Who is Like a Scientist? A Self-Prototype Matching Approach to Women's Underrepresentation in STEM Fields

Martin Ryan

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2014

Reading Committee:

Yuichi Shoda (Chair)

Sapna Cheryan

Nancy Kenney

Program Authorized to Offer Degree:

Psychology

University of Washington

Abstract

Who is Like a Scientist? A Self-Prototype Matching Approach to Women's Underrepresentation in STEM Fields

Martin Ryan

Chair of the Supervisory Committee:

Yuichi Shoda

Psychology

Why are women approaching parity in some science, technology, engineering, and math (STEM) majors but still lagging behind in others? I construct a multidimensional scaling map to understand which clusters of majors appear most unattractive to female students, and found that fields seen as scientific rather than cultural and as related to artificial kinds rather than nature were seen as least interesting by women. I then investigated the stereotypes which make those majors seem unappealing. I approach this from a self-prototype matching context, meaning that I predicted that participants would be most interested in majors whose stereotypes matched their self-image. A multi-level model was employed to examine the within-subjects patterns of response to various majors given stereotypes about those majors. Participants did in fact prefer majors whose stereotypes were similar to their self-stereotypes, and this tendency mediated gender differences in interest. In particular, stereotypes about agency, communion, and technological or “geeky” recreational interests were strong mediators. I also examined the consistency across subjects of these effects. Some stereotypes (such as agency) were high-

consensus, meaning that subjects consistently took them into account, expressing more interest in majors that were similar to them in those stereotypes and less in majors which were not. Others (such as sociability) were low-consensus, meaning that some subjects took them into account very strongly and others very weakly. Taken together, these results suggest that researchers studying stereotypes should expect the effects of those stereotypes to vary between subjects, and to depend particularly on those subjects' self-perception. Researchers studying the low-consensus stereotypes here should be particularly aware of this possibility, but all should take into account that the same intervention may have opposite effects for participants with opposite self-ratings. This dissertation advocates for interventions which emphasize the diversity of STEM fields and suggest that more than one prototypical class of student can succeed in them.

Chapter 1 Introduction

Gender inequality in the science, technology, engineering, and math (STEM) fields has been an enduring social problem. While the workforce in general becomes more gender integrated, some particular technical fields have lagged behind, perpetuating a division in some of the most prestigious and highly influential endeavors in our society. This dissertation will explore some of the possible reasons why college men and women might choose these STEM majors at different rates.

To understand this, I will explore how college students make decisions about majors. There is a vast array of majors to choose from, and most students will have only a small amount of actual experience with which to decide among them. So, how can one make such a decision? I'd like to put forward one possible approach: that students will be likely to select majors in which they feel that they will fit in or belong, majors in which people similar to them are to be found. I'll explore the effect of this sort of strategy for choosing majors on unequal representation in some majors, particularly among the STEM fields.

I will be focusing on gender differences in interest in STEM that relate to the decision to study or not to study STEM fields. What factors influence women to decide to study STEM subjects less often than men? What might make them feel like they belong less in those fields than men do? Success in STEM fields requires more than just deciding to study them, of course, but that decision is generally a minimal requirement.

The Gender Gap in STEM Fields

There is a substantial gender gap in women's participation in STEM in the US. It starts early in life and continues to become more pronounced as one progresses through the educational system. This is a serious inequality, because STEM training may lead to prestigious and well-paid jobs, and considerable effort has already been made to combat it.

The STEM fields are not universally unequal, however. Women make up 47% of US undergraduate degrees awarded in the life sciences, and 42% in math. On the other hand, they receive only 14% of degrees in engineering, 30% in physical sciences, and 25% in computer science. (NCES, 1997). I would like to understand the reasons why some STEM fields achieve increased gender parity and others do not. I will focus particularly on computer science as a field which has a strong bias and about which a body of literature exists.

