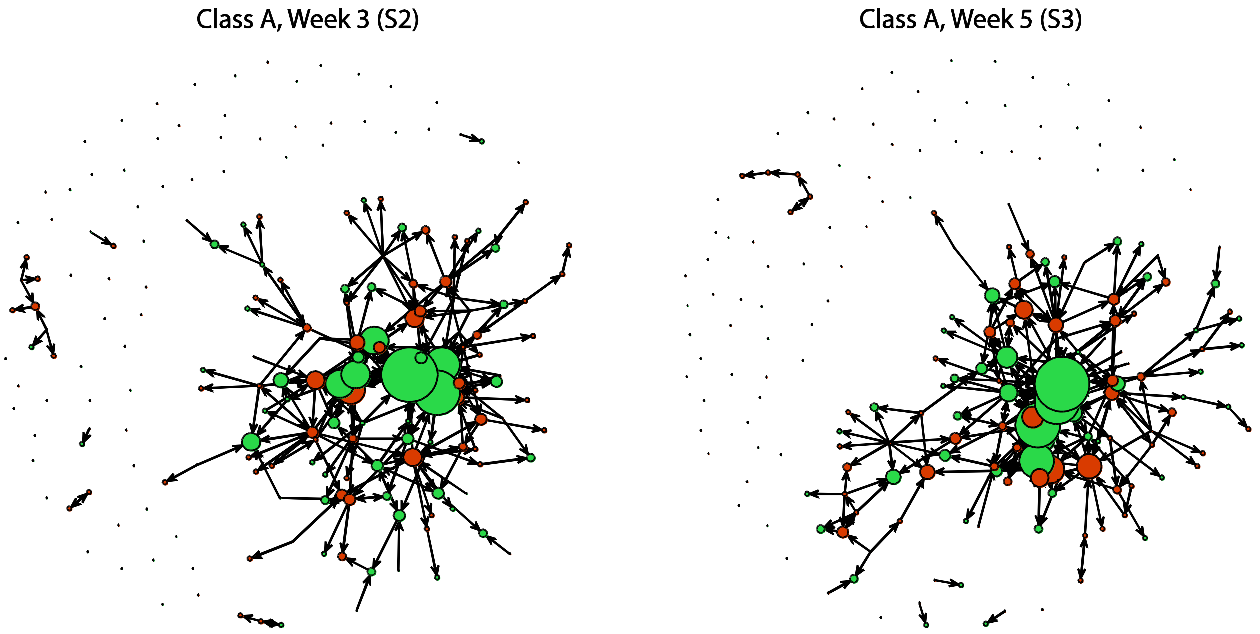


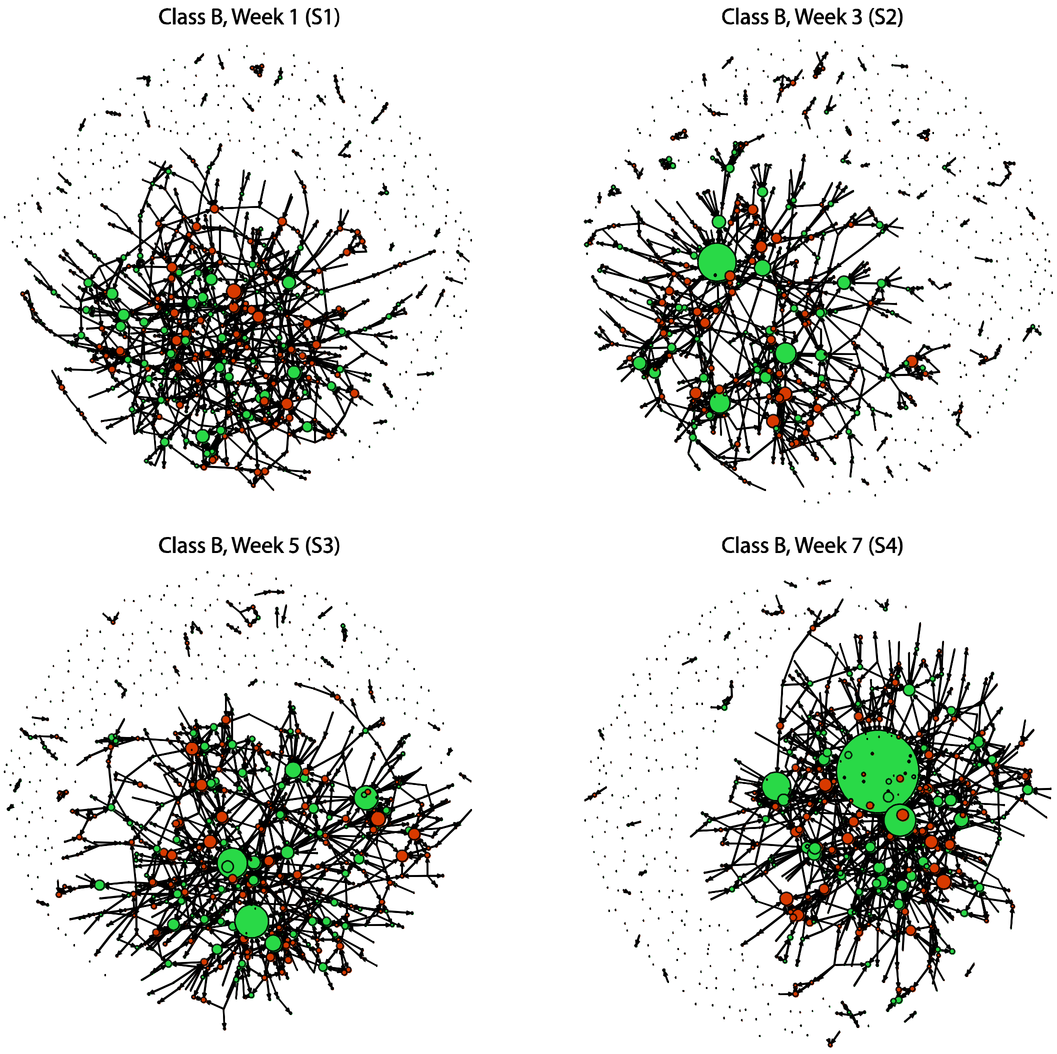
Figure 2.1: Unequal distribution of peer perception of mastery of content among genders grows over the term. Sociographs at the beginning of course (S1) and after exam 3 (S4) in class B. Male students are represented by green circles and females by orange circles. The size of nodes correlates with how many nominations each student received. Arrows show direction from the nominator to the nominee.



Class A, Week 3 (S2)

Class A, Week 5 (S3)

Figure 2.2: Unequal distribution of peer perception of mastery of content among genders grows over the term in Class A. Sociographs after the first exam (S2) and after exam 2 (S3) in class A. Male students are represented by green circles and females by orange circles. The size of nodes correlates with how many nominations each student received. Arrows show direction from the nominator to the nominee.



Class B, Week 1 (S1)

Class B, Week 3 (S2)

Class B, Week 5 (S3)

Class B, Week 7 (S4)

Figure 2.3: Unequal distribution of peer perception of mastery of content among genders grows over the term in Class B. Sociographs at the beginning of the class and after the first three exams in class B. Male students are represented by green circles and females by orange circles. The size of nodes correlates with how many nominations each student received. Arrows show direction from the nominator to the nominee.

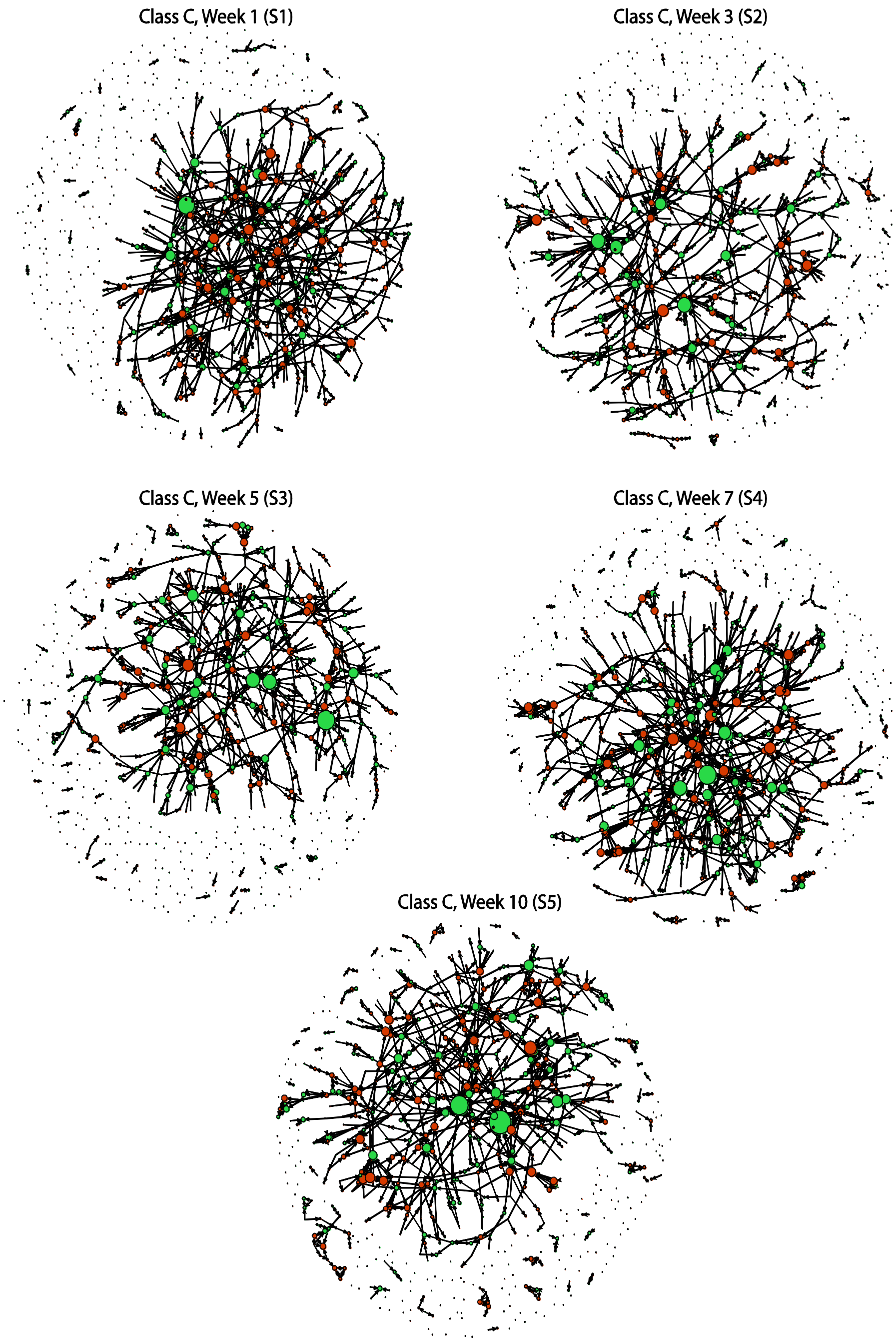


Figure 2.4: Unequal distribution of peer perception of mastery of content among genders grows over the term in Class C. Sociographs at the beginning until the end of class C. Male students are represented by green circles and females by orange circles. The size of nodes correlates with how many nominations each student received. Arrows show direction from the nominator to the nominee.

Table 2.2: ERGM results fro mbase model.

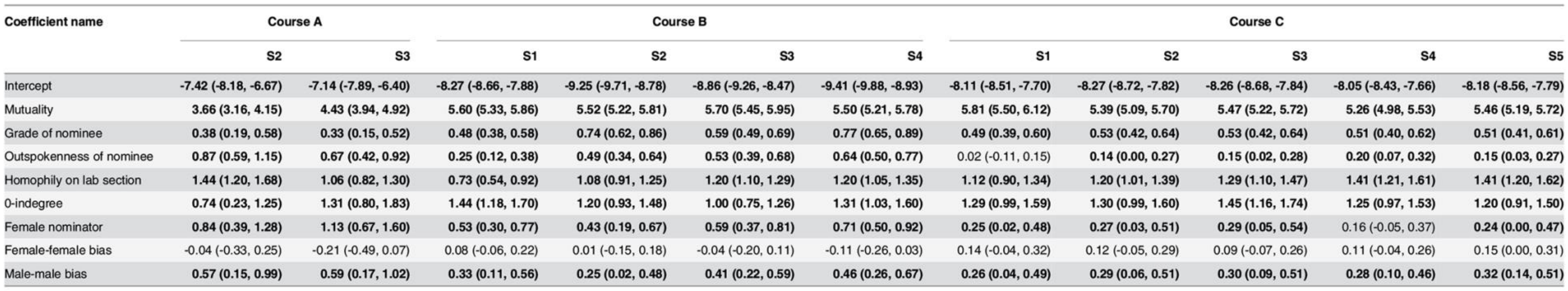
Model 1 shows ERGM results showing the effect of gender on the likelihood of a nomination for all 11 networks. This model controls for mutuality, and thus takes into account the increased likelihood of a nomination from student A to student B, given a nomination from B to A. This model shows the gender bias in nominations before taking into account outspokenness and class performance.

Coefficient name	Course A		Course B				Course C				
	S2	S3	S1	S2	S3	S4	S1	S2	S3	S4	S5
Intercept	**-6.03**	**-6.22**	**-7.19**	**-7.19**	**-7.11**	**-7.16**	**-6.89**	**-6.99**	**-7.01**	**-6.71**	**-6.84**
Mutuality	**4.10**	**4.59**	**5.60**	**5.77**	**6.04**	**5.64**	**5.84**	**5.53**	**5.67**	**5.42**	**5.74**
Female nominator	**1.14**	**1.43**	**0.63**	**0.57**	**0.70**	**0.90**	**0.28**	**0.37**	**0.41**	**0.28**	**0.33**
Female-female bias	**-0.35**	**-0.56**	-0.01	-0.13	**-0.21**	**-0.32**	0.11	0.07	0.01	-0.01	0.06
Male-male bias	**0.88**	**0.90**	**0.48**	**0.44**	**0.42**	**0.66**	0.23	**0.35**	**0.48**	**0.44**	**0.42**

S = survey number

Table 2.3: Coefficients from exponential random graph models from all 11 networks across all three courses, demonstrating that female bias towards nominating other females is not significant in any survey, while male bias towards nominating males is significant in all 11 surveys.

Each column represents coefficients from a different survey (S): S1 surveys were taken the first week, S2 after the first exam, S3 after the second exam, and so on. Coefficients represent the influence on the log-odds of a nomination for each predictor; each is formally defined in the Methods section. Bolded coefficients indicate significance at α of 0.05. Positive coefficients indicate that ties are more likely to occur, while negative coefficients indicate that ties are less likely to occur. Values in parentheses represent 95% confidence intervals.



Coefficient name	Course A		Course B				Course C				
	S2	S3	S1	S2	S3	S4	S1	S2	S3	S4	S5
Intercept	**-7.42 (-8.18, -6.67)**	**-7.14 (-7.89, -6.40)**	**-8.27 (-8.66, -7.88)**	**-9.25 (-9.71, -8.78)**	**-8.86 (-9.26, -8.47)**	**-9.41 (-9.88, -8.93)**	**-8.11 (-8.51, -7.70)**	**-8.27 (-8.72, -7.82)**	**-8.26 (-8.68, -7.84)**	**-8.05 (-8.43, -7.66)**	**-8.18 (-8.56, -7.79)**
Mutuality	**3.66 (3.16, 4.15)**	**4.43 (3.94, 4.92)**	**5.60 (5.33, 5.86)**	**5.52 (5.22, 5.81)**	**5.70 (5.45, 5.95)**	**5.50 (5.21, 5.78)**	**5.81 (5.50, 6.12)**	**5.39 (5.09, 5.70)**	**5.47 (5.22, 5.72)**	**5.26 (4.98, 5.53)**	**5.46 (5.19, 5.72)**
Grade of nominee	**0.38 (0.19, 0.58)**	**0.33 (0.15, 0.52)**	**0.48 (0.38, 0.58)**	**0.74 (0.62, 0.86)**	**0.59 (0.49, 0.69)**	**0.77 (0.65, 0.89)**	**0.49 (0.39, 0.60)**	**0.53 (0.42, 0.64)**	**0.53 (0.42, 0.64)**	**0.51 (0.40, 0.62)**	**0.51 (0.41, 0.61)**
Outspokenness of nominee	**0.87 (0.59, 1.15)**	**0.67 (0.42, 0.92)**	**0.25 (0.12, 0.38)**	**0.49 (0.34, 0.64)**	**0.53 (0.39, 0.68)**	**0.64 (0.50, 0.77)**	0.02 (-0.11, 0.15)	**0.14 (0.00, 0.27)**	**0.15 (0.02, 0.28)**	**0.20 (0.07, 0.32)**	**0.15 (0.03, 0.27)**
Homophily on lab section	**1.44 (1.20, 1.68)**	**1.06 (0.82, 1.30)**	**0.73 (0.54, 0.92)**	**1.08 (0.91, 1.25)**	**1.20 (1.10, 1.29)**	**1.20 (1.05, 1.35)**	**1.12 (0.90, 1.34)**	**1.20 (1.01, 1.39)**	**1.29 (1.10, 1.47)**	**1.41 (1.21, 1.61)**	**1.41 (1.20, 1.62)**
0-indegree	**0.74 (0.23, 1.25)**	**1.31 (0.80, 1.83)**	**1.44 (1.18, 1.70)**	**1.20 (0.93, 1.48)**	**1.00 (0.75, 1.26)**	**1.31 (1.03, 1.60)**	**1.29 (0.99, 1.59)**	**1.30 (0.99, 1.60)**	**1.45 (1.16, 1.74)**	**1.25 (0.97, 1.53)**	**1.20 (0.91, 1.50)**
Female nominator	**0.84 (0.39, 1.28)**	**1.13 (0.67, 1.60)**	**0.53 (0.30, 0.77)**	**0.43 (0.19, 0.67)**	**0.59 (0.37, 0.81)**	**0.71 (0.50, 0.92)**	**0.25 (0.02, 0.48)**	**0.27 (0.03, 0.51)**	**0.29 (0.05, 0.54)**	0.16 (-0.05, 0.37)	**0.24 (0.00, 0.47)**
Female-female bias	-0.04 (-0.33, 0.25)	-0.21 (-0.49, 0.07)	0.08 (-0.06, 0.22)	0.01 (-0.15, 0.18)	-0.04 (-0.20, 0.11)	-0.11 (-0.26, 0.03)	0.14 (-0.04, 0.32)	0.12 (-0.05, 0.29)	0.09 (-0.07, 0.26)	0.11 (-0.04, 0.26)	0.15 (0.00, 0.31)
Male-male bias	**0.57 (0.15, 0.99)**	**0.59 (0.17, 1.02)**	**0.33 (0.11, 0.56)**	**0.25 (0.02, 0.48)**	**0.41 (0.22, 0.59)**	**0.46 (0.26, 0.67)**	**0.26 (0.04, 0.49)**	**0.29 (0.06, 0.51)**	**0.30 (0.09, 0.51)**	**0.28 (0.10, 0.46)**	**0.32 (0.14, 0.51)**

Cell entry = Point estimate (95% CI); S = survey number; Bold = significant ($p < 0.05$).

Note: for Course C Survey S5, the lower bound of the confidence interval is positive for female nominator and negative for female-female bias, although both round to 0 with only two decimal places.

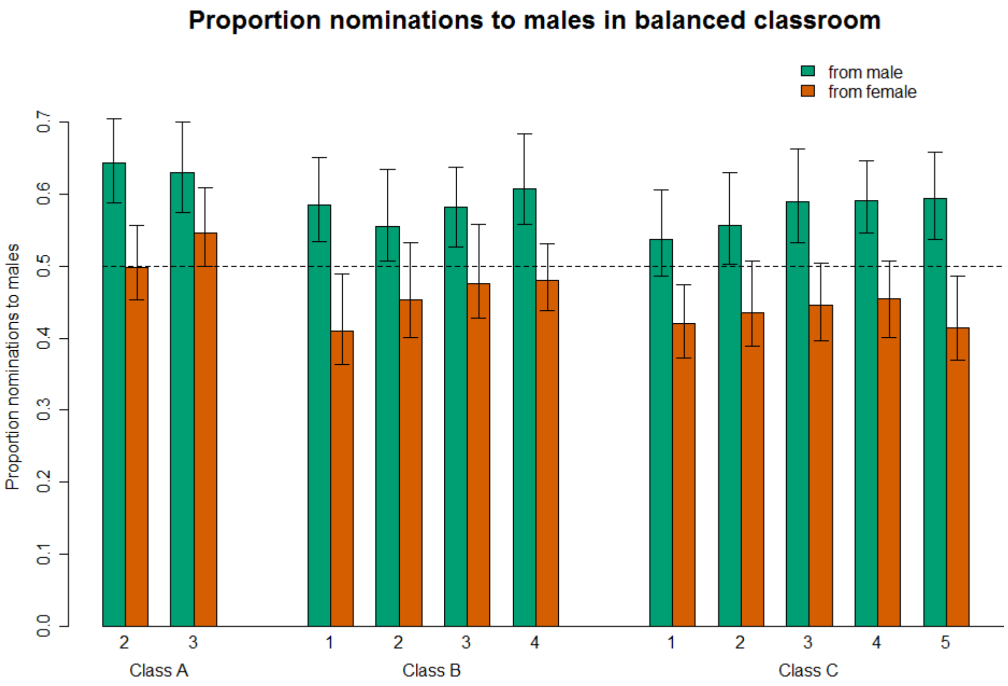


Figure 2.5: Males over-nominate males; females are closer to equitable in their nominations.

Model based predictions for a hypothetical class comprising 50% males and 50% females. To isolate the effect of gender bias this class was also modeled as having an equal grade distribution and level of outspokenness across genders. We plot the results from 100 simulations for each of the models; the main bars represent the mean, and the whiskers reflect the range in which the central 95% of the simulations fall. Even with equal performance and outspokenness in this hypothetical class across all three model predictions, the longitudinal increase in bias of male students to nominate males remains. Female students also demonstrate a pattern of moving from female to male nominations over the course of each class.

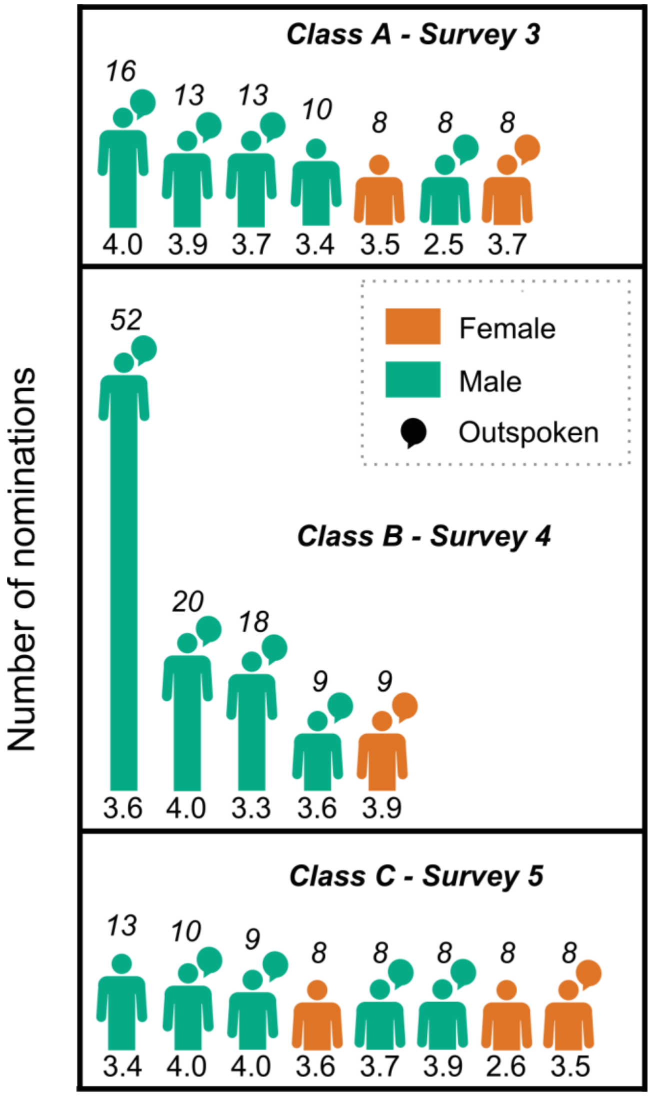


Figure 2.6: The most renowned students in each class tend to be male.