Women are dramatically underrepresented in computer science, in spite of the many job opportunities it presents, and the considerable effort spent in the past 30 years to increase their representation (NCES, 1997). Computer science is unusual in that it has not only failed to advance in gender parity, but has experienced a decrease in female participation in the past 30 years (NCES, 1997). As computers play a growing role in our economy, a gender gap in the computer field both limits the opportunities available to women, and limits the range of perspectives which are available in solving technological problems.

The gender gap in interest and achievement in computer education starts early. Even in middle school, boys are more likely to aspire to science and engineering careers, and are more likely to have spoken to a scientist or used a computer than girls (NCES, 1997). Young girls on average have less experience with science in general and less interest in specific science topics (including computers) than young boys (Jones, Howe, & Rua, 2000). Middle and high school girls tend to perceive school computer classrooms as all-male clubhouses which they entered only with difficulty (Margolis & Fisher, 2002). Margolis et al. (2002) also found that girls were less likely to have home access to a computer for study or recreation purposes, though it is possible that the increasing ubiquity of computers has erased this inequality.

This difference persists in the college environment, where students are making the decisions which will shape their academic futures. Male first year college students listed CS as a major they were considering at five times the rate of female students (NCES, 1997). Women were especially likely to defect from science majors after entering, most often citing competitive environments and unresponsive teachers (Strenta, Elliot, Adair, Matier, & Scott, 1994). Nor does the gap close after graduation. In 2008, women made up between 19.4% and 29.2% of jobs in the CS industry, depending on subfield; they were only 20.6% of tenured CS faculty and 22.8% non-tenured faculty (Hill, Corbett, & St. Rose, 2010). Only 10% of tech company startup founders are female (Mitchell, 2011).

Research Perspectives Examining the factors that discourage women from choosing STEM

A useful point from which to approach the question of why women are less likely to choose STEM fields is Eccles' (1987) Expectancy-Value model. I'd like to start here because the expectancy value model involves an in-depth taxonomy of the factors which may affect women's choices, which makes it easy to integrate it with other perspectives even if they are not explicitly inspired by it.

The model proposes two pathways by which achievement related decisions may be influenced. First, there are the expectancy factors, which influence one's perceived likelihood of one's own success in the domain in question. For instance, if you have often been praised by teachers for your mathematical ability, you might expect that you could study mathematics with some success. On the other hand, if there are negative stereotypes of your gender in mathematics, you might have less such confidence. So for this half of the model, one would look to explain gender differences in choice by looking at factors which might increase or decrease confidence in a gender specific way.

The second possibility is that there may be factors which influence the perceived value of success in the domain in question. In other words, in addition to differing in perceived likelihood of success, people may differ in how good they think it would be to succeed. If you think it unlikely that you would enjoy a job in computer science even if you got it, you are unlikely to pursue it.

There is a considerable body of literature exploring the reasons why young men and women might have different expectancies or values for their STEM success, or in general why they might view STEM fields differently. I give a number of examples of common themes below. Identifying the developmental origins of gender differences is not directly the purpose of this study, however, so I will not dwell on them at great length. They should be seen as potential explanations for my findings; I am interested in seeing *where* differences between men and women occur, and these are possible reasons *why* such differences may have come about in the first place.

- Women may see themselves as worse at science and math than men are, in spite of actually superior performances. (Correll, 2001; Correll, 2004; Eccles, 1987; Lundeberg, Fox, & Puncochar, 1994)
- Girls are less likely than boys to have access to toys and extracurricular activities which encourage them to see science, computers, and math as fun, and give them an early

academic head start. (Kahle & Lakes 1983; Jones & Wheatly 1989; Jones & Wheatly, 1990; Crowley, 2001; Margolis & Fischer 2002)

- Women entering STEM fields can realistically expect to be in male-dominated environments, where stereotype threat and solo status represent threats to their success and wellbeing (Steele, 1997; Spencer, Steele, & Quinn, 1999; Good, Aronson, & Harder, 2007; Sekaquaptewa & Thompson, 2002; Sekaquaptewa & Thompson, 2003; Koch, Müller, & Sieverding, 2008; Adams, Garcia, Purdie-Vaughns, & Steele, 2006)
- They may also expect to experience negative treatment from their outside-STEM peers, because studying STEM fields may be seen as unfeminine and a violation of gender roles, which often leads to retaliation and decreased social opportunities (Deaux & Lewis, 1984; Eagly & Steffen, 1984; Cejka & Eagly, 1999; Eagly 2007; Rudman & Fairchild, 2004; Park, Young, Troisi, Pinkus, 2011)
- Women might simply have STEM field ability as a less central part of their self-concept, lowering the value of the field (Eccles, 1994; Nosek, Banaji, & Greenwald, 2002; Marsh & Yeung, 1998; Van de Gaer, Pustjens, & Van Damme, 2008).
- Women may place lower value on computer science education because they do not believe it fulfills their high level goals. Women are likely to prefer person-centered rather than thing-centered careers, and to prefer communal and family goals over agentic and financial goals (Lippa, 1998; Morgan, Isaac, & Sansone, 2001; Diekman & Eagly, 2008; Diekman, Brown, Johnson, & Clark, 2010).

There is a wealth of possibility in hypothesized explanations for the STEM gender gap, and it's difficult to sort them out, or to understand which of them is a relevant factor for which women, or to determine whether some have been left out. What I would really like to have is a sense of what each student imagines when he or she imagines computer science (for instance), and what features of that imaginary computer science make a real impact on each students' desire to study computer science.

Though expectancy and value are common focuses in the literature, and have been useful in organizing it, they don't appear to exhaust all the possibilities for reasons why one might choose to study a STEM field. I would like to devote particular attention to a recent approach falling outside this model, which focuses on the idea that we are more likely to participate in fields where we feel we belong. This can be understood in terms of social belonging, a feeling of connection to members of the group in question (Walton, Cohen, Cwir, & Spencer, 2011). It may also mean belonging to an academic field, the feeling that one fits in with the people in that field and that one's contributions are likely to be valued by them (Good, Rattan, & Dweck, 2012).

There is evidence that belonging in either sense is related to academic motivation and success. An intervention that minimized threats to belonging in college among African American students increased GPA, self-reported happiness, and self-assessed general health (Walton & Cohen, 2011), and women with a greater sense of belonging in mathematics were more confident and less anxious about math, and more likely to intend to pursue math education (Good et al., 2012). Students may be respond with greater motivation to participate even to small belonging cues, such as the gender balance of a conference (Murphy, Steele, & Gross, 2007) or sharing a trivial personal characteristic with a single peer (Walton et al., 2011).

It also appears that stereotypes associated with STEM fields may play a particular role in reducing interest. Exposure to a stereotypically “geeky” simulated computer science classroom reduced women's feelings that they would belong in computer science, which predicted a decrease women's interest in computer science (Cheryan, Plaut, Davies, & Steele, 2009). Likewise, interacting with a stereotypical computer science role model reduced belonging which resulted in a decrease in interest (Cheryan, Drury, & Vichayapai, 2013), and reading articles counteracting these stereotypes promoted women's interest (Cheryan, Plaut, Handron, & Hudson, 2013).

My goal in this dissertation is not to decide among these explanations, since I view them as compatible with each other. The above discussion is intended to express the breadth of possibilities which have already been considered in explaining women's lack of representation in STEM fields. In the current dissertation, my goal is to look at many of these explanations side by side, to look at which factors matter to which women. And even so, a study of all of the explanations discussed above would be too broad- as discussed in the next section, the approach I am using is best suited to cases where women's interest in STEM fields is related to their perceptions of the kind of person found in those fields, so I will focus particularly on explanations suited to this model.