Students with the five highest numbers of nominations are depicted for each class. The numbers above each student represent how many nominations that student received, while the numbers below each student represent their grade point average earned in the course out of 4 points. These data come from the last surveys administered in Classes A, B, and C, and represent our best estimate for the perceptions developed by the end of each class.

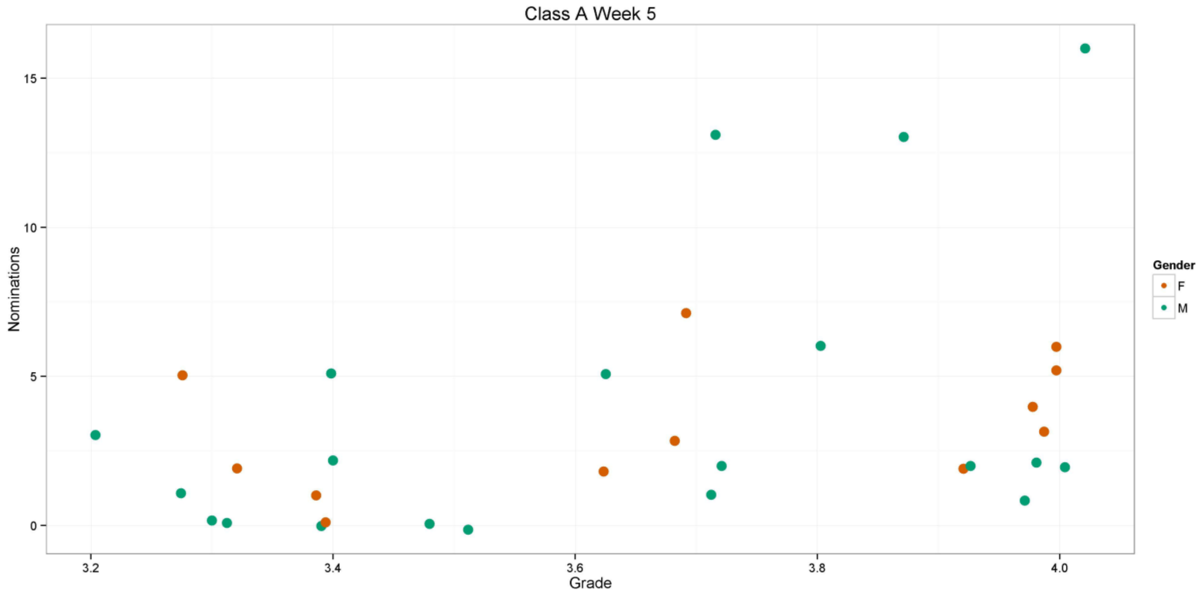


Figure 2.7: Outspoken students who scored in the top 10% of class and nominations earned at the last survey collection in class A.

Even though outspoken females with extremely high scores exist, they fail to reach the same “celebrity” status as their male counterparts.

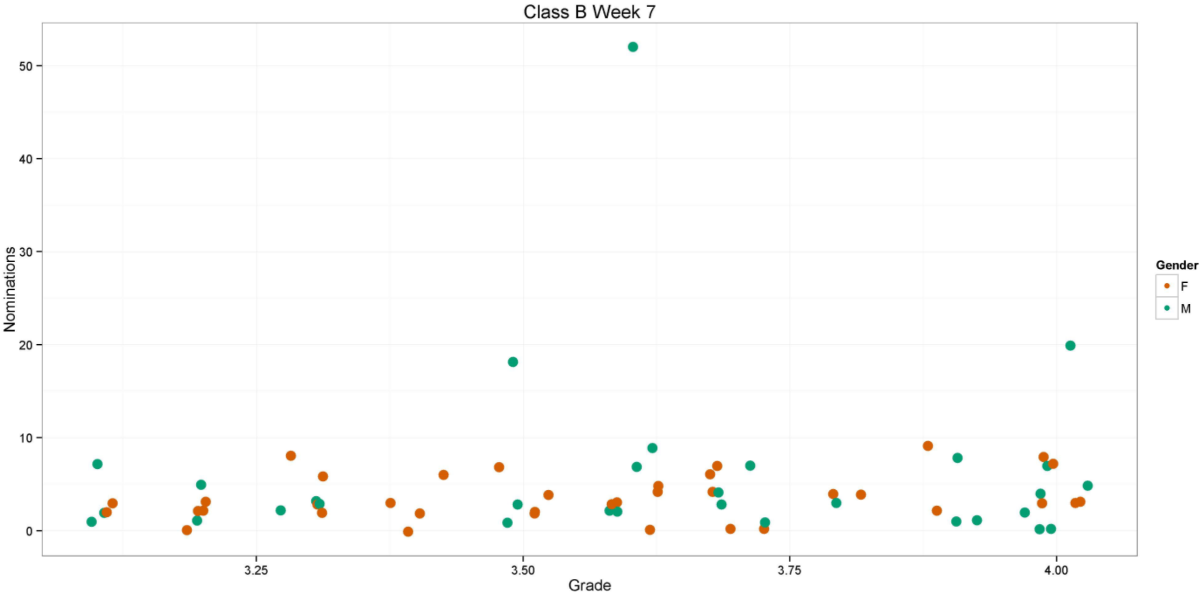


Figure 2.8: Outspoken students who scored in the top 10% of class and nominations earned at the last survey collection in class B.

Even though outspoken females with extremely high scores exist, they fail to reach the same "celebrity" status as their male counterparts.

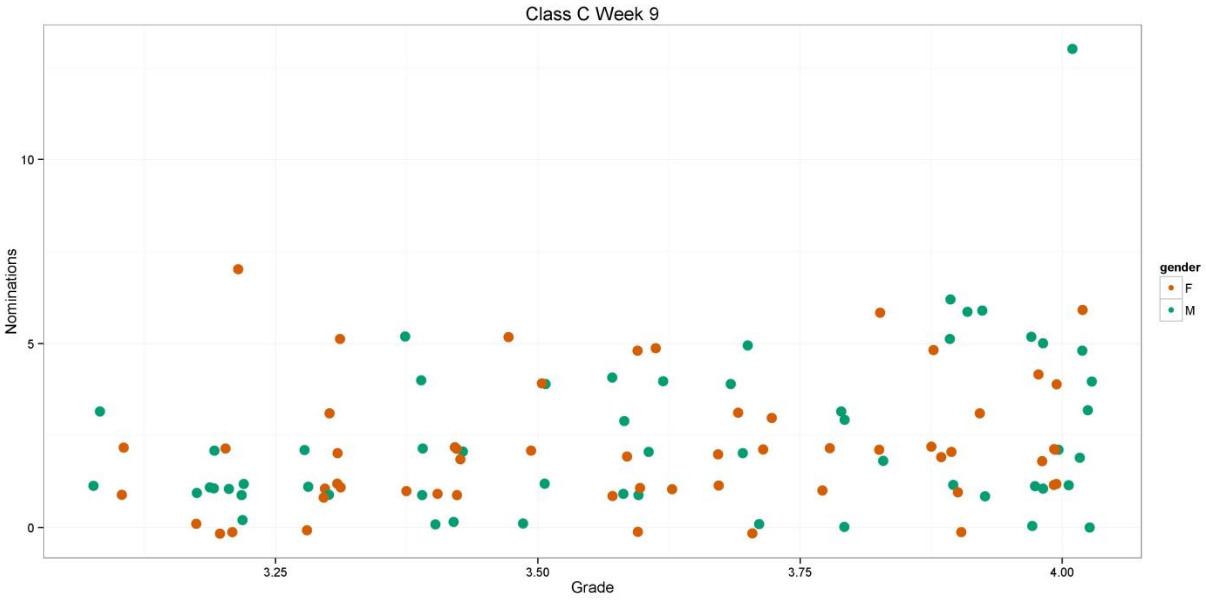


Figure 2.9: Outspoken students who scored in the top 10% of class and nominations earned at the last survey collection in class C

Even though outspoken females with extremely high scores exist, they fail to reach the same "celebrity" status as their male counterparts.

Addendum to Chapter 2

The framing and discussion of the previous chapter focuses on the educational implications for the large and consistent inequity found between the nomination patterns of male and female students. By illuminating a hidden and important bias, the implications of this finding to equity in higher education are made clear and are important. Fully interrupting this bias requires a thorough understanding of why it exists in the first place. Given the complex nature of human psychology, social interactions, and the culture of college classrooms, many fields can offer insight into the proximate and ultimate causes of the observed bias. Here, I take an evolutionary approach to understanding these biases, and reframe these findings within Biocultural Anthropology.

Across most subsistence human societies, males exhibit more political influence, usually through coalitional support (Kitayama, et al. 2016; Low 1992). While it is not clear whether this patriarchal dominance is a human universal, the widespread existence of this power inequity warrants critical examination (Smuts 1995; Strathern, et al. 1981). A more holistic understanding of this power dynamic can be achieved by considering possible evolutionary pressures that have selected, biologically or culturally, for male status seeking. With a more complete understanding of gendered behaviors, we can both identify new as well as support current policies and strategies to eliminate classroom biases.

In this addendum, I argue that the critical importance of status for male reproductive success, both through signaling and coalitional support, has selected for gendered signaling dynamics whereby we may expect the observed male-male bias. I argue that the exhibition of these male biases may be galvanized in culturally or perceptively “male-typical” domains, including science classrooms. After re-framing the current study within evolutionary theory, I briefly discuss several strategies to interrupt this bias which are supported by this evolutionary viewpoint.

Distal Causes of Gendered Behaviors:

Looking at issues of gender inequity within an evolutionary light can appear dangerous on the surface. Invoking biological explanations for different behaviors between males and females as it pertains to status can be shallowly misunderstood as not only asserting inevitability to inequities favoring males, but morally justifying them. However, a careful and complex consideration of deep seeded roots of gendered behaviors and status inequality can help bring to light potentially effective interventions and changes to help achieve social justice. Further, while this discussion focuses on a more biological perspective to the roots of inequity, it does so without claiming any biological determinism to gendered behaviors, and recognizes the complex nature of behavior, biological sex, and gender identity. Before discussing this bias through an evolutionary lens, I first consider perspectives on the distal causes of male and female behavioral norms. That is, how can biological, cultural, and social circumstances contribute to behaviors we associate as being more male, or more female?

Approaches to understanding differences between male and females range from essentialist, where differences are innate and hardwired to biological sex, to a social constructionist approach, where gender is viewed as being socially constructed and associated behaviors are understood as independent of biological sex, and are instead something we learn and prescribe to in society(Bohan 1993). A blended approach between these two extremes represents a more realistic, yet complex approach. The “biosocial model” is one example of an approach which blends biological essentialism with social construction of gender. This model ascribes the origin of human patriarchy to differences in reproductive strategies between males and females and a sex based division of labor, which has selected for greater stature and strength of males (Ingram, et al. 2009; Wood and Eagly 2002). This approach necessitates evolutionary theory for sex differences in morphology and labor roles. However, this model also recognizes the plasticity of male and female behaviors. A biosocial model takes into account that cultural constructs regarding normative gendered behavior become emergent based on these more essentialist sex differences. This approach is able to incorporate important evolutionary pressures acting on males and females, while accounting for the plasticity seen across and within society regarding behaviors. For instance, females gaining access to contraception and divisions of labor breaking down in industrial societies has come with updated definitions of acceptable masculine and feminine behaviors (Wood and Eagly 2002). Germane to the current discussion, the cultural context, and gender norms themselves, may emerge through a cultural evolutionary process – making biocultural anthropology a useful approach even if one assumes a great importance of social construction to gendered differences.

This re-contextualization of the current findings focuses on potential evolutionary roots of gendered behavior, incorporating the importance of socialized gender norms. I draw on predictions from evolutionary theory to understand why male behaviors may differ from females', which may ultimately lead to perception biases in STEM classrooms. I assume that culturally prescribed gender norms which influence behavior have been shaped through cultural evolutionary processes that likely conform to evolutionary pressures. In this way, I remain agnostic as to the relative importance of nature versus nurture on gendered behaviors, and instead assume that both are important enough to warrant consideration on the current issues.

Male Status Seeking:

Status inequality between men and women in Western societies is particularly salient. Only 4.2% of the CEO's of S&P 500 companies are female (Catalyst 2016), the United States has never had a female President or Vice President, and women earn less money than men (Marianne 2011). In hunter gatherer societies, even though material wealth is sparse and egalitarian food sharing is a necessary buffer against starvation (Kaplan, et al. 2005), the political influence of men is nearly always greater than that of women or children (Low 1992).

Why are men, and not women, more likely to assume higher positions in these systematic inequalities? An evolutionary account of gender status differential starts with different reproductive benefits that males and females gain from having higher status. Males and females have different avenues to achieving reproductive success (RS), rooted in respective physiological constraints to reproductive investment. Male RS is constrained by access to mates, while female RS is constrained by access to resources (Trivers 1972), While human males are exceptional in the amount of parental care offered compared to other mammals, they are nonetheless expected to seek more mating opportunities with less importance placed on mate quality (Geary 2000) while females are expected to be more selective and choose high quality mates (Trivers 1972). Because males have a lot to gain by obtaining extra mating opportunities, the variance of male RS ends up being much higher than female RS (Betzig 2012).

Considerable work predicts and empirically tests differences in mating strategies between males and females. Cross culturally, women value social status of their partner more than men, while men place more value on physical attractiveness than women (Buss and Schmitt 1993). Because females seek mates with resources, then a good male reproductive strategy would be to obtain and signal status, as status has associations with access to resources. With high female selectivity and skewed RS in males, male competition for status faces strong sexual selective pressures. Indeed, while status is important to females (Hess and Hagen 2006, Cashdan, 1996 #350), evidence shows that men tend to be more competitive and seek status at higher levels than females (Cashdan 1998).

If evolution selects for status seeking in males, we should expect high status males in non-industrialized societies to sire more children. Indeed, empirical work shows that high status men are more reproductively successful than low status men in these populations (Hopcroft 2006). Among the Tsimane forager-horticulturalists in Amazonian Bolivia, dominance and prestige of men are associated with a variety of outcomes, including (but not limited to) earlier marriage ages, higher survival rates of offspring, and more coalitional support (Von Rueden, et al. 2008). Other anthropological work with the Yomut Turkmen has shown that male wealth is associated with number of offspring sired, adjusted for age (Irons 1979). This also holds true for non-human primates, where the association between high status and reproductive success is stronger for males than females (Cowlishaw and Dunbar 1991).

Two primary avenues to achieve high status for humans have been acknowledged: dominance and prestige (Henrich and Gil-White 2001). Dominance is gained through physical exertion or winning dyadic competitions, while prestige is differentiated as earned recognition and deference based on skills, success, or knowledge (Henrich and Gil-White 2001). They argue that freely granting prestige to others has evolved as a strategy to gain access to highly skilled and knowledgeable others. Individuals may offer coalitional support, share food and resources, or confer other benefits to an alter they wish to learn from. By providing deference, they increase the prestige of the alter, and may gain access to useful cultural knowledge. Prestige is a form of status unique to humans, and may drive males to signal or "show off." For example, it has been documented that good hunters are able to garner high status by sharing food, (Gurven and Hill 2009; Smith 2004) a way to gain status without inflicting harm on others. In the case of a college classroom, offering answers in front of the classroom can be perceived as a means to show off and gain status as a knowledgeable peer, who is likely to advance in a prestigious profession and career. In this study and previous work in similar classrooms (Eddy, et al. 2014), it was shown that males are more likely to speak in class than females.

The current study accounts for this inequitable class participation through the binary "outspoken" label. However, this variable fails to capture the tenor of responses in the classroom and the variation in volume of responses. While the outspoken variable does attenuate the male bias, suggesting male "showing-off" is likely a part of the pathway towards male specific prestige, a more confident response pattern or a higher volume of correct answers being offered by some males may account for some of the observed male bias. However, showing off is not the entire story, since it is only males and not females who over-praise their male peers. A complete understanding of the current findings must account for this male specific behavior.

Perceiving other males:

Human psychology and intelligence has likely been shaped by evolutionary pressures to navigate our complex social world, including social hierarchies (Herrmann, et al. 2007). Males and females benefit from knowing and navigating their own place within dyadic and group relationships, as well as the hierarchies surrounding them. In other words, paying attention to others is important. In the case of the classroom perception bias, it appears that students pay attention to their peers, with males particularly sensitive to recognizing their male peers, while females remain neutral based on gender.

This finding is somewhat confounding - females consistently list status as an important attribute in a mate (Buss and Schmitt 1993; Sadalla, et al. 1987), and may be expected to recognize any "showing off" of their male peers. Why are males and not females doling out excess recognition to males? I consider three possible explanations which suggest selection for a gendered psychology leading to male specific recognition of peer prestige. First, high costs to male-male competition may have selected for males to be especially attuned to the status of their competitors. Second, while males compete within groups, they tend to be more coalitional than females. Lastly, the content of the classroom may impact the directions of the gender bias.

Male human and non-human primates participate in more frequent and intense physical confrontations than females, where losers face serious injury, death, and loss of access to resources and mates (Sell, et al. 2009). For this reason, avoiding conflicts where there is a high likelihood of defeat is an important strategy to avoid the potential costs of losing. Evidence suggests that human cognition is accurate at identifying dominance, but with some variance between individuals (Watkins, et al. 2010). College aged males, who are at the end of adolescence, may behave as if in competition for status within a college classroom setting, while females do not come in with this same competitive motive. In this case, a male psychology for contest avoidance may drive greater attentional attunement between male students. This increased within sex attention could lead to greater biases in nominating male peers.

While males are expected to partake in intra-sexual competition, intergroup competition appears to be a major driving force in human evolution (Bowles 2009; Manson, et al. 1991). An importance of group competition makes coalition forming a potentially critical male behavior. Evidence for the importance of coalitions to males comes from studies on early childhood friendship differences between males and females. Research performed in Western contexts shows that boys are more likely to form large same sex play groups compared to girls (Rose and Rudolph 2006). Interestingly, the type of play within these larger male groups include a considerable amount of rough and tumble play (Rose and Rudolph 2006), which results in stronger hierarchies than seen in female groups (Omark, et al. 1975; Savin-Williams 1979). While both boys and girls frequently interact within dyadic friendships, females spend longer durations in these one-on-one interactions (Benenson, et al.

1997). Thus, the social landscape for males and females seem to reflect an increased importance of coalition forming in males, along with hierarchy within coalition groups.

The content and cultural context of the classroom may also impact gender biases in perception. Due to historical circumstances, science carries a male stereotype (Nosek, et al. 2002b; Nosek, et al. 2009). A field that starts out as male dominant creates greater barriers to females than to males, and quickly creates a perception that masculinity is important to this profession. Research in Fijian villages showed that individuals held sex based biases according to the domain of knowledge when listing others from whom they would prefer to learn (Henrich and Broesch 2011). Males were more likely to be listed for fishing and yam growing, while females were more likely to be listed for learning about growing medicinal plants, matching the gendered division of labor within the population. We may expect a similar “male-science” bias in our similar name generating task. Henrich and Broesch (2011) don’t find that males are more likely to list other males, or that females are more likely to list other females, though, which they attribute to a possible desire to learn unique information held by dissimilar others. The inconsistency between the studies in this regard likely stems from the difference in the survey question; In Fiji, individuals were questioned about to whom they would go for advice, while the current study asked only to identify knowledgeable peers. The difference between desired social interaction and nomination of knowledgeable others may be responsible for the unique results. It would be important to further study how content matter alters sex based biases in knowledge perception, especially since evidence shows that the level at which fields are associated with maleness scales to the gender ratios within that field (Carli, et al. 2016).

Further support of the current findings:

Prescribed gender norms may also be playing into the classroom and nominations on these surveys. Research on how individuals ascribe status, which takes a social constructivist approach to gendered behavior, finds that male undergraduates, but not females, are more likely to use unreliable information about status (Ridgeway, et al. 2009). Specifically, they had males and females interact in subsequent encounters with individuals who had experimentally manipulated status – either low or high. Individuals were told the status of the first interaction, but not the status of subsequent partners. Compared to females, males whose first encounter was with a high status individual were more likely to treat future un-identified partners as if they were high status, even if they were actually low status. . The authors attribute this to gender typical norm prescriptions; females are expected to act communal and nice (Eagly and Karau 2002), making the act of falsely attributing status to others incongruous with gender roles. Previous work has shown that women can face negative backlashes by males and females for not behaving in accordance with their prescribed identity, specifically in relation to asserting superiority (Rudman and Glick 2001). The risk of acting atypically female may drive females to be more conservative in their behavior when nominating peers compared to males.

Other work has examined whether males and females recall networks with the same fidelity. When given both fictive kin-based networks, as well as non-kin network relations, undergraduate females performed better at a network memory task (Brashears, et al. 2016), and likely offer a more accurate recall of network relations. While the current study did not ask students to remember a pre-determined network structure, nor to include any information about how listed individuals are connected to each other, such as was measured in Brashears, et al. (2016), females do appear to report more accurately which students were performing well in the course without inflating individuals based on attributes unrelated to knowledge.

Interrupting bias:

If we wish to interrupt biases in education, understanding the biological, cultural, and social roots of these inequities are important. There are reasons to believe that showing off knowledge, while being more attuned to other male peers, may have been important male behaviors throughout evolutionary history, perhaps even in modern contexts. The cultural context, where male and female students are enculturated into gender typical norms and are operating in a field stereotyped as "male", are also likely galvanizing the current male bias. From the current discussion, several hypotheses for effective strategies to eliminate the current bias emerge. While using evolutionary logic is not a pre-requisite to identifying these, it does lend further support for their examination.

1) Work to break the science-as-male stereotype:

   If perception is tainted by gender associations with different domains of knowledge, then breaking the stereotype that science is a male field can help erase perception biases. This movement is certainly underway, with major toy manufacturers, such as LEGO® introducing female scientist mini-figures to their historically male line.

2) Be aware of the power held by the highest status individual in a classroom community - the instructor:

   Within hierarchies, the highest status individual holds formal and informal power to award status to others. The structure of a classroom is no different, and automatically grants an alpha role to the instructor. Thus, the actions and perceptions of the instructor are likely very salient for any prestige granted within the class. Instructors who display favor or disfavor to students may greatly impact how those students are perceived by their peers, and can be a powerful way to indoctrinate implicit biases in students. Instructors must be particularly mindful of their actions.

3) Structure matters:

   A recent paper lays out ways to increase equity in classrooms through classroom structure (Tanner 2013). The suggestions within this article are not derived from evolutionary thought, but many of them are ways to interrupt emergent inequitable interactions that occur in classrooms. These suggestions include systematically limiting students' ability to "show off" in class by making sure voices are heard equitably in the class, avoiding overpraising students, and building a cooperative classroom community, among others. Avoiding a competitive environment is particularly salient here. Forming a classroom based on cooperation where the goal is to help one another, as opposed to compete with one another, gets away from rewarding generally male-normative behaviors of status seeking in classrooms, while aiding learning opportunities for all students.