The Role of Self-Prototype Matching

When we are trying to decide whether to engage with some community, one natural way to make the decision is to imagine the kind of person who would be found in that community, and ask whether that person is similar to oneself. This would suggest that women would be most interested in fields where they believe they will find others similar to themselves, on whatever dimensions they consider important.

What I am proposing, in other words, is a self-prototype matching model. The idea of such a model is that people have a prototypical image of themselves (a self-prototype, Kihlstrom & Cantor, 1984) and a prototypical image of the kind of person who might be found in a particular environment or situation (a person-in-situation prototype, Cantor, Mischel, & Schwartz, 1982). Thus, when making a choice between two or more environments or situations, people may compare their self-prototype with their person-in-situation prototype for each of the options, in order to determine which option is most likely to contain people most similar to them.

Further, the self-prototype matching approach argues that people are motivated in general to enter situations in which they think people similar to themselves are to be found, both because doing so is a good predictor of one's own success in those situations, and because it produces a feeling of self-consistency to behave in accordance with one's expectations for people like oneself (Niedenthal, Cantor, & Kihlstrom, 1985). Thus, people will be motivated to choose the option where the degree of similarity between the self and the prototype of those who choose that option is greatest. For instance, students who saw themselves as similar to their prototypical graduate student were more likely to have an interest in graduate education, while those who saw themselves as similar to their prototype of a student who enters the workforce after graduation were more likely to do that (Burke & Reitzes, 1981). Self-prototype matching effects have been found in other decisions relevant to students, such as the decision to smoke cigarettes (Chassin, Presson, Bensenberg, Corty, Olshavik, & Sherman, 1981), and the decision between various kinds of student housing (Niedenthal et al., 1985) and between various models of car (Setterlund & Niedenthal, 1993).

Thus, to understand the different outcomes for biology and computer science, for instance, we would want to understand what women see in biology that is similar to their self-concept, or what women see in computer science that clashes with it. We want to understand women's preferences for various majors in terms of their perceived similarity to the kind of person who would be found in that major.

One fact which will have a good deal of influence on the mathematical underpinnings of this dissertation is that self-prototype matching proposes a nonlinear relation between the traits ascribed to a major and participants interest in it. Participant interest should rise with increasing major trait rating until it reaches a maximum when the major and self-ratings are the same, and decrease thereafter, forming a parabola peaked at the self-rating. Since linear models are the typical case in social psychology, this will occasionally involve some unusual predictions and statistical models, which I will cover in the relevant chapters.

Of course, since undergraduate students typically have limited information about what life is really like for STEM majors, what must really shape their decisions is their stereotypes and impressions of that life. Thus, I am interested in looking at how stereotypes about majors compare to male and female students' self-image, and how that comparison relates to their interest in pursuing those majors.

This approach is compatible with the expectancy-value approach discussed above. Students are likely to determine whether or not they have the skills necessary to succeed in a STEM field by comparing their own skills to those of someone who could succeed- that is, if I can do what computer science students can do, I can study computer science successfully. Likewise, if I am similar to people who enjoy studying computer science, it's reasonable to think that I would enjoy it too. So we should expect that factors that play a role in the Expectancy-Value model would carry over into a self-prototype matching model.

In addition, the set of available hypotheses is broader in a self-prototype model. The expectancy value model might almost be thought of as an economic, success based model in which one computes the chances of success and the value of success and multiplies them. However, not all factors in decision making may be related to at all to success. Students may seek a major with others like them because they desire it for its own sake, in addition to seeking it because it will help them succeed. We might expect them to vary in their expectancy of similarity to computer scientists and the extent to which they value similarity, in addition to success. So I hope to supplement the expectancy-value approach by looking in directions which are not included in its model.

Likewise, this approach is compatible with the belonging approach discussed above, but not identical with it. It participants are likely to believe they would belong with people to whom they are similar (and similarity has been used as a belonging manipulation by e.g. Walton et al., 2012), but belonging also includes a sense of being valued and connected which similarity does not necessarily imply. Thus, one could feel belonging without similarity (if one believes one's difference has been embraced) and feel similar without belonging (if one feels similar to the target group while still anticipating rejection).