Conclusion

Finding that males over-rate their male peers in Biology classrooms suggests large impacts for equity in STEM fields. This finding can be understood through many lenses, including a biocultural one. Stemming initially from reproductive constraints and mating strategies, social behaviors can be expected to differ between males and females. While differences in showing off are visible and accounted for in the current analyses, any perceptive differences between genders are not. I argued here that these differences can be understood as having evolutionary roots, which may be either innately or culturally grounded. Given evidence in the literature, I briefly examined some suggestions for eliminating the observed bias, both at the societal and classroom level.

Chapter 3: Examining the use and utility of model-based biases in large lecture classrooms

Abstract

Formal evolutionary models predict that humans should exhibit biases in whom they choose to learn from. Two classes of these social learning biases include *payoff-based* biases, where individuals are expected to preferentially learn from successful or prestigious others, and *homophily-based* biases, where individuals are expected to preferentially learn from similar others. Despite considerable theoretical work on learning biases and outcomes from using them when in learning situations, empirical work has been lacking. By using longitudinal social network analysis in large lecture classrooms, I test whether these biases are used by students when forming new study partnerships. I further test whether studying with peers perceived as strong in the course material leads to improved exam performance. I find evidence that students preferentially form study relationships with peers previously identified as strong in the material, and that students assort into same ethnicity study partnerships. I find no evidence that studying with peers identified as strong in the material is associated with higher exam scores, but do find evidence that students do better on exams when reporting many study partners compared to studying alone. This study provides empirical evidence for the use of these learning biases in a novel research environment.

Introduction

Human populations have colonized almost every available ecosystem on Earth – a dizzying array of landscapes ranging from dense tropical forests to arctic tundra (Henn, et al. 2012). This incredible feat occurred before the advent of agriculture, and was made possible through the accumulation of complex cultural knowledge that enabled populations to thrive in such disparate ecosystems (Richerson and Boyd 2008). The high degree of cultural ingenuity and high fidelity in transmitting cultural are also unique to the human species (Tomasello, et al. 1993). Thus, understanding how humans became such a dominant species on earth requires understanding the evolution of our cultural capacities.

The framework of gene culture coevolution (GCC) takes to heart this fundamental importance of culture to human success. A major piece of this framework generates hypotheses about how natural selection may have led to a human psychology which preferentially finds and learns adaptive culture. This transmission of culture has been referred to as social learning, and is an important part of the definition of culture used in the GCC framework: *socially learned information capable of affecting the behavior of those who contain it* (Richerson and Boyd 2008). Because behavioral choices affect reproductive fitness, individuals should be prudent in how they arrive at the information which underlies their behavior. Using a suite of mathematical models, researchers in CGC have theorized a set of heuristics that guide this prudence. These heuristics are referred to as *Social learning strategies* (Boyd and Richerson 1988; Rendell, et al. 2010; Rendell, et al. 2011).

Social learning strategies can be categorized as content or context dependent. Content dependent strategies are based on something about the actual cultural information being learned. This may include a preference for culture that is inherently attractive or that evokes strong emotions (Bangerter and Heath 2004; Heath, et al. 2001). Context dependent strategies, on the other hand, are based on the social context of the information. The context can be further organized into state-based, frequency-dependent, and model-based strategies (Rendell, et al. 2011). State-based strategies refer to the state of the learner herself, and may, for example, include the level of uncertainty she has in her own cultural knowledge or other dependencies based on her current state (Galef 2009). Frequency-dependent strategies focus on the proportion of behavioral practitioners in a population or the frequency with which a behavior is used by a single individual being observed (Boyd and Richerson 1988; Pike and Laland 2010).

The last class of context biases, model-based strategies, is the focus of this chapter. Model based strategies are based on the characteristics of *who* a learner seeks to learn from (Boyd and Richerson 1995; Enquist, et al. 2007; Laland 2004). In this literature, the person being learned from is somewhat confusingly referred to as a "cultural model". Some cultural models may be easier or cheaper to access for potential learners than others. At the same time, cultural models may vary in the quality and relevance of their cultural information. Unfortunately for potential learners, the net benefit of learning from any potential model can often be inconspicuous. Evolutionary models thus predict that humans should have evolved cognitive biases for preferentially learning from cultural models who display traits that correlate with high quality cultural information (Boyd and Richerson 1988). Some examples of model based biases include learning from models who are successful or prestigious (Henrich and Gil-White 2001), older (Dugatkin and Godin 1993), or related to the learner (Henrich and Henrich 2010). In each of these cases, the learner is able to use an observable or known trait to minimize learning costs while maximizing the benefits associated with the information learned.

There are other important considerations to human cultural learning. First, not all learning is social. Learning through trial and error, often referred to as learning individually or asocially, is a critical part of cultural evolution, as it is thought to be the main source of new cultural variation (Boyd and Richerson 1988; Richerson and Boyd 2008). Second, social learning strategies are in and of themselves behaviors which can be culturally learned (Acerbi, et al. 2009; Ghirlanda, et al. 2006; Mesoudi, et al. 2015a). That is, learning behaviors may be strictly under genetic control, strictly under cultural control (we learn how to learn), or most likely, outcomes of both genetic and cultural influences. Third, different social learning strategies are not mutually exclusive from one another. The *when, how,* and *from whom* to learn can be path dependent and require complex cognition to properly navigate (Laland 2004). GCC theorizes that the cognition and/or ability to navigate towards the best learning outcome has been selected for, genetically or culturally, throughout human evolutionary history.

Payoff-based biases:

One of the earliest classes of bias considered by evolutionary theorists within model based biases are *payoff-based* biases (Boyd and Richerson 1988). These biases are predicated on locating and copying a more successful model than yourself, and can include various heuristics to decide on whether or not an alter is successful enough to warrant being copied (Schlag 1998). As an example, learning from a fisher who averages a yield of six fish per outing represents a good strategy if your current fishing method results in an average of one fish. However, this learning choice becomes more difficult if your method yields five fish. In order to use this bias, an individual needs to be able to accurately assess the successes of other fishers, which is not always a trivial task in learning situations. Perhaps a less costly heuristic is to preferentially learn from an individual who others have discerned as a good cultural model. If this individual is a strong enough model to warrant high access costs, one might expect others to grant prestige to this person and provide favors and gifts in return for learning opportunities. This logic is the basis for the evolution of a *prestige bias* (Henrich and Gil-White 2001). As an example, learning from a fisher who has six individuals trying to learn their methods instead of learning from the fisher who only has one represents using a prestige bias. Both prestige and success biases aim to discriminate something about the quality of the individual being learned from, and are thus payoff-based.

Assortative social learning - homophily bias:

Extensive evidence shows that humans tend to affiliate with similar others, a phenomenon referred to as homophily (Lazarsfeld and Merton 1954; McPherson, et al. 2001; Rogers 2010). These homophilic interactions are seen starting at a very young age (Fawcett and Markson 2010; Kinzler, et al. 2007), appear across human populations (Cohen and Haun 2013; Kinzler, et al. 2012), and are seen less frequently in non-human primates (Haun and Over 2015). Further, these types of interactions create and stabilize cultural diversity between distinct groups (Centola 2011). Despite the ubiquity and clear importance of human social homophily, an evolutionary basis for assortative social learning has been studied far less than other model based biases within a cultural evolution context. Recent work suggests its importance in situations where payoff-based strategies are too difficult to use, and when there is no singular “right” behavior for all individuals (Katsnelson, et al. 2014). In these instances, it may behoove individuals to copy similar others. For example, in cultures with a strong division of labor between genders, females would presumably find less reproductive success by copying male behaviors and vice versa. Further, the success of dyadic and group cooperative interactions often rely on shared norm signaling (Richerson 2003; Sosis 2000). Assortative social learning helps spread shared norms, and may ultimately have selected for heuristics based on homophily when learning (Haun and Over 2015). This can be the case if adopting shared norms is more likely when learning from similar others, and displaying shared norms results in more beneficial cooperative interactions with others.

Empirical tests of learning strategies:

To date, the field of GCC has primarily been propelled by formal evolutionary models to understand how the dual inheritance of genes and culture has driven human evolution, including human learning biases. Empirical research about how predicted learning biases transpire in real life has lagged behind theory, and often occurs in laboratory settings (Chudek, et al. 2012; Efferson, et al. 2008; Mesoudi 2011b; Molleman, et al. 2014; Morgan, et al. 2012). Work in Fijian fishing villages provides some of the best non-laboratory evidence of success and prestige-based biases to date. Henrich and Henrich (2010) found that food taboos are preferentially learned from older prestigious women, evidence of payoff based biases. Interestingly, this also indicated domain specificity in who to learn from. In a study across three villages in Fiji, participants were asked to list who they would go to for advice for several culturally relevant skills (Henrich and Broesch 2011). By examining the emergent network, the researchers were able to test whether individuals in this ecology preferentially learned from prestigious and successful individuals. The results showed that individuals would prefer to learn from knowledgeable and prestigious others. Further, they found that their desired learning networks were fairly centralized, suggesting that learning about their culturally relevant tasks does not simply happen within a nuclear family, but individuals tend to seek out the same knowledgeable individuals.

Social learning in classrooms:

The current study expands upon the work of Henrich and Broesch (2011) within the context of an undergraduate classroom by longitudinally tracking the formation of study partnerships, and the impacts that different partnerships may have on student success. Undergraduate classes aim to transmit cultural information to students, who are then assessed on their mastery of the material. Students gain informational capital through self-formed study networks, where having learning relationships may impact student performance in a positive manner. This may occur if peers 1) understand a greater quantity of material, or 2) understand material at a higher level or with more clarity, or 3) use more suitable methods of studying and preparing for exams, or 4) know of additional useful learning resources, and if these peers transmit any of these forms of informational capital to study partners.

To date, the formation of study networks and their impacts on students in undergraduate classrooms are understudied (Grunspan, et al. 2014). Concurrently, empirical tests of learning strategies in natural settings are lacking. The current study tests whether undergraduate social behaviors throughout a 10 week course conform to model-based context biases. Specifically, this work tests whether students preferentially use payoff-based and homophily biases, and whether or not the use of a payoff bias impacts the performance of students in the course.

Hypotheses:

Hypothesis 1a) *Students will be more likely to form learning relationships with peers they previously identified as knowledgeable (payoff biased); 1b) Students will be more likely to break study relationships with peers they do not identify as knowledgeable (payoff biased).*

Study ties between students should be more likely to form between students who had previously identified one another as strong in the course material. Because students gain information about their peers as a class goes on, updated information on the content knowledge of peers will aid the decision of whether to persist studying with the same individuals. Following this logic, students should be more likely to discontinue studying with each other if neither student identifies the other as knowledgeable.

Hypothesis 2a) *Students will be more likely to study with peers that are similar to them; 2b) Students will be more likely to break study relationships with peers that are different from them.*

Study ties between students should be more likely to form if both students are of the same ethnicity and/or gender. Further, as students find new potential partners, non-homophilous study partnerships may be more likely to break compared to homophilous partnerships over time. The evolutionary roots of this bias are hypothesized to relate to the importance of adopting identity specific norms. In the case of a classroom, though, all students are aiming for the same correct answers. Thus, homophily by gender or ethnicity should provide no informational advantage independent of other traits. However, this learning bias is likely to persist into new environments where it may not necessarily be adaptive for individuals. In this sense, this bias may represent disequilibrium between a previously adaptive evolved learning bias and a new environment where the adaptive properties may no longer exist.

Hypothesis 3) *Students will perform better on exams when they use payoff-based biases.*

Payoff-based biases are predicted to emerge because they optimize gains from cultural learning processes. In the case of social study behaviors, students should perform better on exams when their study group includes a peer who they identified as strong in the material compared to exam s where this isn't the case.

Methods

Setting

Data collection took place across three different iterations of the same large-enrollment undergraduate Biology course (classes A, B, and C). This was the second in a series of three introductory classes; many students came into the course knowing some of their peers. Lectures took place four times a week for 50 minutes, and smaller recitation lab sections of a maximum of 24 students met once per week for two and a half hours. Each class contained four exams, each separated by 2-3 weeks of class material. Classes A and B were each taught by one instructor, while class C had three total instructors teach throughout the course. Two of the three instructors in class C were the instructors from classes A and B.

Data

Students were asked to list peers in the class who they perceived were strong with class material (prestige), and peers with whom they studied for each exam. The collection of these network data occurred longitudinally throughout the quarter, centering around 4 exams in each class. Class C had the most collection points (all four study networks), while classes A and B only had three study networks collected. Further, networks regarding who students felt were strong in the class material were only collected twice in class A. For this reason, class C allows more repetition in analyses compared to B or A. Student gender, ethnicity, and college GPA were collected from the Office of Educational Assessment.

Logistic models of formation and dissolution

Estimating the likelihood of study partnership formation and persistence over time is of central interest to hypotheses 1 and 2. One way to model these data is through separate logistic regression models, one estimating the likelihood of a new tie between two students forming, here referred to as a “formation” model, and a separate model for currently existing ties persisting to the next exam, here called a “dissolution model. Repeating these analyses separately over each time panel affords a fine grained causal analysis than could be afforded by pooling across all time panels. This allows any potential temporal differences in the use of context biases to be considered (i.e. do students use these biases early in the class, but not later?). Further, by separating analyses into newly formed and persisting ties this approach captures whether patterns in observed networks are driven primarily by forces of formation or dissolution.

The models were fit with the temporal exponential random graph models (TERGM) package as part of the statnet suite of packages in *R (Handcock, et al. 2008; Krivitsky and Handcock 2015)*. STERGMs are an extension of exponential random graph models which capture longitudinal dynamics of network formation (Krivitsky and Goodreau 2015; Krivitsky and Handcock 2014). When these models lack network level parameters, such as in the current analysis, they conform to simple logistic regression models. Thus, the use of STERGM models makes these parametrically complex models easy to compute while remaining analytically identical to logistic regression models.

To test hypotheses 1 and 2, the models were parameterized as follows:

Formation model:

1) **Payoff bias: Nomination in perception network previous time step**: If students preferentially studied with successful or prestigious peers, their previous identification of a peer as knowledgeable should predict a future study partnership. This tendency may be even stronger if two individuals mutually identified each other as knowledgeable. This variable takes the structure of a matrix with the same dimensions as the study networks. Each cell of the matrix takes values 0, 1, or 2 for each possible tie that could exist. These values are determined by the state of the perception network at the beginning of the time

panel of interest. 0 represents no nomination from student *i* to student *j* or student *j* to *i*, a 1 if student *i* nominated student *j* or if student *j* nominated student *i*, and a 2 if there were mutual nominations between students *i* and *j*.

2) **Homophily by ethnicity and gender:** If students preferentially study with similar peers, then student *i* and student *j* should be more likely to initiate a studying relationship between the start and end of the time panel if they are of the same ethnicity or gender.

3) **Homophily by lab section:** The class structure included smaller lab sections, where students often make friends and form study groups. To control for this structure and the importance of lab sections in brokering relationships, homophily by lab section was included as a control variable.

Dissolution model

1) **Payoff bias: Nomination in perception network at previous time step*:*** If students preferred to study with successful or prestigious peers, then study partnerships may be less likely to have persisted if neither student nominated the other in the perception network despite studying together. Likewise, they would have been most likely to continue their study relationship during the following exam if both students nominated each other. This variable is thus parameterized as described in the formation model above.

2) **Homophily by ethnicity and gender:** If students preferentially studied with similar peers, then a student *i* and student *j* who shared an edge at $t = 1$ should be less likely to have had their edge persist until $t = 2$ if they are of different ethnicity or gender than if they are the same.

3) **Homophily by lab section:** Similar to the formation model, lab section was used as a control for the social structure of the course.

All analyses were run with the *statnet* package (Handcock, et al. 2008; Krivitsky and Handcock 2015) in *R (R Core Team 2013).*

Data selection:

The number of students in classes A, B, and C were 196,759, and 760, respectively. Student attribute data was missing from 0 students in class A, 33 students in class B and 17 students in class C who either dropped the course or were unmatchable to registrar data. Because STERGM models require complete nodal attribute data, subsets without these 33 and 17 students were used in models for classes B and C.

Mixed effects models:

Mixed effects models were used to test the value added for students using payoff-based biases when studying for exams. The data include multiple exam scores from the same students who varied in their study groups and perception of peers across exams. To control for shared variance within each student and class, student ID and class (B or C) were modeled as second and third level random effects, respectively. This hierarchical model has exam scores within students, who are in either Class B or C. Grade in the previous course, exam number, and a categorical variable for number of study partners (0, 1, 2-3, or >3) were modeled as fixed effects, and control for student aptitude, exam difficulty, and study group size, respectively. To further control for any differences on exam difficulty, Z-scores on exams were used as the dependent variable. The independent measure of interest is whether a student studied with a peer who they perceived as strong with the course material or not. This was modeled as a fixed binary variable indicating whether a student studied with at least one peer who they nominated as prestigious (at any point leading up to the exam of interest), or not. Students who either always studied with a prestigious peer or never studied with a prestigious peer were removed from these analyses, as there is no way to assess the impact of this payoff bias in these students. Analyses were run using the *lme4* (Bates, et al. 2012; Bates, et al. 2007) package in *R* (R Core Team 2013).

Results

Description of study networks:

In all three classes, the mean number of study partnerships increased from the first exam until the last exam. The only exception was on the 2nd exam in class C, where there were fewer study partnerships than for any other exam in this course. While there was a net gain in number of study partnerships, it is important to note the large number of study partnerships that did not persist between exams (**Error! Reference source not found.**). Sociographs of each study network are shown in Figures 3.1, 3.2, and 3.3. These visual inspections of the structure of the network show some evidence of homophily by ethnicity. They also show a large number of isolated students outside of a somewhat dense network.

Biases in study networks:

To assess whether students used predicted biases in forming new study partnerships, we examine model results from separable temporal exponential random graph models (**Error! Reference source not found.**). These are interpreted in the same fashion as logistic regression. In the formation model for Class A (t_2 -> t_3), the edges coefficient of -5.905 is the log odds of a study partnership forming between two students who are different by ethnicity and gender, and neither of whom previously said the other was knowledgeable with class material during preparation for the second exam; the estimated probability of a study partnership forming between two students with these attributes would be $exp(-5.905)/(1 + exp(-5.905))$, or 0.003. If one of these students indicated that the other student was knowledgeable at t_2, the odds that they studied together at t_3 would then be estimated as $exp(-5.905+ 2.510/(1 + exp(-5.905 + 2.510))$ or 0.032. If both of the students listed each other as knowledgeable, the probability increases to $exp(-5.905 + 2*2.510)/(1 + exp(-5.905 + 2*2.510))$, or 0.292. If these students also shared ethnicity, the probability would increase to $exp(-5.905 + 2*2.510+ 0.500)/(1 + exp(-5.905 + 2*2.510+ 0.500))$, or 0.405. Both payoff bias and homophily by ethnicity were consistently significant predictors for the formation of new study partnerships across every exam within all of these classes. Shared gender significantly predicted the formation of study partnerships in two out of the six models.

The coefficients in the dissolution models are interpreted differently. Here, the edges coefficient in Class A of 0.783 is the log odds that a study partnership persists (between two students of different ethnicity, gender, and who neither partner identified as knowledgeable), from the second exam to the third exam. The low payoff coefficient is consistently significant and negative. This indicates that the likelihood of a study partnership persisting from t_1 to t_2 was significantly less likely if neither student identified the other as knowledgeable at t_1.

Biases and exam performance:

Payoff-based biases are predicted to evolve because their use should optimize learning outcomes. Mixed effects models were used to test whether students' classroom performance was positively impacted by the use of payoff-based biases in the formation of study partnerships. These models utilized the repeated measures of exam scores within students, while accounting for the hierarchical structure of the data where students were either in class B or class C. Model results showed no significant effect of studying with a peer identified as strong with the course material on exam grades. That is, a student did not do significantly better or worse on exams when they studied with a peer identified as strong with the material compared to exams where they didn't. However, the number of peers a student studied with did seem to matter, where having more study partners was associated with higher exam scores for students (**Error! Reference source not found.**).

Discussion

CGC predicts that humans should rely on context biases when socially learning. These include payoff and homophily based biases. To date, these social learning heuristics have not been examined within the setting of an undergraduate classroom, where learning relationships are an important part of student experiences. By longitudinally following study partnerships in three iterations of an introductory biology course, along with how students perceive one another, this study finds that students preferentially form partnerships with peers they previously identified as strong with the course material and peers who share the same ethnicity as them. There is some evidence that students preferentially form new study partnerships with peers of the same gender, as well. Students are less likely to continue studying with a peer if there was no recognition of content knowledge between the students, but are equally likely to continue partnerships with students who share, and don't share, the same ethnicity or gender.

The formation of new study partnerships largely conform to predictions from CGC's context biases, adding additional empirical support to the current literature testing these biases. Payoff-biases not only predicted formation, but also whether relationships persisted from one exam to another. Once someone finds themselves with a peer they think is strong with the content, they are significantly more likely to keep studying with that peer. There is no strong indication that the importance of payoff biases changed in any coherent or consistent way over time from t_1 to t_4.

While previous empirical work has largely focused on laboratory studies (Mesoudi 2011b) or desired learning relationships (Henrich and Broesch 2011), this work manages to capture the formation and persistence of learning relationships over time in a population where these relationships can have tangible and important outcomes. In doing so, the networks formed include the constraint of access costs to models, an important theoretical constraint within payoff-based learning strategies (Henrich and Gil-White 2001). In this case, high aptitude students in the course may represent desirable study partners, but they are unable, and likely disinterested, in studying with every student who is interested in studying with them.

Some of the effect of students preferentially studying with peers previously nominated as knowledgeable may be driven by these peers existing within a small friend group. For instance, imagine that Pat knows five students in the course – Corey, Jamie, Glenn, Jean, and Jessie. Pat studies with Jessie, but identifies Corey and Jamie as strong with the material at time t_1. When predicting who Pat will study with at time t_2, Corey and Jamie already represent likely study partners because they are in their local network, and are also likely to be listed as strong in course material due to this same local effect. The effect size of payoff bias influencing study partner choices may be inflated due to the inability to capture these smaller local networks in the model. It is also worth considering that students are likely not choosing study partners based solely on a desire to maximize their learning. There are clearly other reasons why college students would prefer to spend time together that are not captured here, including student interest groups and belonging to Greek organizations, along with pursuit of romantic relationships.

Social learning strategies are predicted to emerge based on the assumption that they are adaptations for learning cultural information. In a classroom environment, if study partner choice is guided by adaptive cognition for social learning, then using payoff biases should be expected to result in higher exam scores. Students did not have any value added to exam scores when they studied with peers identified as strong in the course material compared to when they did not. The lack of a relationship between studying with a peer identified as strong in the material is somewhat surprising. However, exam scores are a notoriously noisy variable which can situationally vary by circumstances which are difficult to capture (Jargowsky and El Komi). Further, effect sizes in educational settings are often much smaller than other fields or in controlled laboratory studies (Lipsey, et al. 2012). Capturing possible impacts on student achievement may require either a more controlled setting, or considerably larger sample size.

While exam scores weren't higher for students when studying with a peer previously nominated as knowledgeable, there was an effect of simply studying with peers. When studying alone, students did significantly worse on exams than when they had studied with peers. This effect was particularly strong for students having greater than three study partners compared to having none. While working with peers almost certainly helps student success on exams, there is likely some correlation between studying with peers and effort; a student who studies alone on one exam, but with peers on another, is also likely putting in different amounts of effort between the exams. Thus, a student who reported having many study partners may also be indicating that their overall effort was higher. This makes it hard to say with certainty that studying with peers has a purely causal impact on exam scores in this case. More work would be needed to investigate any causal relationship between study relationships and learning gains, as well as what precise impact peers have on each other. For example, are students transferring classroom content, study techniques, or having positive effects by boosting peer confidence?

The importance of sharing ethnicity for the formation of study partnerships, indicating that students follow the well-studied and evolutionarily predicted homophily bias. Similar to payoff biased relationships, there is no clear trend that can be understood regarding the effect size of homophily in partnership formation. Interestingly, homophily was not important for the persistence of study relationships. This suggests that student identity is mostly important as a social barrier in these classrooms. It is possible that students realize a value in diverse study groups once they are in such a group. Research in higher educational settings shows that working with diverse peers helps lead to solutions of complex problems, and can have lasting positive impacts on student outcomes (Denson and Chang 2009; Denson and Zhang 2010; Gurin, et al. 2002; Laird 2005). The importance of diversity and inclusion in higher education makes the propensity for same-ethnicity partnerships founds in these classrooms worrisome.

A tendency for homophily by ethnicity suggests that students from minority groups may find it harder to locate study partners within the class, which my ultimately impact their social capital. Previous work has shown that existing inequalities in social capital, combined with homophilic interactions, can negatively impact attainment of social capital across generations (Cross and Lin 2008). A similar operational force may be at work within science, technology, engineering, and mathematical fields, where existing achievement gaps (Barton and Coley 2009) operating alongside homophily can compound these issues. Thus, understanding homophily in this setting, and the possible evolutionary underpinnings of this social tendency, represents an important research focus that may help to close achievement gaps.

Conclusion

Undergraduate students followed context biases that emerge from CGC models, lending empirical support for context biases guiding social learning behaviors. Specifically, students preferentially form study partnerships with peers they had previously identified as strong with the course material, as well as peers who were of the same ethnicity. Further, study relationships between students who recognize each other as strong in the material were more likely to persist compared to relationships where this recognition was not granted. While theory predicts that these sorts of learning biases may emerge because they lead to increased fitness, there was no evidence that students actually performed better on exams for studying with prestigious peers. Instead, the results showed evidence that simply studying with peers was an important behavior that correlated with better exam performance.

This research also provides an overview of what informal study networks look like in this particular large lecture classroom, and how they relate to student identities and perception. It is apparent that students have a propensity to study with peers they think highly of, as previous identification of a peer as strong in the course material predicts a future study partnership. It is also clear that students in these classes assort based on ethnicity. The finding that students performed better on exams when they were studying with peers compared to studying alone corroborates with previous work (Tinto 1997) that the social environment in college classrooms is an important arena for educational research. The work presented here further suggests that an evolutionary perspective to these social behaviors represents one way to understand these behaviors and their impacts.

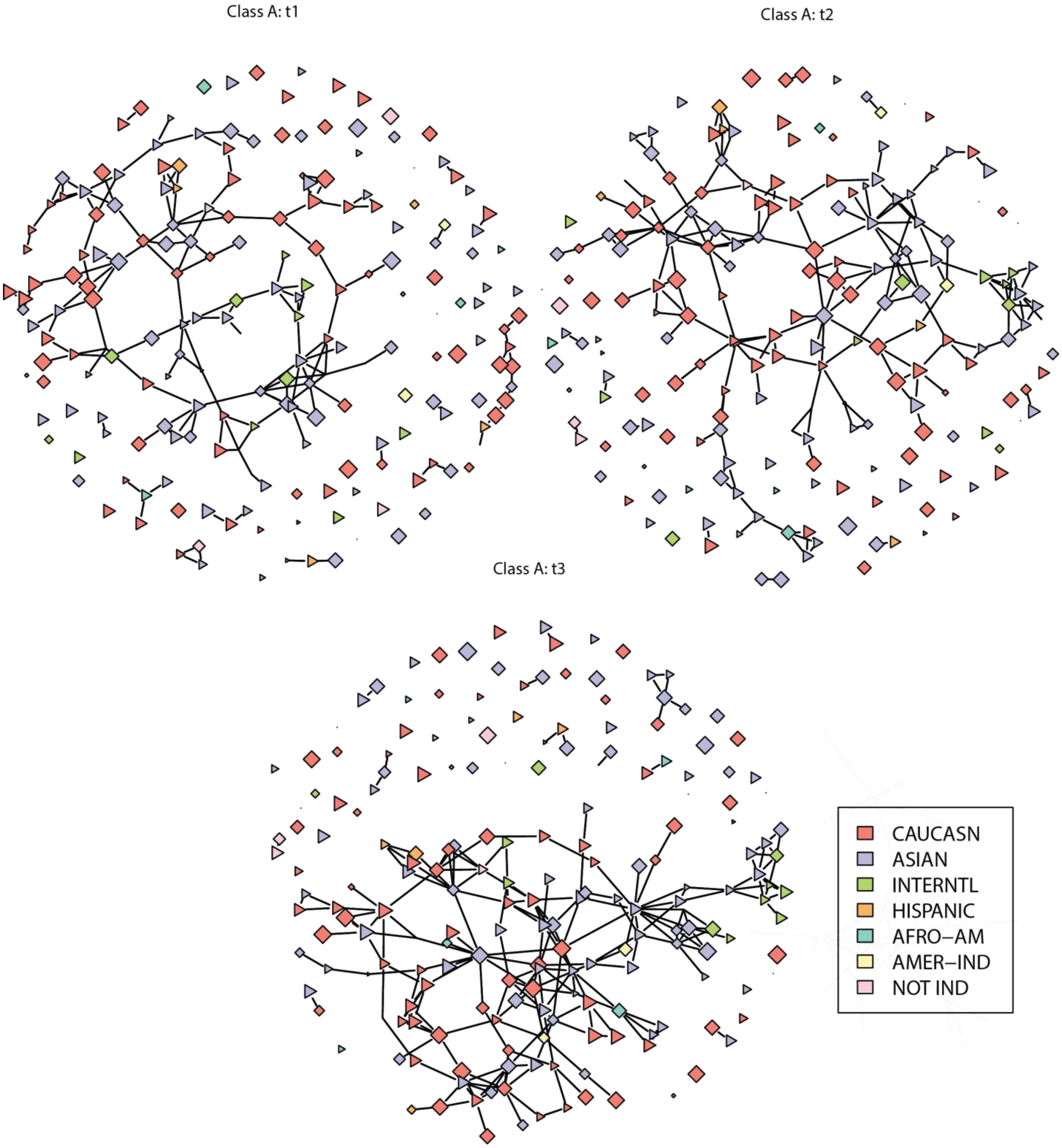


Figure 3.1: Class A Study Networks.

Class A study networks at times *t=1, t=2,* and *t=3*. Sizes of nodes correlates with grade in course, so larger nodes are students with higher grades. Triangles are male students, while squares are female students. The color of the node denotes student ethnicity.

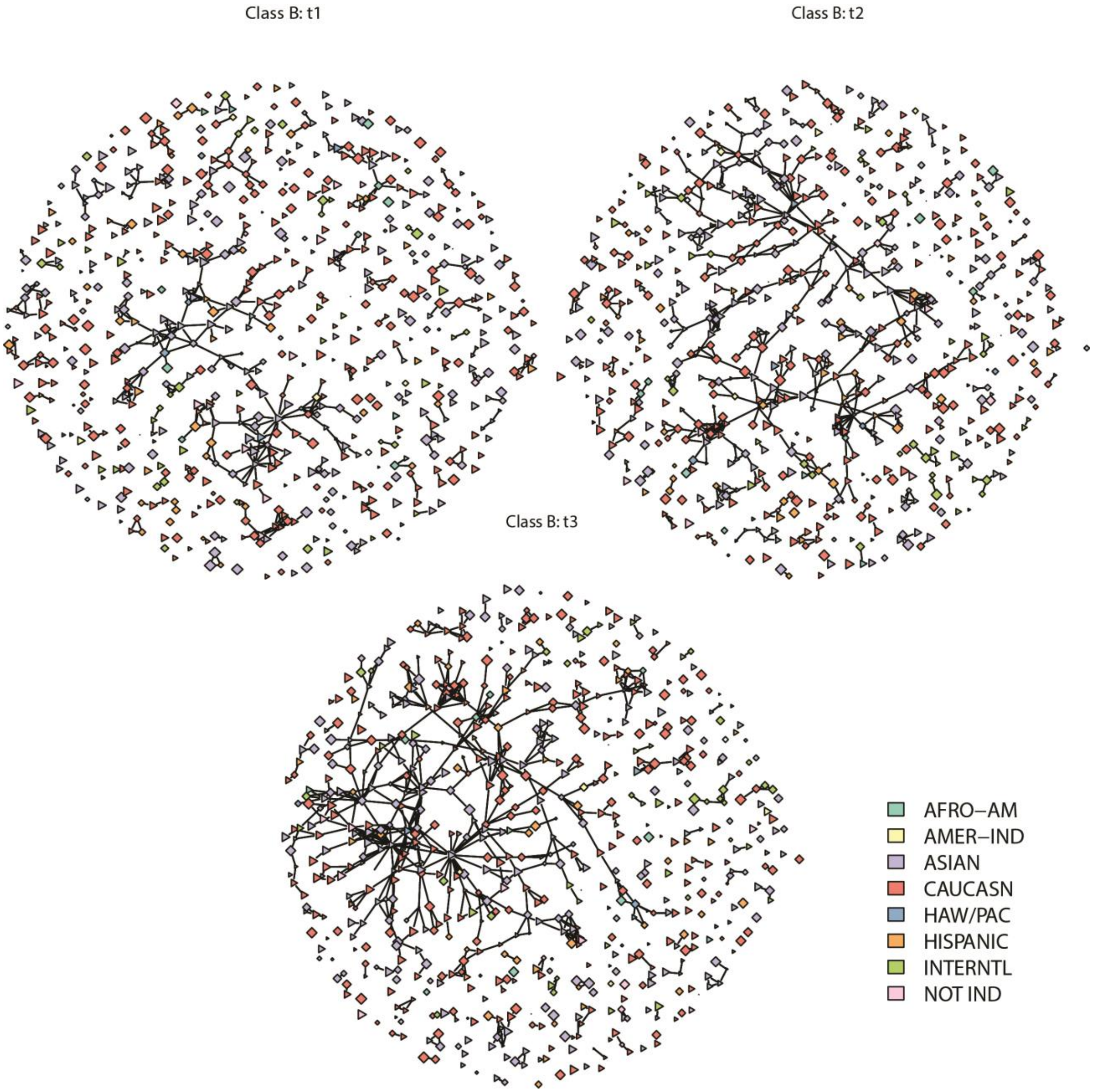


Figure 3.2 Error! No text of specified style in document.**:** Class B study networks. Class B study networks at times *t=1*, *t=2*, and *t=3*. Sizes of nodes correlates with grade in course, so larger nodes are students with higher grades. Triangles are male students, while squares are female students. The color of the node denotes student ethnicity.

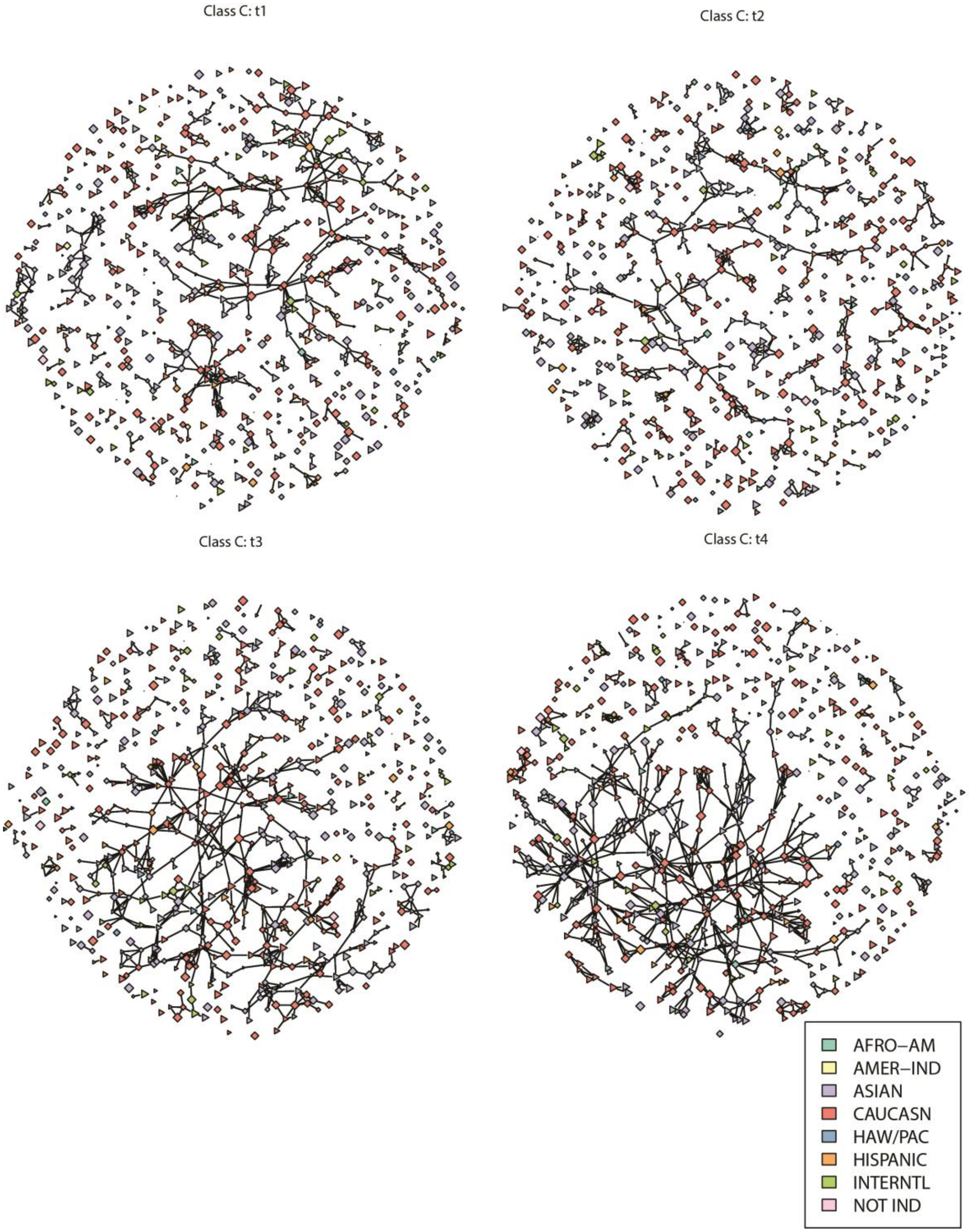


Figure 3.3 Error! No text of specified style in document.**:** Class C study networks.

Class C study networks at times *t=1*, *t=2*, and *t=3*. Sizes of nodes correlates with grade in course, so larger nodes are students with higher grades. Triangles are male students, while squares are female students. The color of the node denotes student ethnicity**.**

Table 3.1: The number of study partnerships (edges) for each exam in all three classes. The number of new study partnerships, as well as the number of study partnerships that did not persist between exams, are shown in the offset columns for each time step.

<table>
<tr><th></th><th colspan="5">Class A - 196 Students</th><th colspan="5">Class B - 759 Students</th><th colspan="5">Class C - 760 Students</th></tr>
<tr><th>Exam</th><th>Study partnerships</th><th colspan="2">Newly formed partnerships</th><th colspan="2">Ended partnerships</th><th>Study partnerships</th><th colspan="2">Newly formed partnerships</th><th colspan="2">Ended partnerships</th><th>Study partnerships</th><th colspan="2">Newly formed partnerships</th><th colspan="2">Ended partnerships</th></tr>
<tr><td>1</td><td>169</td><td rowspan="2">89</td><td></td><td rowspan="2">63</td><td></td><td>495</td><td rowspan="2">267</td><td></td><td rowspan="2">115</td><td></td><td>622</td><td rowspan="2">176</td><td></td><td rowspan="2">246</td><td></td></tr>
<tr><td>2</td><td>195</td><td rowspan="2">80</td><td rowspan="2">47</td><td>647</td><td rowspan="2">273</td><td rowspan="2">227</td><td>552</td><td rowspan="2">343</td><td rowspan="2">159</td></tr>
<tr><td>3</td><td>227</td><td rowspan="2">-</td><td rowspan="2">-</td><td>693</td><td rowspan="2">-</td><td rowspan="2">-</td><td>736</td><td rowspan="2">269</td><td rowspan="2">177</td></tr>
<tr><td>4</td><td>-</td><td></td><td></td><td>-</td><td></td><td></td><td>828</td><td></td><td></td></tr>
</table>

Table 3.2: Coefficients from separable temporal exponential random graph models for the formation of new study partnerships between exams (top), and the persistence of study partnerships between exams (bottom). Times correspond to exam numbers in each class, so t_2 -> t_3 results are for a model looking at the formation and persistence of study partnerships between exams two and three. Positive coefficients in the formation model (top) indicate that ties are more likely to form, while negative coefficients indicate that ties are less likely to form. Positive coefficients in the dissolution model (bottom) indicate that ties are more likely to persist, while negative coefficients indicate that ties are less likely to persist. Parentheses represent standard errors on estimates.

	Class A	Class B		Class C		
Formation Variables	t_2 -> t_3	t_1 -> t_2	t_2 -> t_3	t_1 -> t_2	t_2 -> t_3	t_3 -> t_4
edges	-5.905*** (0.206)	-7.951*** (0.122)	-7.373*** (0.111)	-7.734*** (0.136)	-7.789*** (0.121)	-7.512*** (0.116)
Same lab	0.814** (0.271)	1.408*** (0.157)	1.498*** (0.148)	1.095*** (0.267)	1.539*** (0.165)	1.886*** (0.163)
Same ethnicity	0.500* (0.228)	0.535*** (0.126)	0.544*** (0.122)	0.585*** (0.156)	1.045*** (0.114)	0.512*** (0.126)
Same gender	-0.0148 (0.228)	0.453*** (0.130)	0.153 (0.122)	-0.075 (0.155)	0.672*** (0.118)	0.191 (0.126)
Payoff bias	2.510*** (0.324)	1.904*** (0.244)	1.693*** (0.231)	4.327*** (0.263)	4.055*** (0.200)	4.350*** (0.224)

	Class A	Class B		Class C		
Dissolution Variables	t_2 -> t_3	t_1 -> t_2	t_2 -> t_3	t_1 -> t_2	t_2 -> t_3	t_3 -> t_4
edges	0.783* (0.365)	0.976*** (0.206)	0.538** (0.169)	0.074 (0.168)	0.462* (0.182)	0.807*** (0.190)
Same lab	0.035 (0.3654)	0.180 (0.298)	0.066 (0.221)	0.805** (0.258)	0.236 (0.269)	0.108 (0.244)
Same ethnicity	0.435 (0.348)	0.186 (0.222)	0.228 (0.174)	0.050 (-0.168)	0.103 (0.191)	0.226 (0.175)
Same gender	-0.188 (0.356)	0.161 (0.225)	0.021 (0.179)	0.077 (0.172)	0.319 (0.192)	0.048 (0.180)
Payoff bias	0.894* (0.351)	0.595 (0.484)	0.643* (0.249)	0.400** (0.128)	0.359** (0.138)	0.316* (0.124)

***p<0.001; **p<0.01; *p<0.05

Table 3.3: Coefficient estimates and standard errors for fixed effects from mixed effects model on exam Z-Score. Random intercepts are assigned for individual students and classes (B, and C).

Fixed effects of mixed-model for Exam Z-Scores (N=1491 exams)

	β	**SE**
Previous Class Grade	0.990**	0.048
Exam number		
2 versus 1	0.044	0.041
3 versus 1	0.050	0.045
4 versus 1	0.045	0.055
Number of study partners		
1 versus 0	0.096	0.058
2-3 versus 0	0.096	0.062
>3 versus 0	0.192*	0.078
Use of payoff-bias	-0.007	0.049
Constant	-3.128**	0.150

$^{**}p < .01$; $^{*}p < .05$

Chapter 4: Evidence that *DRD4* 48 base pair VNTR differentially impacts student preferences for highly active learning activities

Abstract

The 48 base pair variable number tandem repeat (VNTR) polymorphism of Dopamine Receptor D4 (*DRD4*) has been extensively studied, specifically the 2R and 7R alleles. These alleles have been linked to inattention and hyperactivity, but show evidence of positive selection in populations historically migratory populations. Evolutionary models of social learning behaviors suggest that quickly changing environments may select for learning styles conducive to informational exploration, which may explain the endophenotypes associated with 2R and 7R. Other research has shown these alleles to be associated with norm sensitivity; individuals who carry these alleles are sensitive to the environment of their upbringing. I test whether *DRD4* impacts learning preferences for pedagogy that ranges in the amount of passive listening required, either directly or through an interaction with student ethnicity. International students who are carriers of 2R and 7R alleles prefer more passive classroom days compared to non-carriers, while the effect of these alleles is opposite for American born Asian students and Caucasian students. This work lends support for the norm-sensitivity hypothesis of *DRD4*, and indicates that these alleles may impact learning styles.

Introduction

An impressive ability for socially learning and innovating new information is one of *Homo sapiens'* most distinguishing traits. Culture use has been instrumental to human range expansions across the planet, and continues to be a hallmark of humanity (Henn, et al. 2012; Henrich and McElreath 2003). Gene-culture co-evolutionary theory (GCC) takes an evolutionary approach to cultural change and human learning behaviors, acknowledging the dependencies each have on the other (Richerson and Boyd 2008). This framework suggests that some basic social learning behaviors are innate to human learning psychology, and that cultural evolution is directed partially by these learning behaviors (Laland 2004; Rendell, et al. 2010; Rendell, et al. 2011). Further, it posits that human genetic evolution responds to the emergent cultural landscapes (Gerbault, et al. 2011; Richerson and Boyd 2008; Tishkoff, et al. 2007). Advances in behavioral and population genetics make GCC a useful paradigm to understand potential selective pressures underlying recent changes in allele frequencies. This is especially true for genes whose impacts are well understood (Laland, et al. 2010; Richerson, et al. 2010) and are linked to learning behaviors. A gene that fits both of these conditions is the Dopamine Receptor D4 (*DRD4*).

DRD4:

The dopaminergic system is important in reward based behaviors, and *DRD4* is one of several receptors important to the regulation of dopamine signaling. *DRD4* exhibits a well-studied 48 base-pair variable number tandem repeat (VNTR) within the third exon (Ding, et al. 2002; Grady, et al. 2003; Van Tol, et al. 1992a). Globally, the most common *DRD4* 48 bp VNTR allele is 4R, named after its four repetitions of the 48 bp sequence. The most common variants other than 4R are the 2R and 7R variants, also named for their copy numbers. Combined, these three alleles account for more than 90% of the global variation of this VNTR (Chang, et al. 1996; Ding, et al. 2002; Matthews and Butler 2011).

Compared to 4R, the 2R and 7R alleles have a blunted response to dopamine, with 7R having a more blunted response than 2R. This altered response results in a lowered intracellular secondary messenger response (Asghari, et al. 1995). *DRD4* is highly expressed in the pre-frontal cortex, an area important for cognition and attention, making this altered response to dopamine behaviorally relevant (Oak, et al. 2000). Both 2R and 7R show links to Attention Deficit Hyperactivity Disorder (ADHD) (Faraone, et al. 2014; Gizer, et al. 2009; Grady, et al. 2003; Leung, et al. 2005a). 7R, which has been more extensively studied than 2R, has been linked to reward based behaviors such as addiction (McGeary 2009) financial risk taking (Dreber, et al. 2009), and has been investigated for potential links to novelty seeking (Munafò, et al. 2008).