The most direct application of the self-prototyping model to STEM I am aware of is by Hannover and Kessels (2004), who asked German high school students to describe a student whose favorite class was either English, German, math, or science in terms of trait adjectives, and found that students were most likely to favor classes where they described the prototypical student as similar to themselves. This will serve as a jumping off point for my own studies; I would like use an approach similar to Hannover and Kessels', but to extend it as deeply as possible.

Specifically, I'd like to expand in three directions. First, I'd like to try to evaluate the perceptions students have of STEM fields as broadly as possible, not just in terms of trait adjectives. To do this, I will begin with items collected from the literature above, attending to span all of the currently hypothesized ways in which STEM fields may be stereotyped. I will also use open-ended questions to generate additional potential features which may be seen differently by men and women and which have not been included in previous research. This will enable us to get as complete a picture as possible of how students see STEM fields.

Second, I will not assume that “similarity” is a unitary construct. Participants may see themselves as similar to computer science students in some ways and not in others. Some sorts of similarity may impact interest in STEM while others do not. And the weighting of these factors may be higher for some students than for others. For instance, some students may care very much about the perception whether or not students in a given major are seen as sociable, while another student might not care very much about that factor at all. By taking a multi-factor, idiographic approach, we hope to discover not just whether similarity matters, but what kind of similarity and to whom.

And third, I would like to devote some time first to figuring out where the STEM fields fall in relation to each other. It's clear that, though my topic here is STEM fields, not all of them have the same problem, and we would particularly like to counteract gender bias in those fields where it is most present. It might be that it's been random which fields turn out to suffer from bias, but I think it's more likely that subjects will see the male dominated fields as conceptually related. As such, I will spend some time developing a conceptual map that defines how closely each field is to each other field in the eyes of subjects, and will attempt to look at what concepts subjects are using to determine those relationships.

Notes, Caveats, and Definitions

The words “stereotype” and “prototype” are used in sufficiently similar contexts that the difference between them often becomes obscure. I will be following Hannover and Kessels (2004) in understanding a stereotype to be a description of a group, while a prototype is a description of a typical member of that group. The difference between the two is sufficiently fine as to be for the most part irrelevant to our goals here, but for the sake of fidelity to convention I will use the word “prototype” when discussing research which also uses that word, and likewise for “stereotype”. Readers should not assume I am making a strong distinction between the two.

I'd also like to mention that when I talk about “choosing” to study computer science or biology, I am not talking about a unitary decision made on a particular day. Rather, there are countless minor decisions to take the next step on the path through a college major, including the decisions necessary to enter a major and those necessary to remain in it. However, I will focus initially on the perspective of young students who are not majoring in computer science but might decide to do so, both because they seem particularly likely to be malleable in their attitude toward computer science and because they represent a larger and more diverse pool of subjects.

The focus on choices I make here may suggest that I think inequality results only from men's and women's different choices, and of course this is not the case. I will mention some possible alternatives here, following Ceci and Williams' (2010) division of the possibilities. First, some have suggested innate differences in men's and women's mathematical ability, particularly among highly mathematically talented students, may be the source of the STEM gap (e.g. Benbow & Stanley, 1982, Hedges and Nowell, 1995). However, Ceci, Williams, and Barnett (2009) review data on direct hormonal or structural effects on mathematical performance and find them inconclusive; in addition, they note that if one examines these trends internationally, they are more pronounced in some nations and reduced or even inverted in others, suggesting that they are not likely to be universally innate differences in men and women. Thus I will neglect such explanations hereafter. A second class of explanation is that in spite of a high degree of interest in these fields, biased evaluations or gatekeepers prevent women from entering them. Gender bias in evaluation has been found (for instance) in peer review (Wenneras & Wold, 1997), letters of recommendation for faculty (Trix & Psenka, 2003), and on the evaluation of academic job applicants and tenure candidates (Steinpreis, Anders, & Ritzke, 1999). These are potentially very important in determining who will become a professor, but have little to tell us about who will take a first year college mathematics course, before any of these gatekeepers hold sway. So for our current study of the decision to study STEM courses, these factors will also be left out. But note that the expectation of sexism in the field, or the stereotypical association of the field with sexism, is within the scope of this paper.