While some research shows a direct effect of 2R and 7R on behavior, other research suggests that these alleles, along with other polymorphic genes in Dopaminergic pathways (such as *DRD2*) affect behavior through gene by environment interactions (Belsky, et al. 2007). That is, these alleles are not *risk* alleles, but instead moderate both “positive” and “negative” behavioral effects dependent on the childhood environment of the carrier. A meta-analysis of gene by environment studies on dopamine-related genes provides perhaps the best evidence for this. Covering 12 different studies on dopaminergic genes, Bakermans-Kranenburg and van Ijzendoorn (2011) show that children who carry “risk” alleles are not only more susceptible to negative outcomes if they experienced an adverse upbringing, but they are also more susceptible to positive outcomes in more nurturing upbringings. For example, research has shown that individuals who carry the 7R allele (7R+) who grew up with parental problems report higher unresolved loss and trauma compared to non-carriers, but report lower unresolved loss and trauma than non-carriers if they experienced few parental problems (Bakermans-Kranenburg, et al. 2011).

A recent study examined the self-construal of Americans of European ancestry and Asian-born Asian undergraduates in the United States (Kitayama, et al. 2014). Results indicated that participants who were carriers of 7R and 2R alleles were more congruent with cultural norms. When surveyed for “interdependence” versus “independence”, self-construals traditionally associated with Eastern and Western cultures, respectively (Markus and Kitayama 1991; Nisbett 2010), Asian carriers of 2R or 7R were more interdependent, while European-American carriers were more independent (Kitayama, et al. 2014). These interactions have led some to suggest that these dopamine alleles often associated with risk should really be understood as being associated with “differential susceptibility” or “norm sensitivity” (Bakermans-Kranenburg and van Ijzendoorn 2011; Kitayama, et al. 2016; Kitayama, et al. 2014)

ADHD, DRD4 and learning behaviors:

Based on linkage disequilibrium patterns, 7R is estimated to have arisen from a single mutational event 40,000-50,000 years ago, and the 2R allele around 10,000 years ago (Ding, et al. 2002; Matthews and Butler 2011). Despite their implication in “negative” endophenotypes such as ADHD, the 2R and 7R alleles show a pattern of positive selection (Ding, et al. 2002; Matthews and Butler 2011; Wang, et al. 2004) (but see (Kidd, et al. 2014; Naka, et al. 2011)). Population-specific frequencies of these alleles are highly correlated with migration rates (Chen, et al. 1999; Matthews and Butler 2011). This population-level pattern suggests that these alleles led to higher likelihoods of migration, and/or migrating populations with higher frequencies of these alleles had an adaptive advantage over populations with lower frequencies (Chen, et al. 1999).

These coalescent date estimates for the 2R and 7R alleles occur at times critical in human migration. While early human migrations out of Africa are of great debate, recent evidence suggests that non-African populations stem from a population which started its migration around 55,000 years ago (Posth, et al. 2016). Similarly, this timing aligns closely with the beginning of the Upper Paleolithic, where major cultural advances in tool complexity, along with abstract culture such as art are found (Bar-Yosef 2002). Given the nature of 7R and 2R's behavioral links, the timing of coalescence for these alleles, and the population pattern reflecting migratory history, I argue that these alleles may be a product of selection for diverse learning behaviors. This possibility is supported by the environment where selection for these alleles likely occurred: frontier environments encountered by migrating populations.

Learning behaviors and environmental context:

The importance of the environmental context to learning behaviors is an area of investigation within GCC. Results from GCC models show that the utility of copying someone else (social learning), and trial and error learning (individual learning), differs based on the stability of the environment (Feldman, et al. 1996). Social learning is more adaptive when environments are stable, while individual learning is particularly adaptive when environments are unstable (Feldman, et al. 1996). The reasoning is as follows: socially available information can become obsolete when an environment changes, making individual learning a potentially better learning strategy in changing environments. If an environment never changes, then socially available knowledge never goes obsolete, and social learning is reliable. Thus, the frequency and types of learning strategies may differ between populations that experience stable environments compared to populations experiencing more frequent environmental changes (Chang, et al. 2011). In fact, experimental evidence suggests that college students in mainland China use social information more than British college students, with Chinese students in college in the United Kingdom displaying an intermediate use of social information (Mesoudi, et al. 2015a). This work suggests that we culturally learn how to learn, but does not exclude potential impacts of genetic influences.

Some of the symptoms associated with ADHD appear to be a behavioral realization of individual learning as defined by GCC. That is, inattention and hyperactivity represent possible ways to happen upon new information or new ways to solve problems. These traits could lead to more frequent, error prone, or creative trial and error, increasing opportunities for novel information to be uncovered. Indeed, studies have shown that despite some associated cognitive deficiencies, ADHD is associated with some advantages during creativity tasks (Abraham, et al. 2006). For example, individuals with ADHD have a heightened ability to arrive at several solutions or generate multiple ideas compared to peers without ADHD (White and Shah 2006; White and Shah 2011).

Evidence for the value of inattentiveness for problem solving comes from a cooperative learning task, where triads of middle school children were formed to complete a complex task. Half of the triads included a student who displayed symptoms associated with ADHD. Compared to groups without a student with these symptoms, these triads were much more successful at solving the task. This occurred despite more distractive behavior occurring in the groups with the inattentive students (Zentall, et al. 2011).

In frontier environments encountered during range expansions, increased levels of individual learning could have facilitated cultural adaptation to novel ecological challenges.

Migratory populations may have benefited from having individuals with lowered inhibitory control, increased creativity, and increased inattention. By benefitting the group, alleles that led to these exploratory behaviors could have experienced group selection. A simulation model of a changing foods task, where unpredictable exploration by an individual can enhance group level fitness, lends evidence that group-selective dynamics in a similar learning task could lead to increased frequencies of individually costly alleles, which could help explain the high rates of seemingly maladaptive behaviors such as ADHD (Williams and Taylor 2006). While there is considerable skepticism about the prevalence of and power of group selection in biology (Hamilton 1963), increasing evidence suggests that the group structure and culturally mediated behaviors of our species may make group selection a more powerful force in humans than most other mammals (Soltis, et al. 1995; Wilson 1975; Wilson and Sober 1994). Further, game theoretic models and empirical work have shown that group selective forces are indeed stronger in expanding populations due to assorting cooperators together towards the ranges (Van Dyken, et al. 2013).

With links to ADHD, as well as increased frequencies in migratory populations, both 7R and 2R may have been selected in frontier settings. In this account, carriers of these alleles may be thought of as prone to individual learning behaviors. However, given growing evidence that the behavioral impacts of these alleles occur through differential susceptibility, learning styles of 2R+ and 7R+ individuals may depend on the norms of their cultural upbringing. Given the opposite directions of impact in the case of differential susceptibility, an evolutionary understanding for DRD4 is difficult to untangle in this scenario. In either case, if *DRD4* is implicated in learning styles, then students should have different learning experiences based on their *DRD4* genotype.

DRD4, active learning, and the classroom:

Pedagogy aimed at incorporating exploration and active engagement with material, which values divergent thinking and requires personal discovery of material, should cater to "individual" learners. This type of pedagogy is at the core of active learning (Bonwell and Eison 1991), where the structure of the classroom forces the construction of knowledge onto the students. This contrasts to passive instruction, where students sit quietly and listen to content delivered by the instructor with limited effort to ensure students work through problem solving. Indeed, there is evidence that the main deficit for school aged students with ADHD are during passive learning behaviors, such as quietly reading or listening to a lecture, and not active learning behaviors, such as talking about class material or writing (Junod, et al. 2006).

If being a carrier of the 2R or 7R allele (2R+ and 7R+) is linked to "individual" learning, then learning gains and engagement for 2R+ and 7R+ students are likely to be higher in this kind of pedagogical environment. If 2R and 7R work through norm-sensitivity, but are still connected to learning behaviors, then this link may be mediated by the student's cultural background.

Here I examine whether 7R+ and 2R+ individuals differ in their learning gains and engagement during intensively active learning days compared to passive days in a large lecture undergraduate Biology classroom. In some ways, undergraduate classrooms represent the pinnacle of human social learning, where complex information is aimed to be disseminated to hundreds of individuals. Traditionally, classes at the collegiate level rely on passive lecturing. Within the United States, a national push to move away from passive instruction in favor of active learning is gaining momentum (Olson and Riordan 2012). For this reason, this examination not only can inform our understanding of human evolution and learning, but how pedagogical choices may have positive or negative impact for students whose diversity is not visible to instructors.

Hypotheses:

Hypothesis 1 *– 7R+ and 2R+ students will report higher engagement than other genotypes on days taught through intensive active learning. This difference may vary based on whether or not the activity is highly social and by the ethnicity of the student.*

Hypothesis 2 *– 7R+ and 2R+ students will experience higher learning gains than other genotypes on days taught through intensive active learning. This difference may vary based on whether or not the activity is highly social and by the ethnicity of the student.*

Methods

Active learning activities and student demographics:

Data come from a 759 student introductory biology classroom at an R1 University. Lecture sections took place twice a day and were taught by the same instructor, with half of the students attending an early lecture, and the other half attending an afternoon lecture. On four occasions throughout the quarter, students were taught through a 50-minute long active learning module. The content taught in these active learning modules was translation, eukaryotic gene regulation (EGR), polymerase chain reaction (PCR), and cell cycle (CC). These four days were either taught through a *constructive* or *collaborative* version of the active learning module. These modules were designed according to the *ICAP* (Interactive-Constructive-Active-Passive) framework (Chi and Wylie 2014). In this framework, interactive (hereafter referred to as collaborative) activities are expected to provide the deepest learning by having students dialogue and build on each other's knowledge. The second deepest learning gains are expected by constructive activities, which rely on generating and producing knowledge, but without a necessary collaborative component.

The administration of constructive and collaborative days took advantage of the split lecture design. On each of the four days, one lecture received a constructive activity, while the other received an equivalent activity designed to be collaborative by requiring peer exchange of ideas. On three of the four occasions, the morning lecture received a collaborative activity while the afternoon section received a constructive activity, and on one of the days this was reversed. By design, all students experienced at least one day of constructive active learning and one of collaborative.

The night before each of these activity days, students took an online pre-quiz of eight test-level questions on the material to being taught. The night after the activities, students took the same eight question quiz to assess learning gains. Along with this post-quiz, students filled out a piloted survey to assess their level of engagement compared to normal class days, along with other measures of interest. The engagement survey was administered as part of an ongoing survey development. The survey included, among other questions, a set of eight dichotomous questions which specifically ask the student to contrast their engagement during active learning days to normal lecture days.

Student self-reported ethnicity was obtained from the university registrar. All research procedures were approved through the University of Washington IRB.

DNA extraction and genotyping:

398 students provided saliva samples via Oragene Saliva kits OG-500 (DNA Genotek, ON, Canada), with extraction conducted according to the manufacturer (Birnboim 2004).

Of the 398 samples, 393 (98.7%) DNA samples were extracted at high enough quantity and quality for PCR amplification. Following Eisenberg, et al. (2009), the PCR consisted of 1× Q-Solution (Qiagen), 1× buffer (Qiagen), 1 µM primer 1 (5′GCGACTACGTGGTCTACTCG 3′), 1 µM primer 2 (5′AGGACCCTCATGGCCTTG 3′), 200 µM dATP, 200 µM dTTP, 200 µM dCTP, 100 µM dITP, 100 µM dGTP, 0.3 U of HotStar Taq (Qiagen), and 1 µl of DNA template, in a total volume of 10 µl. The PCR profile began with 15 min at 95°C for enzyme activation and denaturing of template DNA followed by 40 cycles consisting of 1-min denaturation at 94°C, 1-min annealing at 55°C, and 1.5-min extension at 72°C; it finished with a 10-min extension at 72°C. Amplicons were electrophoresed through 2% agarose gels containing SYBR Safe (Invitrogen). Genotypes were determined after being compared to 100 bp ladders.

To counteract common problems with allelic dropout at the DRD4 locus, (Eisenberg, et al. 2008; Eisenberg, et al. 2009; Eisenberg, et al. 2007; Hamarman, et al. 2004), several measures were taken to obtain accurate genotypes. To account for the high GC content in the amplified region, the PCR reaction included nucleotide analog dITP in the dNTP mix (Motz, et al. 2000) and Q-Solution (Qiagen). To confirm that samples scored as homozygotes were not heterozygous, all homozygotes were reanalyzed at 1/40X and 1/80X DNA template dilutions. If heterozygosity was found in any amplification of the same sample this was assumed to be the true genotype. Hardy–Weinberg (HW) equilibrium was tested with the Markov chain algorithm (Guo and Thompson 1992) in *R* with the *gap* package (R Core Team 2013; Zhao 2007). The genotype frequencies from the whole sample departed significantly from HW equilibrium (Total sample: $p = 0.015$). The non-equilibrium may be caused by the mixed ethnicities within the population. When the analysis was stratified by the two largest ethnic groups - students who identified as Caucasian (n=177) and students who identified as Asian (n=148), the genotypes of both subsamples were consistent with HW equilibrium (Caucasian: p = 0.2447; Asian: p = 0.7393).

Following previous work with mixed ethnicity populations, students were scored as carriers if they had at least one copy of either the 7R or 2R allele (Kitayama, et al. 2014; Reist, et al. 2007; Sasaki, et al. 2011). This reflects the physiological difference of higher dopamine signaling in carriers compared to non-carriers (Asghari, et al. 1995).

Statistical Analyses

For several reasons, analyses focused only on students in the course who were listed as “Asian”, “Caucasian”, or “International” by the registrar. First, an overlap of research on differences in social learning behaviors and the behavioral effects of DRD4 exists between populations from East Asian and European or American descent (Chang, et al. 2011; Kitayama, et al. 2014; Leung, et al. 2005a; Mesoudi, et al. 2015a; Zhang, et al. 2012). Second, the demographics of the class where collection took place are majority White and Asian, with a large International Asian population.

To assess dimensionality of the eight engagement survey questions, a polychoric correlation matrix of the dichotomous data was calculated using the *Polycor* package in *R* (Fox and Fox 2010) for each of the four activities. A priori, it was hypothesized that these questions all underlie a single latent variable of student engagement during the activity compared to normal class days. Confirmatory factor analysis with a varimax rotation and parallel analysis was used to assess whether a single factor represents a reasonable model for these questions (Hayton, et al. 2004).

To reduce these questions into a single latent score, the eight responses from each of the four activity days were fit to dichotomous Item Response Theory Models. These models assume an underlying latent ability of subjects that underlies their responses to items. Three different nested models are possible – including a one-parameter logistic model (commonly referred to as a Rasch Model), a two-parameter logistic model (2PL), and a three-parameter logistic model (3PL). Rasch models only parameterize the difficulty of different items, 2PL models add a parameter for how much discrimination is offered by each item, and the 3PL model adds a parameter that accounts for guessing in subject responses (Embretson and Reise 2013). Observed response patterns from the four administrations of the survey were fit to these three models, and the best fit model was assessed through likelihood ratio tests. The best fit model for each survey was used to obtain estimates of students' latent abilities (referred to as engagement scores hereafter) using the *ltm* package in *R* (Rizopoulos 2006). Because engagement scores are model estimates they have associated standard errors which are accounted for in regression methods as explained below.

Mixed effects models were used to assess the relation between DRD4 genotype and student engagement scores. These data include repeated measures on the same students across the four different activity days, breaking independence assumptions of typical linear regression models. Mixed effects models were thus used to take into account the variance within repeated responses from each individual by modeling random intercepts for individuals. The model tested was:

Engagement score ~ Activity.day + Activity.Type + DRD4 + Ethnicity +
Activity.Type*DRD4*Ethnicity + (1|Student.ID)

The latent engagement scores from the IRT models were the dependent variable of interest. Activity.day was modeled as a fixed effect to control for any differences between the content. Activity type (collaborative or constructive), along with whether a student is a carrier of 2R or 7R, student ethnicity, and interactions between these three variables were dependent variables of interest. To account for error surrounding the engagement score estimates, 10 iterations of this model were run, each with an imputed engagement score drawn from a normal distribution with individual estimates as the mean and the associated errors. Fixed effects from these models were pooled to test for an association between DRD4 and student engagement.

For hypothesis 2, learning gains (post-test score – pre-test score) were measured as a continuous dependent variable of interest. Similar to the first hypothesis, Student ID was modeled as a random effect to control for the shared within-subject variation. Similar to hypothesis one, the model tested had the same fixed effects of interest:

Learning gains ~ Activity.day + Activity.Type + DRD4 + Ethnicity + Activity.Type*DRD4*Ethnicity (1|Student.ID)

The current hypotheses have substantive interests in different cultural backgrounds of students. Beyond the intrinsic interest in the impact of student ethnicity, including ethnicity in these models was also important for the sake of population stratification. Population stratification occurs when behavioral, allelic frequency, and ancestry co-vary, making it critical to control for shared ancestry (Hamer 2000). While ancestry informative markers would be a more precise control for genetic ancestry, these markers have not been analyzed for this sample, and registrar data are used instead.

Significance testing:

Parameter-specific p-values are difficult to obtain in mixed models, largely due to difficulties with calculating degrees of freedom (Bolker, et al. 2009). We calculate degrees of freedom using the Satterthwaite method (Satterthwaite 1946) using the *lmerTest* package in *R* (Kuznetsova, et al. 2013).

Results

Allele and genotype frequencies in students of European and Asian ancestry:

The allele and genotype frequencies from the analyzed sample are shown in **Error! Reference source not found.**. In line with global allelic frequencies, 4R is the most common allele in all populations; 7R is the second most common allele in Caucasian students (19%), while 2R is the second most common allele in Asian students (17%). The majority of international students attending this university are Asian (84.7%), and over half of all International students are from China (52.7%) (University of Washington International Student Services, 2014). While the exact proportion of Asian International students in this sample is unknown, the high frequency of the 2R alleles among this sample is consistent with a majority Asian population of International students (a recent study with 2,191 Han Chinese research participants had a 2R allele frequency of 20.7% (Jiang, et al. 2015)).

Confirmatory Factor Analysis and Factor Scores:

Confirmatory factor analyses were run separately on the four engagement surveys administered in this course to assess whether a one-factor model fit the eight dichotomous questions of interest. All students who completed the survey were included in these analyses unless they took less than 2 minutes to complete the entire survey (which consisted of 20 questions total) or were not present in class the day of the activity. Several goodness-of-fit indices are used to assess fit. A chi-squared test, root mean square error of approximation (Browne, et al. 1993; Steiger and Lind 1980), and goodness of fit index (Bentler 1990) were considered here in examining whether one factor is appropriate. Frequently used cutoffs include values lower than 0.1 for RMSEA and greater than 0.9 for GFI (Browne, et al. 1993; Hair 2010). Table 4.2 shows these metrics for the data from the four different surveys.

While two-factor models represent better fits for these data, a one-factor model is reasonable. The research focus is on the overall engagement of students which is encompassed in the nature of all eight questions. Further, pairwise comparisons between every combination of the eight items show a consistently high and significant association on every survey. Thus, these eight items were accepted as unidimensional.

Survey responses from the PCR, EGR, and CC activities best fit the 2PL IRT model, while responses from Translation best fit the 3PL model. The difficulty of the items tended to be low; students were more likely to agree with any of the eight items than they were to disagree. Latent scores for each student were estimated from these models. Latent scores take the form of estimated Z-scores, and measure a student's relative engagement compared to their peers. Because the items tended to be easy, this survey did not perform well at discriminating between students who were at the higher end of enjoying the active learning activities. The distributional patterns of the estimates from each survey indicate a mode at the maximum factor score (**Error! Reference source not found.**). This is possibly driven by the tendency for "easy" items, but also seems to indicate that most students were more engaged during active learning days than normal class days.

Exploratory analyses:

Analyses focused only on students who had amplified DNA and were listed as Caucasian, Asian, or International from the registrar. After this filtering, the analyses for student engagement focused on 346 total students with 1286 total surveys across the four activities. Likewise, the analyses on student learning gains only focused on the 319 students who completed both the pre-quiz and the post-quiz for activity days, totaling 747 total pre-post data points.

Across the engagement surveys, the mean latent score for the 346 included students was -0.021 (sd = 0.727). Higher scores indicate stronger preferences for active learning days compared to non-active learning days. These data include repeated measures and are estimates with associated standard errors. **Error! Reference source not found.** shows the mean and standard errors of the estimated engagement scores by ethnicity, genotype, and activity type, but does not take into account the repeated measures or error in engagement score estimates. For this reason, this figure should be viewed as providing a general understanding of response patterns which can aid in understanding model results.

Across the pre-quizzes, students averaged 4.49 out of 8 points (sd = 1.56). The average only increases to 4.86 out of 8 points on the identical post-quizzes (sd = 1.78). These data are similarly hard to visualize due to the repeated measure aspect. **Error! Reference source not found.** shows the mean learning gains by ethnicity, genotype, and activity type, but does not separate out the impacts of repeated measures within students.

Engagement score model results

The fixed effects at the level of the student are of primary interest. Pooled estimates along with 95% confidence intervals are displayed in **Error! Reference source not found.**. Coefficients that do not bound zero are significant. The only significant variable was the interaction between being 2R+/7R+ and being an International student, which had a negative impact on engagement. This signifies that the direction of the effect of 2R and 7R alleles on engagement was in the opposite direction for International students as it was for Caucasian students. No significant association existed between engagement score and any of the other fixed effects.

Learning gains model results

The fixed effects from the mixed model predicting student learning gains are shown in **Error! Reference source not found.**. There was no direct or interactive effect of DRD4 genotype on learning gains. International students had very low, in fact negative, learning gains on average for constructive active learning modules. The only other significant predictor of student learning gains was the activity day, where learning gains on the PCR activity were lower than on the reference Cell Cycle day. A visual inspection of the residuals from this model did not reveal any deviations from normality or homoscedasticity.

Discussion

This study examines whether *DRD4* genotype impacts social learning behaviors. Specifically, do 2R+ and 7R+ students differ from non-carriers in engagement and learning gains in highly active learning classroom structures? Underlying the current hypothesis is an assumption that learning styles which thrive under active learning are likely similar to an individual learning type that would have been beneficial in frontier settings. 2R and 7R alleles may have experienced positive selection in frontier settings due to associated increases in individual exploration and divergent thinking; if so, students who are carriers of these alleles may be predisposed for a more positive experience in active classroom environments than non-carrier peers. However, evidence suggests that any association between DRD4 and learning style may be moderated by cultural differences, in which case this association may depend on student ethnicity.

DRD4 genotype does not have a significant direct impact on engagement scores or learning gains. Ethnicity also shows no significant effect for either of the dependent variables of interest. However, there is a significant interaction between *DRD4* and International status. This result implies a cross-over interaction between ethnicity and genotype. That is, while 2R+/7R+ Caucasian students reported greater engagement than non-carrier Caucasian students for both constructive and collaborative days (although, this effect was not significant, $p = 0.18$), 2R+/7R+ International students reported significantly lower engagement than International students without these alleles ($p = 0.05$).

This result is in line with previous work suggesting that *DRD4* operates through differential susceptibility (Belsky, et al. 2007), and more specifically, with Kitayama et. al.'s (2016) Norm Sensitivity hypothesis. The Norm Sensitivity hypothesis suggests certain genetic polymorphisms may modulate how readily individuals acquire cultural norms and behaviors (Kitayama, et al. 2016). Empirical evidence for this hypothesis comes from heightened independent self construals in 2R+/7R+ students of European descent, and heightened interdependence in 2R+/7R+ Asian-born Asian students (Kitayama, et al. 2014). Based on normative differences between Eastern and Western educational cultures, the current findings lend further support to this hypothesis.

Clear differences in educational practices between Eastern and Western cultures (Chinese and American specifically) exist (Dai, et al. 2011; Ming-li 2012). Chinese educational norms emphasize deference to the instructor, less peer interaction, and a higher valuation of memorization compared to innovative thought (Zhang and Dai 2004). The active learning days under examination involved almost no lecture from an instructor, required group work (especially in collaborative days), and put no focus on rote memorization. If 2R and 7R alleles make individuals keenly susceptible to environmental influences, then 2R+/7R+ individuals from Chinese schooling may find a greater mismatch on active learning days compared to non-carriers. This reflects the finding among International students, where the lowest engagement for 2R+/7R+ students appears to take place in collaborative learning activities which required the most peer interaction. The effect of DRD4 on Caucasian students, while not significant, was in the opposite direction; 2R+/7R+ individuals reported higher engagement. Importantly, the direction of the effect of DRD4 for American born Asian students is more similar to that of Caucasian students than International students. This implies that the root of this difference is likely the difference in the educational system one is brought up in, and perhaps less to do with simply being of an Asian ethnicity.

If the 48 bp VNTR in *DRD4* operates through norm-sensitivity, this could explain inconsistent findings between 7R and novelty seeking. 7R has previously been implicated with novelty seeking behaviors (Lusher, et al. 2001), and is frequently cited as a “Novelty Seeking” allele. However, more recent meta-analyses examining this link show that this association is not significant (Kluger, et al. 2002; Munafò, et al. 2008). Under the norm-sensitivity paradigm, we may expect 7R to be associated with novelty seeking in some cultural contexts but not others. This is a hypothesis that likely warrants further investigation.

While the engagement data suggests an interesting effect of *DRD4*, there is no effect of *DRD4* on learning gains. Instead, the main finding is that international students have very low, in fact negative, learning gains following constructive activities. It is difficult to render a clear interpretation of this pattern. There were only four iterations of pre- and post-tests and 18

International students in the sample, so these data are sensitive to any tests taken with low effort which could introduce considerable noise to the dependent variable. Further, many factors can impact student learning in any single class day, which make measuring acute learning gains in environments such as these messy. These factors can include the peers students worked with, attitudes towards the classroom content, and individual circumstances that do not pertain to the classroom of interest (i.e. outside stressors) which are not captured here.

The lack of an effect of *DRD4* on learning gains is not necessarily surprising. The history of research looking at the effects of genotypes on educational gains has a history of very small effect sizes. Studies lacking extremely large sample sizes often fail to find genetic links to educational outcomes (Chabris, et al. 2013). When educational studies do have enough power, the variation explained by genetics is small. A recent genome-wide association study with 294,000 individuals found that only 3% of the variation in educational attainment was explained by their 74 identified significant loci (Okbay, et al. 2016). While the classroom sizes here are large, the student population in these models does not exceed 346, with only 19 International students. While this sample size has been sufficient in other settings, it may not be enough to identify potentially small effects in the experimentally noisy setting of a classroom.

Conclusion

This study provides some support for a norm sensitivity role of *DRD4*, ultimately impacting the engagement of International students in a highly active learning environment. These results are intriguing, but are by no means conclusive. Further work would be required to confirm whether *DRD4* modulates pedagogical preferences, and whether or not selection for 2R and 7R alleles took place due to complex learning dynamics. A more controlled study would be useful to elicit if *DRD4* impacts learning styles. For example, studying the use of social and individual information in a controlled laboratory environment, with more pointed data on student backgrounds would be one way to further test DRD4's impact on learning style.

Table 4.1: Allele and genotype frequencies from the class, also broken down by students listed as Caucasian, Asian, Under-represented minority (URM), or International.

	Total		*Caucasian*		*Asian*		*URM*		*International*	
Allele	n	%	N	%	n	%	n	%	n	%
2R	114	15	42	12	51	17	12	13	9	25
3R	11	1	7	2	3	1	1	1	0	0
4R	528	68	228	64	218	74	57	62	25	69
5R	14	2	4	1	10	3	0	0	0	0
6R	10	1	4	1	3	1	3	3	0	0
7R	100	13	68	19	11	4	19	21	2	6
10R	1	0.1	1	0.3	0	0	0	0	0	0
Total	778	100.1	354	99.3	296	100	92	100	36	100
Genotype	n	%	N	%	n	%	n	%	n	%
2/2	15	4	5	3	7	5	2	4	1	6
2/3	1	0.3	1	1	0	0	0	0	0	0
2/4	74	19	27	15	34	23	6	13	7	39
2/6	1	0.3	0	0	0	0	1	2	0	0
2/7	8	2	4	2	3	2	1	2	0	0
3/4	10	3	6	3	3	2	1	2	0	0
4/4	180	46	73	41	80	54	18	39	9	50
4/5	12	3	2	1	10	7	0	0	0	0
4/6	9	2	4	2	3	2	2	4	0	0
4/7	63	16	43	24	8	5	12	26	0	0
5/7	2	1	2	1	0	0	0	0	0	0
7/7	13	3	9	5	0	0	3	7	1	6
7/10	1	0.3	1	1	0	0	0	0	0	0
Total	389	99.9	177	100	148	100	46	100	18	100

Table 4.2: Fit measurements for one and two factor models for each of the four administered surveys, including Chi-squared, goodness of fit index, and root mean square error of approximation.

Model	X^2	df	X^2/df	X^2 prob	GFI	RMSEA
Translation (n=693)						
Single Factor	140	20	7	<0.001	0.82	0.095
Two Factor	72	13	5.53	<0.001	0.86	0.081
EGR (n=704)						
Single Factor	270	20	13.5	<0.001	0.79	0.133
Two Factor	98	13	7.54	<0.001	0.86	0.097
CC (n=685)						
Single Factor	180	20	9	<0.001	0.80	0.108
Two Factor	35	13	2.69	<0.001	0.87	0.050
PCR (n=641)						
Single Factor	260	20	13	<0.001	0.82	0.136
Two Factor	58	13	4.46	<0.001	0.90	0.074

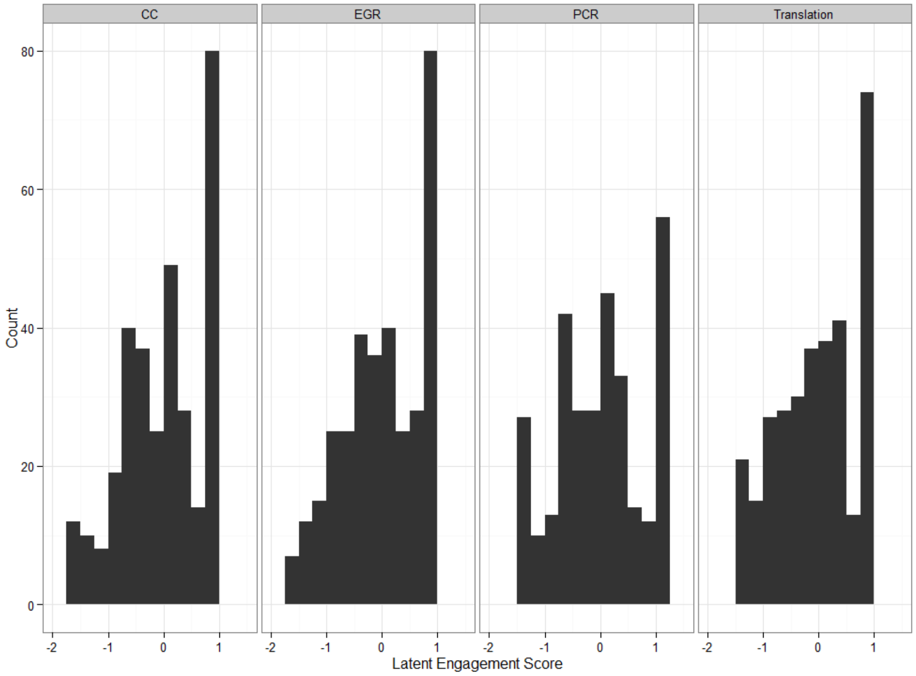


Figure 4.1: Distribution of engagement scores from best fit item response models. Distributions for the engagement scores from for the cell cycle activity (CC), eukaryotic gene regulation activity (EGR), PCR activity (polymerase chain reaction), and the protein translation activity.

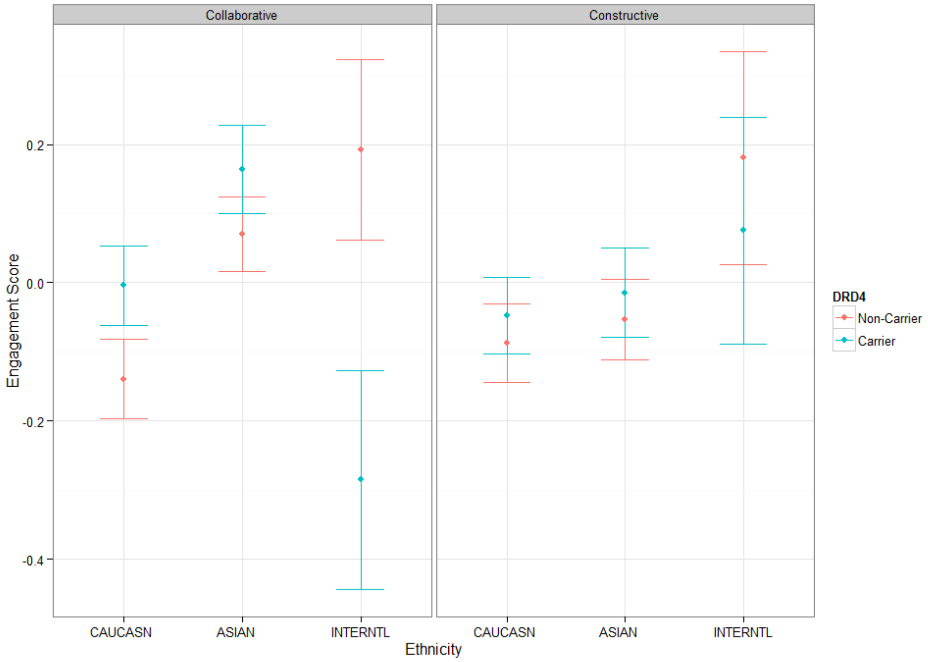


Figure 4.2: Student engagement by ethnicity, activity type, and DRD4 genotype.

The mean and standard errors of estimated engagement scores for carriers and non-carriers of *DRD4-2R* and/or *DRD4-7R*. This does not control for repeated measures within these data, or for the error surrounding the estimates. For this reason, this plot is best used to gain an understanding of the direction of engagement, but underestimates the standard errors surrounding estimates.

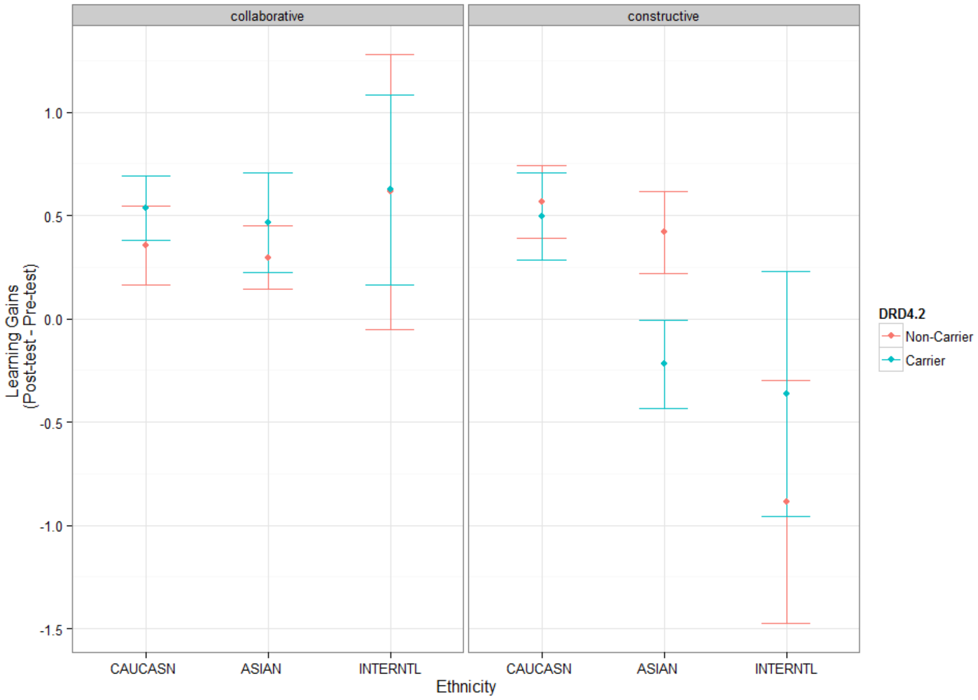


Figure 4.3: Student learning gains by ethnicity, activity type, and DRD4 genotype.

The mean and standard errors of learning gains (measured by the post-test scores minus the pre-test scores) for carriers and non-carriers of *DRD4-2R* and/or *DRD4-7R*. This does not correct for repeated measures within these data, so underestimates the standard errors surrounding estimates.

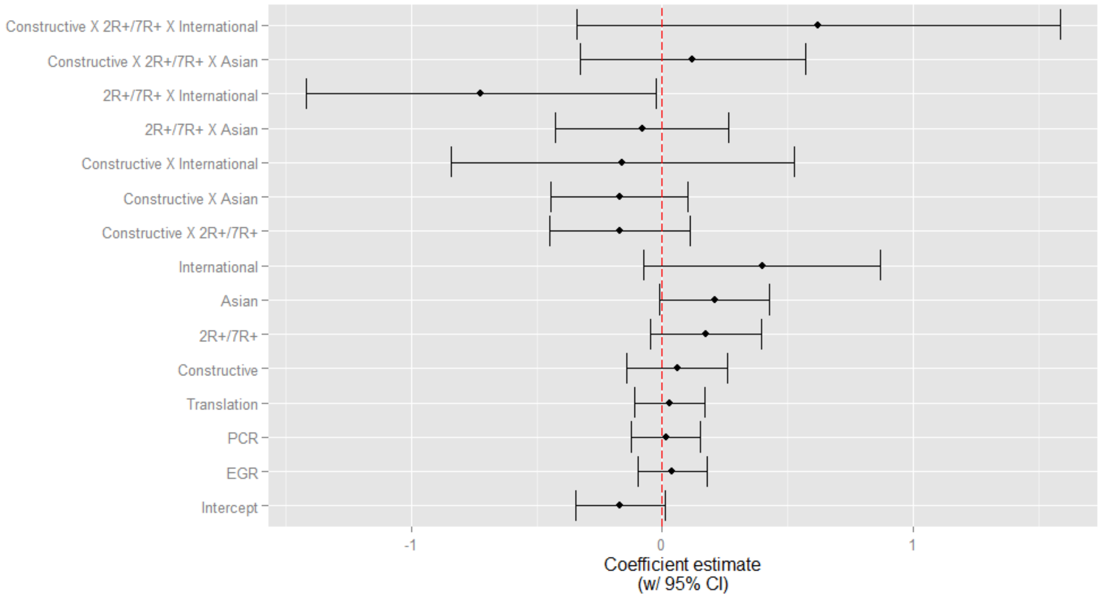


Figure 4.4: Coefficient estimates and 95% confidence intervals from the pooled fixed effects of the mixed effects model for engagement scores.

CI's that cross the red line are not significant at $\alpha = 0.05$. All of the variables in this model are factors, or interactions between factors. CC is the reference for activity day (EGR, PCR, Translation), Collaborative activity is the reference for activity type (Constructive), Non-carrier is the baseline for *DRD4* genotype (2R+/7R+), and Caucasian is the reference for ethnicity (Asian and International).

Table 4.3: Fixed effects results for the mixed model for learning gains. Significant effects are in bold.

Learning Gains Model		
	β	**SE**
Activity Day		
EGR (*Ref: Cell Cycle*)	-0.143	0.178
PCR (*Ref: Cell Cycle*)	**-0.467***	**0.184**
Trans (*Ref: Cell Cycle*)	0.073	0.189
Activity Type		
Constructive Activity (*Ref: Collaborative*)	0.286	0.273
DRD4		
2R/7R Carrier (*Ref: Non-carrier*)	0.214	0.249
Ethnicity		
Asian (*Ref: Caucasian*)	-0.049	0.250
International (*Ref: Caucasian*)	0.254	0.536
Interactions		
Constructive X 2R/7R Carrier	-0.316	0.370
Constructive X Asian	-0.125	0.371
Constructive X International	**-1.636***	**0.819**
2R/7R Carrier X Asian	-0.057	0.371
2R/7R Carrier X International	-0.152	0.849
Constructive X 2R/7R Carrier X Asian	-0.420	0.560
Constructive X 2R/7R Carrier X International	0.665	1.194
Constant	**0.451***	**0.212**
N	747	
Log Likelihood	-1493.640	
AIC	3021.279	
BIC	3099.752	

$^{**}p < .01$; $^{*}p < .05$; $^{.}$

Chapter 5: Discussion and Conclusion

This dissertation contributes to the GCC literature by empirically examining 1) how perceptions of others form in classrooms, 2) how these perceptions and other demographic characteristics impact the formation and persistence of learning relationships, 3) whether the use of a payoff bias increases learning gains in classrooms, and 4) how genes and culture can interact to impact learning styles. In doing so, it was shown that male students hold a perception bias towards their male peers; a finding that likely has complex social and biological roots. It was also shown that students tend to form study partnerships with peers they perceive as knowledgeable, and assort into same ethnicity study groups. Lastly, this dissertation provided some evidence for how genetic and cultural diversity can interact to predict student learning styles. Together, this work examines the diversity of learning behaviors within a college classroom.

The primary focus of this work is on student social behaviors, which remains a critical aspect of college classrooms to understand. The classroom networks of students, and the quality and level of interaction in these social ties may impact learning gains, engagement, and feelings of belonging. All of these can alter student trajectories, which not only impact individual experiences, but can result in inequity in fields if experiences differ depending on a student's background and identity. The first two studies revealed social processes in classrooms that may galvanize inequity in science fields. First, male students showed a gendered bias in who they viewed as knowledgeable in the class. Second, students preferentially studied with peers of the same ethnicity, which may create a situation where the social capital available to each student depends on their identity. Both cases provide further disadvantages to students from historically underrepresented populations in scientific fields. As previously discussed, a more comprehensive understanding of these processes is made available through a biocultural anthropological lens.

The last study investigated the differential impacts *DRD4* genotypes have on student learning preferences based on cultural background. While this adds another thread to the growing line of research on the 48 bp VNTR, supporting the norm sensitivity hypothesis, it also reveals that diversity in the classroom is not always visible. Here, the difference at one locus led to different learning preferences dependent on cultural background. Recognizing the diversity of learning preferences of students in classrooms is critical for effective instruction. Further, evidence that the ADHD-linked 2R and 7R alleles were selected for in past populations suggests that ADHD is not dysregulation, but a mismatch to the current learning environment.

Limitations and Future directions:

Each of the three studies in this dissertation has implications for future research. The first study leaves open several unanswered questions. Answering these questions would fill in current gaps in this study. First, are the observed gendered perceptions reflective of other behaviors that may impact retention in STEM? That is, do students who fail to receive recognition become less likely to earn a degree in Biology? Second, what impacts the gender bias? Repetition of this work in classrooms from diverse subjects and colleges, with diverse instructors and pedagogies would be needed to further understand what impacts these biases. Third, does similar bias impact individuals from diverse racial and ethnic backgrounds? Lastly, given answers to these questions, how can we, both as a society and college educators, work to eliminate these biases? Work on implicit biases and intentioned structuring of classrooms will be important to bringing diversity to fields where it currently is lacking.

The second study creates a basis for studying the impacts of study networks in classrooms. While the formation and persistence of groups was fairly well elucidated, there is likely some amount of error in the models due to missing data issues that are common in census networks. Further, the implications of having study partners remain an important area of investigation. While the current research used exam scores as the dependent variable of interest, these are messy outcome variables. Other outcomes such as study skills, enthusiasm for material, and persistence within a field may all be impacted by getting in with the right crowd in a classroom. Future network studies in classrooms with specified interests and longitudinal designs may be able to address these questions. It would also be interesting to know the importance of relationships outside of the classroom to these networks. As an example, does the expertise of a student in one class lead them to become a desired study partner in a class where they are not as strong? This could elucidate more clearly whether students are utilizing prestige bias or success bias in forming their study partnerships.

While the final study lends support for the norm sensitivity hypothesis regarding the *DRD4* 48 bp VNTR, it does not resolve the reason the 2R or 7R alleles reached higher frequencies in migratory populations. International student carriers had a lower preference for highly active learning days over passive days compared to non-carriers, while the effect of carrying 2R or 7R was in the opposite direction for American born Asians and Caucasians. While this suggests that a link to learning strategies may have played a role in its selection, more stringent tests of learning behaviors would be necessary to confirm this phenotypic difference.

Evolution and education:

While the three studies within this dissertation ask different questions, they each focus on the intersection of student attributes and the social environment of college classrooms. They further converge on their evolutionary approach to understanding learning behavior and human cognition. Thus, beyond the implications of these research findings already discussed within each chapter, this dissertation represents a larger contribution to a small field studying education through an evolutionary lens. This field focuses on understanding the functionally adaptive origins of human learning psychologies, and seeks to use this understanding to guide educational practices and interventions.

To date, evolutionary approaches on education have focused on the structure of social learning environments (Bjorklund and Bering 2000; Gray 2011; Wilson, et al. 2011), human cognition for mathematical abilities (Geary 1995; Geary 2007; Halpern, et al. 2007), underlying causes of bullying (Ellis and Bjorklund 2012; Volk, et al. 2012), instructional techniques to cope with cognitive load (Paas and Sweller 2012), and the physical act of teaching (Csibra and Gergely 2009; Kline 2015; Lancy 2010; Sterelny 2012). None of this work to date, however, has been directly applied to a higher educational setting. Thus, this dissertation represents, to my knowledge, the first evolutionary approach to higher educational research and policy.

GCC, higher education, and lecture: other directions for evolution and higher education research:

While Gene culture-coevolution was used in this dissertation to study social learning behaviors in classrooms, this is only one application of this paradigm to improving higher education. While an evolutionary approach to education should by no means be a sole approach to education research, it offers a unique perspective to several issues facing higher education. For example, moving away from strictly lecture based classrooms is a goal for many education specialists. However, while evidence has clearly shown that active learning classrooms are much more effective than lecture (Freeman, et al. 2014), changing the pedagogical behaviors of current instructors is extremely hard. Here I argue that GCC, with a focus on understanding cultural change in relation to transmission dynamics, can help us understand how a maladaptive practice such as lecture has stuck around for so long.

Lecture is the predominant pedagogical technique used in higher education, and has been since medieval times (Brubacher and Rudy 1997). Initially, lecture was used as a means to transcribe books, and was simply coopted as the way to teach college students the information in these books. Thus, the use of lecture did not start because it proved to be an effective method of teaching; it was simply a practice being used to transfer information from one page to another. One would both hope and imagine, then, that its current widespread use 700 years later is because it has since been shown to be the optimal means to train students. However, considerable evidence shows that active pedagogy, which involves considerably less passive lecture, leads to drastically greater student outcomes (Freeman, et al. 2014). This has been known by college educators for some time; lecture is a maladaptive means to teach students. How has lecture remained unchanged for approximately 700 years?