Finally, looking at how men and women choose differently really means looking at the factors which influence men and women to choose differently. Choices are constrained by social situations, and women may experience negative consequences for making choices not in keeping with assigned gender roles (e.g. Rudman & Fairchild, 2004). Men and women may be culturally influenced to develop different sorts of self-concept as well. For instance, women may derive negative beliefs about their own competence from cultural stereotypes about women (Correll, 2001, Correll, 2004), and in particular may have a negative implicit association between math/science and femaleness (Nosek, et al., 2002). All of these things are malleable, however, and correcting inequality means understanding the mechanisms by which it propagates, which is the goal of this dissertation.

Studies in Brief

In Study 1, I will begin to examine the relationships among majors by constructing multidimensional scaling map, in which subjects' ratings of similarity are used to place each major on a grid with other, similar majors near it. This will give insight into which ways of classifying majors are most relevant to our subjects.

In Study 2, I will test the feasibility of our self-prototype matching model by asking questions participants to rate both themselves and computer science students on our list of stereotypical traits, to see if participants who see themselves as similar to CS students will also be more interested in CS.

In Study 3, we will apply the model from Study 2 to multiple majors. This will enable a multilevel modeling approach, allowing us to see whether (for each subject) some sorts of similarity judgments consistently predict interest, and to make inferences about which kinds of similarity are most relevant to which groups of people.

Chapter 2 The Shape of College Majors

STEM fields are not all created equal. Participants respond quite differently to biology than they do to computer science, for instance. So in addition to knowing how students feel about STEM, or about individual majors, it will be useful to know how students think that various majors relate to each other, or which they see as similar or not similar. In this study, I would like to create a map of the relationships among majors as it appears to students. We can then observe the parts of that map which are most appealing to women or to men.

Two phenomena are potentially of interest to us here. First, we can find the criteria which subjects are most prone to use in classifying majors, without having to give the subjects any prior information about what we're looking for, and we can see where each major falls on these criteria. Second, we can look at the ways men and women's interest in majors depends on these criteria, as a first step in understanding what differences drive some majors to have a gender gap.

There are several reasons why it will be useful to know the relationships among majors. First, it allows us to compact the long list of college majors down into fewer, more manageable units. Here the map metaphor is very useful, since we can think in terms of majors which are neighbors, or part of clusters of related majors. We might plausibly say then that some phenomenon applies to the majors in one cluster but not another. So in the future, we will be able to see where claims about one major might be generalized to related majors, allowing us to make more broadly useful claims.

A second advantage derives from the fact that a map requires compass directions. That is, it requires multiple dimensions on which majors can be different or similar. We derive these dimensions in a bottom up way, from the subject's own responses. Subjects are not told what basis to use for their comparisons; rather we allow them to show us what basis they choose to use. Thus, this method enables us to make statements about how the subjects themselves divide their academic options. For instance, if the two most prominent dimensions appear to relate to solitary vs. team work and to level of monetary compensation, we can say that these two are constructs which our subjects use to parse the world. We can even go farther, and examine individual differences in the extent to which subjects rely on each dimension. Thus, the process of creating the map will also teach us about what is important to our subjects.

There are also several uses to which this information can be put, some of which we will return to in future chapters. First, it allows us to talk about categories among STEM majors, rather than individual majors or STEM as a whole. Second, it motivates future exploration of the dimensions that subjects use to divide the majors. Third, if there are individual differences in the importance of some dimensions, it tells us which people are likely to care about which features of majors, enabling future interventions to be better targeted.