In their seminal book, Richerson and Boyd (2008), address how cultural practices which lower the reproductive fitness of its practitioners can spread. Ironically, they use the example of becoming a college instructor at the cost of adopting the maladaptive practice of delayed marriage and lowered reproduction. Here, if the toil involved with becoming a college instructor is only possible if you also delay marriage, a cultural practice of delaying marriage may spread even if only a few individuals are able to become teachers. This happens if teachers are prestigious and influential – which is often the case. Students, then, may be influenced by instructors and tacitly adopt the belief that delaying marriage is a good belief, because they spent half a year trusting the other information disseminated by their instructor. In the case of delayed marriage, the belief is maladaptive because it hurts reproductive success. If we instead consider the spread of pedagogical beliefs, where reproductive success is (most likely) not impacted, it seems even easier for individuals to quickly adopt beliefs such as “However Professor X teaches ought to be correct.”

While Boyd and Richerson used instructors as one of many possible influential persons from which a maladaptive belief might spread, instructors are clearly germane to the spread of pedagogical beliefs. Students who are particularly influenced by their college instructors are more likely to become the next generation of college instructors than those who were not. If the frequency of classes taught through lecture starts at 100%, then oblique transmission (one generation to the next) of lecture usage from instructors to students creates an evolutionary equilibrium where lecture can remain fixated year after year. Beyond being influenced by their college instructors’ use of lecture, by virtue of becoming college instructors, these individuals were likely very successful in the lecture environment, which may strengthen beliefs in the efficacy of lecture. Through the generational learning dynamics of oblique transmission from teachers to students, one can grasp how the use of lecturing can remain unchanged for 700 years despite evidence supporting its abandonment.

The use of cultural evolutionary dynamics to understand this system can provide several hypotheses. For example, we may expect that students influenced by teachers utilizing active learning pedagogy, who then become instructors themselves, may adopt an active mode of teaching and may be the critical force to a movement towards widespread use of active learning. If this is the case, then it may take a few generations of new college instructors who have experienced gradual change away from lecture to break the historic “lecture equilibrium.” Further, one may expect uncertain novice instructors to socially learn how to teach in the moment, as social learning is often the optimal learning strategy when uncertain (Richerson and Boyd 2008). If frequency based transmission is used, new instructors are likely to become lecturers. If prestige based transmission is used, we may also expect maladaptive pedagogy to spread, as prestige in college departments is more often research based, which may lead to copying someone entrenched in traditional lecture.

Thus, education specialists may have a hard time fighting the strong transmission of lecture, making approaches which take into account the networks of faculty (Andrews, et al. 2016; Quardokus and Henderson 2015) useful, and supported by this cultural evolutionary framework.

Concluding thoughts:

Humans are unique in the complexity of culture we exhibit, and the extreme ability to learn from one another. A large body of research has focused on the evolution of human learning. Under a GCC framework, this work has been extended from formal evolutionary models, to an expanding range of empirical research. This research contributes empirical research guided by GCC. Beyond empirically testing predictions from GCC, this work reveals important social aspects of classroom learning dynamics. In doing so, I hope this work serves as an example, and potential catalyst, for more research aimed to improve educational practice and policy by leveraging work on the evolution of human social learning.

Bibliography

Abraham A, Windmann S, Siefen R, Daum I, Güntürkün O. 2006. Creative thinking in adolescents with attention deficit hyperactivity disorder (ADHD). Child Neuropsychology 12(2):111-123.

Acerbi A, Enquist M, Ghirlanda S. 2009. Cultural evolution and individual development of openness and conservatism. Proceedings of the National Academy of Sciences 106(45):18931-18935.

Andrews T, Conaway E, Zhao J, Dolan E. 2016. Colleagues as Change Agents: How Department Networks and Opinion Leaders Influence Teaching at a Single Research University. CBE-Life Sciences Education 15(2):ar15.

Aoki K, Feldman MW. 2014. Evolution of learning strategies in temporally and spatially variable environments: a review of theory. Theoretical population biology 91:3-19.

Arnett JJ. 2008. The neglected 95%: why American psychology needs to become less American. American Psychologist 63(7):602.

Asch SE. 1951. Effects of group pressure upon the modification and distortion of judgments. Groups, leadership, and men:222-236.

Asghari V, Sanyal S, Buchwaldt S, Paterson A, Jovanovic V, Van Tol HH. 1995. Modulation of intracellular cyclic AMP levels by different human dopamine D4 receptor variants. Journal of neurochemistry 65(3):1157-1165.

Astin AW. 1984. Student involvement: A developmental theory for higher education. Journal of college student personnel 25(4):297-308.

Astin AW. 1993. What matters in college?: Four critical years revisited: Jossey-Bass San Francisco.

Atkisson C, Mesoudi A, O'Brien M. 2012. Adult learners in a novel environment use prestige-biased social learning. Evolutionary psychology: an international journal of evolutionary approaches to psychology and behavior. 10(3):519-537.

Bakermans-Kranenburg MJ, van Ijzendoorn MH. 2011. Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. Development and psychopathology 23(01):39-52.

Bakermans-Kranenburg MJ, van IJzendoorn MH, Caspers K, Philibert R. 2011. DRD4 genotype moderates the impact of parental problems on unresolved loss or trauma. Attachment & human development 13(3):253-269.

Bakermans-Kranenburg MJ, Van IJzendoorn MH, Pijlman FT, Mesman J, Juffer F. 2008. Experimental evidence for differential susceptibility: dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. Developmental psychology 44(1):293.

Bangerter A, Heath C. 2004. The Mozart effect: Tracking the evolution of a scientific legend. British Journal of Social Psychology 43(4):605-623.

Bar-Yosef O. 2002. The upper paleolithic revolution. Annual Review of Anthropology:363-393.

Barton PE, Coley RJ. 2009. Parsing the Achievement Gap II. Policy Information Report. Educational Testing Service.

Bates D, Maechler M, Bolker B. 2012. lme4: Linear mixed-effects models using S4 classes.

Bates D, Sarkar D, Bates MD, Matrix L. 2007. The lme4 package. R package version 2(1).

Becker JR. 1984. The pursuit of graduate education in mathematics: factors that influence women and men. J Educ Equit Leader 4(1):39-53.

Belsky J, Bakermans-Kranenburg MJ, Van IJzendoorn MH. 2007. For better and for worse differential susceptibility to environmental influences. Current Directions in Psychological Science 16(6):300-304.

Benenson JF, Apostoleris NH, Parnass J. 1997. Age and sex differences in dyadic and group interaction. Developmental psychology 33(3):538.

Bentler PM. 1990. Comparative fit indexes in structural models. Psychological bulletin 107(2):238.

Berl RE, Hewlett BS. 2015. Cultural variation in the use of overimitation by the Aka and Ngandu of the Congo Basin. PloS one 10(3):e0120180.

Betzig L. 2012. Means, variances, and ranges in reproductive success: comparative evidence. Evolution and Human Behavior 33(4):309-317.

Birnboim H. 2004. DNA Yield with Oragene• DNA. DNA Genotek.

Bjorklund DF, Bering JM. 2000. The evolved child: Applying evolutionary developmental psychology to modern schooling. Learning and Individual Differences 12(4):347-373.

Board NS. 2014. Science and Engineering Indicators. Arlington VA.

Bohan JS. 1993. Essentialism, constructionism, and feminist psychology. Psychology of women quarterly 17(1):5-21.

Bolker BM, Brooks ME, Clark CJ, Geange SW, Poulsen JR, Stevens MHH, White J-SS. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in ecology & evolution 24(3):127-135.

Bonwell CC, Eison JA. 1991. Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports: ERIC.

Bowles S. 2009. Microeconomics: behavior, institutions, and evolution: Princeton University Press.

Box GE. 1979. All models are wrong, but some are useful. Robustness in Statistics:202.

Boyd R, Richerson PJ. 1976. A Simple Dual inheritance model of the Conflict Between Social and Biological Evolution. Zygon® 11(3):254-262.

Boyd R, Richerson PJ. 1988. Culture and the Evolutionary Process: University of Chicago Press.

Boyd R, Richerson PJ. 1995. Why does culture increase human adaptability? Ethology and Sociobiology 16(2):125-143.

Boyd R, Richerson PJ, Henrich J. 2011. The cultural niche: Why social learning is essential for human adaptation. Proceedings of the National Academy of Sciences 108(Supplement 2):10918-10925.

Brashears ME, Hoagland E, Quintane E. 2016. Sex and network recall accuracy. Social Networks 44:74-84.

Browne MW, Cudeck R, Bollen KA, Long JS. 1993. Alternative ways of assessing model fit. Sage focus editions 154:136-136.

Brubacher JS, Rudy W. 1997. Higher education in transition: A history of American colleges and universities: Transaction Publishers.

Buss DM, Schmitt DP. 1993. Sexual strategies theory: an evolutionary perspective on human mating. Psychological review 100(2):204.

Butts CT. 2008. network: a Package for Managing Relational Data in R. Journal of Statistical Software 24(2):1-36.

Carli LL, Alawa L, Lee Y, Zhao B, Kim E. 2016. Stereotypes About Gender and Science Women≠ Scientists. Psychology of Women Quarterly:0361684315622645.

Cashdan E. 1998. Are men more competitive than women? British Journal of Social Psychology 37(2):213-229.

Catalyst. 2016. Women CEOs of the S&P 500. New York: Catalyst.

Cavalli-Sforza LL, Feldman MW. 1981a. Cultural transmission and evolution: a quantitative approach: Princeton University Press.

Cavalli-Sforza LL, Feldman MW, Chen K-H, Dornbusch SM. 1982. Theory and observation in cultural transmission. Science 218(4567):19-27.

Cavalli-Sforza LLL, Feldman MW. 1981b. Cultural Transmission and Evolution: A Quantitative Approach. (Mpb-16): Princeton University Press.

Cech E, Rubineau B, Silbey S, Seron C. 2011. Professional role confidence and gendered persistence in engineering. American Sociological Review 76(5):641-666.

Ceci SJ, Williams WM. 2011. Understanding current causes of women's underrepresentation in science. Proceedings of the National Academy of Sciences 108(8):3157-3162.

Centola D. 2011. An experimental study of homophily in the adoption of health behavior. Science 334(6060):1269-1272.

Chabris CF, Lee JJ, Benjamin DJ, Beauchamp JP, Glaeser EL, Borst G, Pinker S, Laibson DI. 2013. Why it is hard to find genes associated with social science traits: Theoretical and empirical considerations. American journal of public health 103(S1):S152-S166.

Chang FM, Kidd JR, Livak KJ, Pakstis AJ, Kidd KK. 1996. The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. Hum Genet 98(1):91-101.

Chang L, Mak MC, Li T, Wu BP, Chen BB, Lu HJ. 2011. Cultural adaptations to environmental variability: An evolutionary account of East–West differences. Educational Psychology Review 23(1):99-129.

Chen C, Burton M, Greenberger E, Dmitrieva J. 1999. Population Migration and the Variation of Dopamine D4 Receptor (DRD4) Allele Frequencies Around the Globe. Evolution and Human Behavior 20(5):309-324.

Chi MT, Wylie R. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist 49(4):219-243.

Chudek M, Heller S, Birch S, Henrich J. 2012. Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. Evolution and Human Behavior 33(1):46-56.

Cohen E, Haun D. 2013. The development of tag-based cooperation via a socially acquired trait. Evolution and Human Behavior 34(3):230-235.

Cohen GL, Steele CM, Ross LD. 1999. The mentor’s dilemma: Providing critical feedback across the racial divide. Personality and Social Psychology Bulletin 25(10):1302-1318.

Correa-Chávez M, Rogoff B. 2009. Children’s attention to interactions directed to others: Guatemalan mayan and european american patterns. Developmental Psychology 45(3):630.

Corriveau K, Harris PL. 2009. Choosing your informant: Weighing familiarity and recent accuracy. Developmental science 12(3):426-437.

Cowlishaw G, Dunbar RI. 1991. Dominance rank and mating success in male primates. Animal Behaviour 41(6):1045-1056.

Cross JLM, Lin N. 2008. Access to social capital and status attainment in the United States: Racial/ethnic and gender differences. Social capital: An international research program:364-379.

Csibra G, Gergely G. 2009. Natural pedagogy. Trends in cognitive sciences 13(4):148-153.

Dai DY, Gerbino KA, Daley MJ. 2011. Inquiry-based learning in China: Do teachers practice what they preach, and why? Frontiers of Education in China 6(1):139-157.

Darwin C. 1872. The origin of species: Lulu. com.

Dawkins R. 2006. The selfish gene: Oxford university press.

Denson N, Chang MJ. 2009. Racial diversity matters: The impact of diversity-related student engagement and institutional context. American Educational Research Journal 46(2):322-353.

Denson N, Zhang S. 2010. The impact of student experiences with diversity on developing graduate attributes. Studies in Higher Education 35(5):529-543.

Derex M, Beugin M-P, Godelle B, Raymond M. 2013. Experimental evidence for the influence of group size on cultural complexity. Nature 503(7476):389-391.

DiMaio S, Grizenko N, Joober R. 2003. Dopamine genes and attention-deficit hyperactivity disorder: a review. J Psychiatry Neurosci 28(1):27-38.

Ding YC, Chi HC, Grady DL, Morishima A, Kidd JR, Kidd KK, Flodman P, Spence MA, Schuck S, Swanson JM and others. 2002. Evidence of positive selection acting at the human dopamine receptor D4 gene locus. Proc Natl Acad Sci U S A 99(1):309-14.

Dreber A, Apicella CL, Eisenberg DTA, Garcia JR, Zamore RS, Lum JK, Campbell B. 2009. The 7R polymorphism in the dopamine receptor D4 gene (DRD4) is associated with financial risk taking in men. Evolution and Human Behavior 30(2):85-92.

Dreber A, Rand DG, Wernerfelt N, Garcia JR, Vilar MG, Lum JK, Zeckhauser R. 2011. Dopamine and risk choices in different domains: findings among serious tournament bridge players. Journal of Risk and Uncertainty 43(1):19-38.

Dugatkin LA, Godin J-GJ. 1993. Female mate copying in the guppy {Poecilia reticulata): age-dependent effects. Behavioral Ecology 4(4):289-292.

Eagly AH, Karau SJ. 2002. Role congruity theory of prejudice toward female leaders. Psychological review 109(3):573.

Eddy SL, Brownell SE. 2015. Beneath the numbers: A review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. Physics Review: Special Topics Physics Education Research: Gender in Physics.

Eddy SL, Brownell SE, Thummaphan P, Lan M-C, Wenderoth MP. 2015. Caution, Student Experience May Vary: Social Identities Impact a Student’s Experience in Peer Discussions. CBE-Life Sciences Education 14(4).

Eddy SL, Brownell SE, Wenderoth MP. 2014. Gender Gaps in Achievement and Participation in Multiple Introductory Biology Classrooms. CBE-Life Sciences Education 13(3):478-492.

Efferson C, Lalive R, Richerson PJ, McElreath R, Lubell M. 2008. Conformists and mavericks: the empirics of frequency-dependent cultural transmission. Evolution and Human Behavior 29(1):56-64.

Efferson C, Richerson PJ, McElreath R, Lubell M, Edsten E, Waring TM, Paciotti B, Baum W. 2007. Learning, productivity, and noise: an experimental study of cultural transmission on the Bolivian Altiplano. Evolution and Human Behavior 28(1):11-17.

Eisenberg D, Campbell B, Gray P, Sorenson M. 2008. Dopamine receptor genetic polymorphisms and body composition in undernourished pastoralists: An exploration of nutrition indices among nomadic and recently settled Ariaal men of northern Kenya. BMC Evolutionary Biology 8(1):173.

Eisenberg DT, Apicella CL, Campbell BC, Dreber A, Garcia JR, Lum JK. 2009. Assortative human pair-bonding for partner ancestry and allelic variation of the dopamine receptor D4 (DRD4) gene. Social cognitive and affective neuroscience:nsp026.

Eisenberg DT, MacKillop J, Modi M, Beauchemin J, Dang D, Lisman SA, Lum JK, Wilson DS. 2007. Examining impulsivity as an endophenotype using a behavioral approach: a DRD2 TaqI A and DRD4 48-bp VNTR association study. Behavioral and Brain Functions 3(1):1.

Elia J, Sackett J, Turner T, Schardt M, Tang S-C, Kurtz N, Dunfey M, McFarlane N, Susi A, Danish D and others. 2012. Attention-Deficit/Hyperactivity Disorder Genomics: Update for Clinicians. Current Psychiatry Reports 14(5):579-589.

Ellis BJ, Bjorklund DF. 2012. Beyond mental health: an evolutionary analysis of development under risky and supportive environmental conditions: an introduction to the special section. Developmental psychology 48(3):591.

Embretson SE, Reise SP. 2013. Item response theory: Psychology Press.

Enquist M, Eriksson K, Ghirlanda S. 2007. Critical Social Learning: A Solution to Rogers's Paradox of Nonadaptive Culture. American Anthropologist 109(4):727-734.

Faraone SV, Biederman J, Weiffenbach B, Keith T, Chu MP, Weaver A, Spencer TJ, Wilens TE, Frazier J, Cleves M. 2014. Dopamine D4 gene 7-repeat allele and attention deficit hyperactivity disorder.

Fawcett CA, Markson L. 2010. Similarity predicts liking in 3-year-old children. Journal of experimental child psychology 105(4):345-358.

Feldman MW, Aoki K, Kumm J. 1996. Individual versus social learning: evolutionary analysis in a fluctuating environment. Anthropological Science 104(3):209-231.

Fox J, Fox MJ. 2010. Package ‘polycor’.

Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014. Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences 111(23):8410-8415.

Galef BG. 2009. Strategies for social learning: testing predictions from formal theory. Advances in the Study of Behavior 39:117-151.

Geary DC. 1995. Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. American Psychologist 50(1):24.

Geary DC. 2000. Evolution and proximate expression of human paternal investment. Psychological bulletin 126(1):55.

Geary DC. 2007. An evolutionary perspective on learning disability in mathematics. Developmental neuropsychology 32(1):471-519.

Gerbault P, Liebert A, Itan Y, Powell A, Currat M, Burger J, Swallow DM, Thomas MG. 2011. Evolution of lactase persistence: an example of human niche construction. Philosophical Transactions of the Royal Society B: Biological Sciences 366(1566):863-877.

Ghirlanda S, Enquist M, Nakamaru M. 2006. Cultural evolution develops its own rules: the rise of conservatism and persuasion. Current Anthropology 47(6):1027-1034.

Gillen A, Tanenbaum C. 2014. Exploring Gender Imbalance Among STEM Doctoral Degree Recipients. Washington, DC: American Institutes for Research.

Gizer IR, Ficks C, Waldman ID. 2009. Candidate gene studies of ADHD: a meta-analytic review. Human genetics 126(1):51-90.

Grady DL, Chi HC, Ding YC, Smith M, Wang E, Schuck S, Flodman P, Spence MA, Swanson JM, Moyzis RK. 2003. High prevalence of rare dopamine receptor D4 alleles in children diagnosed with attention-deficit hyperactivity disorder. Mol Psychiatry 8(5):536-45.

Grafen A. 1984. Natural selection, kin selection and group selection. Behavioural ecology: an evolutionary approach 2:62-84.

Gray P. 2011. The Decline of Play and the Rise of Psychopathology in Children and Adolescents. American Journal of Play 3(4):443-463.

Gray RD, Atkinson QD. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. Nature 426(6965):435-439.

Gray RD, Jordan FM. 2000. Language trees support the express-train sequence of Austronesian expansion. Nature 405(6790):1052-1055.

Grunspan DZ, Wiggins BL, Goodreau SM. 2014. Understanding classrooms through social network analysis: A primer for social network analysis in education research. CBE-Life Sciences Education 13(2):167-178.

Guo SW, Thompson E. 1992. A Monte Carlo method for combined segregation and linkage analysis. American journal of human genetics 51(5):1111.

Gurin P, Dey E, Hurtado S, Gurin G. 2002. Diversity and higher education: Theory and impact on educational outcomes. Harvard educational review 72(3):330-367.

Gurven M, Hill K. 2009. Why do men hunt? Current Anthropology 50(1):51-74.

Hair JF. 2010. Multivariate data analysis.

Halpern DF, Benbow CP, Geary DC, Gur RC, Hyde JS, Gernsbacher MA. 2007. The science of sex differences in science and mathematics. Psychological science in the public interest 8(1):1-51.

Hamarman S, Fossella J, Ulger C, Brimacombe M, Dermody J. 2004. Dopamine receptor 4 (DRD4) 7-repeat allele predicts methylphenidate dose response in children with attention deficit hyperactivity disorder: a pharmacogenetic study. Journal of Child & Adolescent Psychopharmacology 14(4):564-574.

Hamer DH. 2000. Beware the chopsticks gene. Molecular psychiatry 5(1):11-13.

Hamilton WD. 1963. The evolution of altruistic behavior. The American Naturalist 97(896):354-356.

Handcock MS, Hunter DR, Butts CT, Goodreau SM, Morris M. 2008. statnet: Software tools for the representation, visualization, analysis and simulation of network data. Journal of statistical software 24(1):1548.

Haun D, Over H. 2015. Like me: a homophily-based account of human culture. Epistemological Dimensions of Evolutionary Psychology: Springer. p 117-130.

Hayton JC, Allen DG, Scarpello V. 2004. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. Organizational research methods 7(2):191-205.

Heath C, Bell C, Sternberg E. 2001. Emotional selection in memes: the case of urban legends. Journal of personality and social psychology 81(6):1028.

Henn BM, Cavalli-Sforza LL, Feldman MW. 2012. The great human expansion. Proceedings of the National Academy of Sciences 109(44):17758-17764.

Henrich J, Broesch J. 2011. On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. Philos Trans R Soc Lond B Biol Sci 366(1567):1139-48.

Henrich J, Gil-White FJ. 2001. The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evolution and Human Behavior 22(3):165-196.

Henrich J, Heine SJ, Norenzayan A. 2010. The weirdest people in the world? Behavioral and brain sciences 33(2-3):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 of London B: Biological Sciences 277(1701):3715-3724.

Henrich J, McElreath R. 2003. The evolution of cultural evolution. Evolutionary Anthropology: Issues, News, and Reviews 12(3):123-135.

Herrmann E, Call J, Hernàndez-Lloreda MV, Hare B, Tomasello M. 2007. Humans Have Evolved Specialized Skills of Social Cognition: The Cultural Intelligence Hypothesis. Science 317(5843):1360-1366.

Hess NH, Hagen EH. 2006. Sex differences in indirect aggression: Psychological evidence from young adults. Evolution and Human Behavior 27(3):231-245.

Hewlett BS, Cavalli-Sforza LL. 1986. Cultural transmission among Aka pygmies. American Anthropologist 88(4):922-934.

Heyes C. 2016. Blackboxing: social learning strategies and cultural evolution. Phil. Trans. R. Soc. B 371(1693):20150369.

Hofstede G, Bond MH. 1988. The Confucius connection: From cultural roots to economic growth. Organizational dynamics 16(4):5-21.

Holdren JP, Lander E. 2012. Report to the President—Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. President’s Council of Advisors on Science and Technology.

Hopcroft RL. 2006. Sex, status, and reproductive success in the contemporary United States. Evolution and Human Behavior 27(2):104-120.

Hunter DR, Handcock MS, Butts CT, Goodreau SM, Morris M. 2008. ergm: A package to fit, simulate and diagnose exponential-family models for networks. Journal of statistical software 24(3):nihpa54860.

Ingram CJ, Mulcare CA, Itan Y, Thomas MG, Swallow DM. 2009. Lactose digestion and the evolutionary genetics of lactase persistence. Human genetics 124(6):579-591.

Irons W. 1979. 1997, Cultural and biological success. L. Betzig, éd.

Isbell LA, Young TP, Harcourt AH. 2012. Stag parties linger: continued gender bias in a female-rich scientific discipline. PLoS One 7(11):e49682.

Jargowsky PA, El Komi M. Before or After the Bell? School Context and Neighborhood Effects on Student Achievement.

Jiang Y, Bachner-Melman R, Chew SH, Ebstein RP. 2015. Dopamine D4 receptor gene and religious affiliation correlate with dictator game altruism in males and not females: evidence for gender-sensitive gene× culture interaction. Frontiers in neuroscience 9.

Junod REV, DuPaul GJ, Jitendra AK, Volpe RJ, Cleary KS. 2006. Classroom observations of students with and without ADHD: Differences across types of engagement. Journal of School Psychology 44(2):87-104.

Kaplan H, Gurven M, Hill K, Hurtado AM. 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.

Katsnelson E, Lotem A, Feldman MW. 2014. ASSORTATIVE SOCIAL LEARNING AND ITS IMPLICATIONS FOR HUMAN (AND ANIMAL?) SOCIETIES. Evolution 68(7):1894-1906.

Kena G, Musu-Gillette L, Robinson J, Wang X, Rathbun A, Zhang J, Wilkinson-Flicker S, Barmer A, Velez EDV. 2015. The Condition of Education 2015. NCES 2015-144. National Center for Education Statistics.

Kidd KK, Pakstis AJ, Yun L. 2014. An historical perspective on "The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus". Human genetics 133(4):431-433.

Kinzler KD, Dupoux E, Spelke ES. 2007. The native language of social cognition. Proceedings of the National Academy of Sciences 104(30):12577-12580.

Kinzler KD, Shutts K, Spelke ES. 2012. Language-based social preferences among children in South Africa. Language learning and development 8(3):215-232.

Kitayama S, King A, Hsu M, Liberzon I, Yoon C. 2016. Dopamine-system genes and cultural acquisition: the norm sensitivity hypothesis. Current Opinion in Psychology 8:167-174.

Kitayama S, King A, Yoon C, Tompson S, Huff S, Liberzon I. 2014. The dopamine D4 receptor gene (DRD4) moderates cultural difference in independent versus interdependent social orientation. Psychological science 25(6):1169-1177.

Kline MA. 2015. How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. Behavioral and Brain Sciences 38:e31.

Kluger A, Siegfried Z, Ebstein R. 2002. A meta-analysis of the association between DRD4 polymorphism and novelty seeking. Molecular psychiatry 7(7):712-717.

Krivitsky P, Handcock M. 2015. tergm: Fit, Simulate and Diagnose Models for Network Evolution Based on Exponential-Family Random Graph Models. The Statnet Project (http://www. statnet. org). R package version 3(1).

Krivitsky PN, Goodreau SM. 2015. STERGM-Separable Temporal ERGMs for modeling discrete relational dynamics with statnet.

Krivitsky PN, Handcock MS. 2014. A separable model for dynamic networks. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76(1):29-46.

Kuznetsova A, Brockhoff PB, Christensen RHB. 2013. lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2(6).

Laird TFN. 2005. College students' experiences with diversity and their effects on academic self-confidence, social agency, and disposition toward critical thinking. Research in Higher Education 46(4):365-387.

Laland KN. 2004. Social learning strategies. Animal Learning & Behavior 32(1):4-14.

Laland KN, Odling-Smee J, Myles S. 2010. How culture shaped the human genome: bringing genetics and the human sciences together. Nature Reviews Genetics 11(2):137-148.

Laland KN, Odling-Smee J, Feldman MW. 2001. Cultural niche construction and human evolution. Journal of evolutionary biology 14(1):22-33.

Lancy DF. 2010. Learning'from nobody': The limited role of teaching in folk models of children's development. Childhood in the Past 3(1):79-106.

Lazarsfeld PF, Merton RK. 1954. Friendship as a social process: A substantive and methodological analysis. Freedom and control in modern society 18(1):18-66.

Leslie S-J, Cimpian A, Meyer M, Freeland E. 2015. Expectations of brilliance underlie gender distributions across academic disciplines. Science 347(6219):262-265.

Leung PW, Lee C, Hung S, Ho T, Tang C, Kwong S, Leung S, Yuen S, Lieh-Mak F, Oosterlaan J. 2005a. Dopamine receptor D4 (DRD4) gene in Han Chinese children with attention-deficit/hyperactivity disorder (ADHD): Increased prevalence of the 2-repeat allele. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 133(1):54-56.

Leung PW, Lee CC, Hung SF, Ho TP, Tang CP, Kwong SL, Leung SY, Yuen ST, Lieh-Mak F, Oosterlaan J and others. 2005b. Dopamine receptor D4 (DRD4) gene in Han Chinese children with attention-deficit/hyperactivity disorder (ADHD): increased prevalence of the 2-repeat allele. Am J Med Genet B Neuropsychiatr Genet 133B(1):54-6.

Lipsey MW, Puzio K, Yun C, Hebert MA, Steinka-Fry K, Cole MW, Roberts M, Anthony KS, Busick MD. 2012. Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. National Center for Special Education Research.

López A, Correa-Chávez M, Rogoff B, Gutiérrez K. 2010. Attention to instruction directed to another by US Mexican-heritage children of varying cultural backgrounds. Developmental Psychology 46(3):593.

Low BS. 1992. Sex, coalitions, and politics in preindustrial societies. Politics and the Life Sciences:63-80.

Lusher J, Chandler C, Ball D. 2001. Dopamine D4 receptor gene (DRD4) is associated with Novelty Seeking (NS) and substance abuse: the saga continues. Molecular psychiatry.

Mace R, Holden CJ. 2005. A phylogenetic approach to cultural evolution. Trends in ecology & evolution 20(3):116-121.

Mace R, Pagel M, Bowen JR, Otterbein KF, Ridley M, Schweizer T, Voland E. 1994. The comparative method in anthropology [and comments and reply]. Current Anthropology 35(5):549-564.

Manson JH, Wrangham RW, Boone JL, Chapais B, Dunbar R, Ember CR, Irons W, Marchant L, McGrew W, Nishida T. 1991. Intergroup Aggression in Chimpanzees and Humans [and Comments and Replies]. Current anthropology 32(4):369-390.

Marianne B. 2011. New perspectives on gender. Handbook of labor economics 4:1543-1590.

Markus HR, Kitayama S. 1991. Culture and the self: Implications for cognition, emotion, and motivation. Psychological review 98(2):224.

Martin SN, Milne C, Scantlebury K. 2006. Eye-rollers, risk-takers, and turn sharks: Target students in a professional science education program. Journal of research in science teaching 43(8):819-851.

Matthews LJ, Butler PM. 2011. Novelty-seeking DRD4 polymorphisms are associated with human migration distance out-of-Africa after controlling for neutral population gene structure. Am J Phys Anthropol 145(3):382-9.

Mazur E. 1997. Peer instruction: Upper Saddle River, NJ: Prentice Hall.

McElreath R, Lubell M, Richerson PJ, Waring TM, Baum W, Edsten E, Efferson C, Paciotti B. 2005. Applying evolutionary models to the laboratory study of social learning. Evolution and Human Behavior 26(6):483-508.

McGeary J. 2009. The DRD4 exon 3 VNTR polymorphism and addiction-related phenotypes: a review. Pharmacology Biochemistry and Behavior 93(3):222-229.

McPherson M, Smith-Lovin L, Cook JM. 2001. Birds of a feather: Homophily in social networks. Annual review of sociology:415-444.

Mesoudi A. 2011a. Cultural evolution: how Darwinian theory can explain human culture and synthesize the social sciences: University of Chicago Press.

Mesoudi A. 2011b. An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. Evolution and Human Behavior 32(5):334-342.

Mesoudi A. 2016. Cultural evolution: Integrating psychology, evolution and culture. Current Opinion in Psychology 7:17-22.

Mesoudi A, Chang L, Dall SR, Thornton A. 2016. The Evolution of Individual and Cultural Variation in Social Learning. Trends in ecology & evolution 31(3):215-225.

Mesoudi A, Chang L, Murray K, Lu HJ. 2015a. 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(1798):20142209.

Mesoudi A, Chang L, Murray K, Lu HJ. 2015b. 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 of London B: Biological Sciences 282(1798):20142209.

Mesoudi A, Whiten A, Laland KN, Harrison R. 2004. Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the perspective of The Origin of Species. Evolution 58(1):1-11.

Milkman KL, Akinola M, Chugh D. 2014. What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. A Field Experiment Exploring How Pay and Representation Differentially Shape Bias on the Pathway into Organizations (July 11, 2014).

Ming-li L. 2012. The problems of senior three students'"Mute-Crux" in English learning in China. Sino-US English Teaching 9(5):1139-1144.

Molleman L, Van den Berg P, Weissing FJ. 2014. Consistent individual differences in human social learning strategies. Nature communications 5.

Morgan LH. 1877. Ancient society; or, researches in the lines of human progress from savagery, through barbarism to civilization: H. Holt.

Morgan T, Rendell L, Ehn M, Hoppitt W, Laland K. 2012. The evolutionary basis of human social learning. Proceedings of the Royal Society of London B: Biological Sciences 279(1729):653-662.

Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. 2012. Science faculty's subtle gender biases favor male students. Proceedings of the National Academy of Sciences 109(41):16474-16479.

Motz M, Pääbo S, Kilger C. 2000. Improved cycle sequencing of GC-rich templates by a combination of nucleotide analogs. Biotechniques 29(2):268-270.

Munafò MR, Yalcin B, Willis-Owen SA, Flint J. 2008. Association of the Dopamine D4 Receptor (DRD4) Gene and Approach-Related Personality Traits: Meta-Analysis and New Data. Biological Psychiatry 63(2):197-206.

Naka I, Nishida N, Ohashi J. 2011. No evidence for strong recent positive selection favoring the 7 repeat allele of VNTR in the DRD4 gene. PloS one 6(8):e24410.

National Science Foundation NCfSaES. 2013. Women, Minorities, and Persons with Disabilities in Science and Engineering: 2013. Special Report NSF 13-304.

Nisbett R. 2010. The Geography of Thought: How Asians and Westerners Think Differently... and: Simon and Schuster.

Nosek BA, Banaji M, Greenwald AG. 2002a. Harvesting implicit group attitudes and beliefs from a demonstration web site. Group Dynamics: Theory, Research, and Practice 6(1):101.

Nosek BA, Banaji MR, Greenwald AG. 2002b. Math= male, me= female, therefore math≠ me. Journal of personality and social psychology 83(1):44.

Nosek BA, Smyth FL. 2011. Implicit social cognitions predict sex differences in math engagement and achievement. American Educational Research Journal 48(5):1125-1156.

Nosek BA, Smyth FL, Sriram N, Lindner NM, Devos T, Ayala A, Bar-Anan Y, Bergh R, Cai H, Gonsalkorale K. 2009. National differences in gender–science stereotypes predict national sex differences in science and math achievement. Proceedings of the National Academy of Sciences 106(26):10593-10597.

O'Brien LT, Blodorn A, Adams G, Garcia DM, Hammer E. 2015. Ethnic variation in gender-STEM stereotypes and STEM participation: An intersectional approach. Cultural Diversity and Ethnic Minority Psychology 21(2):169.

Oak JN, Oldenhof J, Van Tol HH. 2000. The dopamine D 4 receptor: one decade of research. European journal of pharmacology 405(1):303-327.

Odling-Smee FJ, Laland KN, Feldman MW. 2003. Niche construction: the neglected process in evolution: Princeton University Press.

Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA, Turley P, Chen G-B, Emilsson V, Meddens SFW. 2016. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 533(7604):539-542.

Olson S, Riordan DG. 2012. Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. Executive Office of the President.

Omark DR, Omark M, Edelman M. 1975. Formation of dominance hierarchies in young children. Psychological anthropology:289-315.

Paas F, Sweller J. 2012. An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. Educational Psychology Review 24(1):27-45.

Pike TW, Laland KN. 2010. Conformist learning in nine-spined sticklebacks' foraging decisions. Biology letters:rsbl20091014.

Posth C, Renaud G, Mittnik A, Drucker DG, Rougier H, Cupillard C, Valentin F, Thevenet C, Furtwängler A, Wißing C. 2016. Pleistocene mitochondrial genomes suggest a single major dispersal of Non-Africans and a Late Glacial population turnover in Europe. Current Biology 26(6):827-833.

Quardokus K, Henderson C. 2015. Promoting instructional change: using social network analysis to understand the informal structure of academic departments. Higher Education 70(3):315-335.

R Core Team. 2013. R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing.

Reist C, Ozdemir V, Wang E, Hashemzadeh M, Mee S, Moyzis R. 2007. Novelty seeking and the dopamine D4 receptor gene (DRD4) revisited in Asians: Haplotype characterization and relevance of the 2-repeat allele. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 144(4):453-457.

Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, Fogarty L, Ghirlanda S, Lillicrap T, Laland KN. 2010. Why Copy Others? Insights from the Social Learning Strategies Tournament. Science 328(5975):208-213.

Rendell L, Fogarty L, Hoppitt WJ, Morgan TJ, Webster MM, Laland KN. 2011. Cognitive culture: theoretical and empirical insights into social learning strategies. Trends Cogn Sci 15(2):68-76.

Richerson P. 2003. Shared norms and the evolution of ethnic markers. Current anthropology 44(1):122-130.

Richerson PJ, Boyd R. 2008. Not By Genes Alone: How Culture Transformed Human Evolution: University of Chicago Press.

Richerson PJ, Boyd R, Henrich J. 2010. Gene-culture coevolution in the age of genomics. Proceedings of the National Academy of Sciences 107(Supplement 2):8985-8992.

Ridgeway CL, Backor K, Li YE, Tinkler JE, Erickson KG. 2009. How easily does a social difference become a status distinction? Gender matters. American Sociological Review 74(1):44-62.

Rizopoulos D. 2006. ltm: An R package for latent variable modeling and item response theory analyses. Journal of statistical software 17(5):1-25.

Robins G, Pattison P, Kalish Y, Lusher D. 2007. An introduction to exponential random graph (p*) models for social networks. Social networks 29(2):173-191.

Robnett RD, Leaper C. 2013. Friendship Groups, Personal Motivation, and Gender in Relation to High School Students' STEM Career Interest. Journal of Research on Adolescence 23(4):652-664.

Rogers EM. 2010. Diffusion of innovations: Simon and Schuster.

Rose AJ, Rudolph KD. 2006. A review of sex differences in peer relationship processes: potential trade-offs for the emotional and behavioral development of girls and boys. Psychological bulletin 132(1):98.

Rosenthal L, London B, Levy SR, Lobel M. 2011. The roles of perceived identity compatibility and social support for women in a single-sex STEM program at a co-educational university. Sex Roles 65(9-10):725-736.

Rothbart M, Fulero S, Jensen C, Howard J, Birrell P. 1978. From individual to group impressions: Availability heuristics in stereotype formation. Journal of Experimental Social Psychology 14(3):237-255.

Rudman LA, Glick P. 2001. Prescriptive gender stereotypes and backlash toward agentic women. Journal of social issues 57(4):743-762.

Sadalla EK, Kenrick DT, Vershure B. 1987. Dominance and heterosexual attraction. Journal of Personality and Social Psychology 52(4):730.

Sasaki JY, Kim HS, Mojaverian T, Kelley LD, Park IY, Janušonis S. 2011. Religion priming differentially increases prosocial behavior among variants of the dopamine D4 receptor (DRD4) gene. Social Cognitive and Affective Neuroscience:nsr089.

Satterthwaite FE. 1946. An approximate distribution of estimates of variance components. Biometrics bulletin 2(6):110-114.

Savin-Williams RC. 1979. Dominance hierarchies in groups of early adolescents. Child development:923-935.

Schlag KH. 1998. Why imitate, and if so, how?: A boundedly rational approach to multi-armed bandits. Journal of Economic Theory 78(1):130-156.

Schmader T, Johns M, Barquissau M. 2004. The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. Sex Roles 50(11-12):835-850.

Sell A, Tooby J, Cosmides L. 2009. Formidability and the logic of human anger. Proceedings of the National Academy of Sciences 106(35):15073-15078.

University of Washington International Student Services (2014). All International Students.

International Students Profile 2014. Retrieved from http://iss.washington.edu/sites/default/files/u65/Students%20Profile%202014_0.pdf.

Settles IH, Cortina LM, Malley J, Stewart AJ. 2006. The climate for women in academic science: The good, the bad, and the changeable. Psychology of Women Quarterly 30(1):47-58.

Seymour E, Hewitt NM, Friend CM. 1997. Talking about leaving: Why undergraduates leave the sciences: Westview Press Boulder, CO.

Sheltzer JM, Smith JC. 2014. Elite male faculty in the life sciences employ fewer women. Proceedings of the National Academy of Sciences 111(28):10107-10112.

Sherif M. 1936. The psychology of social norms.

Smith EA. 2004. Why do good hunters have higher reproductive success? Human Nature 15(4):343-364.

Smith JM. 1978. Optimization theory in evolution. Annual Review of Ecology and Systematics 9:31-56.

Smuts B. 1995. The evolutionary origins of patriarchy. Human Nature 6(1):1-32.

Soltis J, Boyd R, Richerson PJ. 1995. Can group-functional behaviors evolve by cultural group selection?: An empirical test. Current Anthropology 36(3):473-494.

Sosis R. 2000. Religion and intragroup cooperation: Preliminary results of a comparative analysis of utopian communities. Cross-Cultural Research 34(1):70-87.

Spencer H. 1857. The origin and function of music. Essays, Scientific, Political, and Speculative 2.

Steiger JH, Lind JC. Statistically based tests for the number of common factors; 1980.

Sterelny K. 2012. The evolved apprentice: MIT press.

Strathern M, Ortner SB, Whitehead H. 1981. Sexual Meanings: the cultural construction of gender and sexuality. Sexual Meanings: The Cultural Construction of Gender and Sexuality.

Swim JK, Hyers LL, Cohen LL, Ferguson MJ. 2001. Everyday sexism: Evidence for its incidence, nature, and psychological impact from three daily diary studies. Journal of Social Issues 57(1):31-53.

Tanner KD. 2013. Structure matters: twenty-one teaching strategies to promote student engagement and cultivate classroom equity. CBE-Life Sciences Education 12(3):322-331.

Terenzini PT, Pascarella ET. 1991. Twenty years of research on college students: Lessons for future research. Research in Higher Education 32(1):83-92.

Thompson M, Sekaquaptewa D. 2002. When being different is detrimental: Solo status and the performance of women and racial minorities. Analyses of Social Issues and Public Policy 2(1):183-203.

Tinto V. 1997. Classrooms as communities: Exploring the educational character of student persistence. Journal of higher education:599-623.

Tishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M. 2007. Convergent adaptation of human lactase persistence in Africa and Europe. Nature genetics 39(1):31-40.

Tomasello M, Kruger AC, Ratner HH. 1993. Cultural learning. Behavioral and Brain Sciences 16(03):495-511.

Trivers RL. 1972. Parental Investment and Sexual Selection.

Tweed RG, Lehman DR. 2002. Learning considered within a cultural context: Confucian and Socratic approaches. American Psychologist 57(2):89.

Tylor EB. 1871. Primitive culture: researches into the development of mythology, philosophy, religion, art, and custom: Murray.

Usher EL, Pajares F. 2008. Sources of self-efficacy in school: Critical review of the literature and future directions. Review of Educational Research 78(4):751-796.

Van Dyken JD, Müller MJ, Mack KM, Desai MM. 2013. Spatial population expansion promotes the evolution of cooperation in an experimental prisoner’s dilemma. Current Biology 23(10):919-923.

Van Tol HH, Wu CM, Guan H-C, Ohara K, Bunzow JR, Civelli O, Kennedy J, Seeman P, Niznik HB, Jovanovic V. 1992a. Multiple dopamine D4 receptor variants in the human population. Nature.

Van Tol HH, Wu CM, Guan HC, Ohara K, Bunzow JR, Civelli O, Kennedy J, Seeman P, Niznik HB, Jovanovic V. 1992b. Multiple dopamine D4 receptor variants in the human population. Nature 358(6382):149-52.

Volk AA, Camilleri JA, Dane AV, Marini ZA. 2012. Is adolescent bullying an evolutionary adaptation? Aggressive Behavior 38(3):222-238.

Von Rueden C, Gurven M, Kaplan H. 2008. The multiple dimensions of male social status in an Amazonian society. Evolution and Human Behavior 29(6):402-415.

Wang E, Ding Y-C, Flodman P, Kidd J, Kidd K, Grady D, Ryder O, Spence M, Swanson J, Moyzis R. 2004. The genetic architecture of selection at the human dopamine receptor D4 (DRD4) gene locus. American journal of human genetics 74(5):931.

Watkins CD, Jones BC, DeBruine LM. 2010. Individual differences in dominance perception: Dominant men are less sensitive to facial cues of male dominance. Personality and Individual Differences 49(8):967-971.

White HA, Shah P. 2006. Uninhibited imaginations: creativity in adults with attention-deficit/hyperactivity disorder. Personality and Individual Differences 40(6):1121-1131.

White HA, Shah P. 2011. Creative style and achievement in adults with attention-deficit/hyperactivity disorder. Personality and Individual Differences 50(5):673-677.

Whiten A, Hinde RA, Laland KN, Stringer CB. 2011. Culture evolves. Philosophical Transactions of the Royal Society B: Biological Sciences 366(1567):938-948.

Wiens DJ, Depping DJ, Wallerich SR, Van Laar ES, Juhl AL. 2003. Gender Matters: Factors Influencing Biology Interest among Undergraduate Majors. Journal of College Science Teaching 33(1):32-36.

Williams J, Taylor E. 2006. The evolution of hyperactivity, impulsivity and cognitive diversity. J R Soc Interface 3(8):399-413.

Wilson DS. 1975. A theory of group selection. Proceedings of the national academy of sciences 72(1):143-146.

Wilson DS, Kauffman Jr RA, Purdy MS. 2011. A program for at-risk high school students informed by evolutionary science. PLoS One 6(11):e27826.

Wilson DS, Sober E. 1994. Reintroducing group selection to the human behavioral sciences. Behavioral and Brain Sciences 17(04):585-608.

Wood W, Eagly AH. 2002. A cross-cultural analysis of the behavior of women and men: implications for the origins of sex differences. Psychological bulletin 128(5):699.

Yaveroglu IS, Donthu N. 2002. Cultural influences on the diffusion of new products. Journal of International Consumer Marketing 14(4):49-63.

Young DM, Rudman LA, Buettner HM, McLean MC. 2013. The influence of female role models on women's implicit science cognitions. Psychology of Women Quarterly 37(3):283-292.

Zeldin AL, Pajares F. 2000. Against the odds: Self-efficacy beliefs of women in mathematical, scientific, and technological careers. American Educational Research Journal 37(1):215-246.

Zentall SS, Kuester DA, Craig BA. 2011. Social behavior in cooperative groups: Students at risk for ADHD and their peers. The Journal of Educational Research 104(1):28-41.

Zhang D, Dai Z. 2004. The'Two Basics' Mathematics Teaching Approach and the Open Ended Problem Solving in China. Research in Mathematical Education 8(3):123-144.

Zhang L, Chang S, Li Z, Zhang K, Du Y, Ott J, Wang J. 2012. ADHDgene: a genetic database for attention deficit hyperactivity disorder. Nucleic Acids Res 40(Database issue):D1003-9.

Zhao JH. 2007. gap: Genetic analysis package. J Stat Softw 23(8):1-18.