Since gender is my focus here, I will be particularly interested in how gender differences might emerge in this process. It may be that men and women weight different things as important in the construction of the map, and thus will produce very different maps. Even if the maps are the same, it may also turn out that men prefer majors in one part of the map, and women prefer majors in another. Either result would indicate that the map we are creating here is significant for understanding the relationship between gender and choice of majors.

In a multidimensional scaling (MDS) model, similarities between pairs of items are used to create a map of the relationship among those items. If we were to perform multidimensional scaling on the locations of US cities, we would begin with a list of the distances between each pair cities, and the algorithm would try to recreate a map of the US using only those distances. In our case, participants will rate each pair of majors for similarity, and the similarities will create a map that reflects their picture of the college major space.

Like factor analysis, MDS allows the researcher to specify the number of dimensions to be used to build the map. It becomes prohibitively difficult to visualize many-dimensional maps, which limits their usefulness as illustrations and the ease of interpreting them, but the dimensions chosen may still have psychological meaning. In this study, I will base the decision of how many dimensions to use on the additional variance which will be accounted for by the next dimension, but the higher numbered dimensions should be viewed as to some extent tentative.

MDS could be used separately for each participant to create a library of idiographic maps, one for each subject. But we would like to be able to integrate them into an overall model. For this purpose I will use the individual difference scaling model (INDSCAL), which combines the participants' ratings to create a consensus map. At the same time, INDSCAL will tell us how each individual's map differs from the consensus map, in the form of individual difference weight coefficients. Thus, we will have two things to analyze: the consensus map, and weight coefficients for each subject.

There have been some previous attempts at classifying academic disciplines. The classic result in this area is Biglan's (1973) use of MDS for a dichotomous divisions of majors into physical/nonphysical, applied/theoretical, and life/not-life categories. This is more or less also our goal, but Biglan's work was focused on the views of professional academics with an eye towards understanding publishing patterns. Coupled with the fact that his analysis is over 40 years old, and does not does not mention gender at all, I believe my current analysis is a useful addition. Gender is more relevant to Diekman et al.'s (2010) division of career paths into STEM, female stereotypical, and male stereotypical categories; but their study focused on classifying careers rather than academic disciplines, and used a priori categorization based on gender percentages rather than allowing subjects to provide their own judgments based on their own categories.

My goals here are exploratory, and I do not advance a hypothesis about this data. There are several questions to be answered by this analysis: 1) What dimensions do subjects use to organize the space of college majors, 2) do men and women differ in their structuring of the major space, 3) how do participants' levels of interest in a field depend on where they place it in their cognitive map?

Method

In this study, 45 participants (20 women) rated all pairwise combinations of 24 majors for similarity. Three participants identified as Black/African American, 25 as Asian/Asian American, 1 as Latino/Hispanic, 3 Native American/American Indian, and 15 as White. Majors were selected by the researcher from those available at UW[1], and were intended to include a wide range of both STEM and non-STEM majors, and to focus on majors likely to be familiar to subjects. In some cases the names of the majors were edited to be for clarity or to avoid double-barreled questions; for instance, "Earth and Space Sciences" was shortened to "Earth Sciences", "Business Administration" (which includes finance as a subfield) was changed to "Business and

1 The majors used were Biochemistry, Biology, Chemistry, Communication, Education, Economics, English, History, Political Science , Psychology, Sociology, Business and Finance, Earth Sciences, Oceanography, Civil and Environmental Engineering, Astronomy, Computer Science, Medicine or pre-med), Mathematics, Mechanical Engineering, Anthropology, Architecture, Electrical Engineering, Visual and Performing Arts. Due to a programming oversight, Physics was excluded from the original sample in spite of its theoretical interest; a followup sample was run to determine its appropriate position on the map of majors. The map including physics is included in Appendix A; none of the other results were substantively altered by its inclusion.

Finance", and "Early Childhood and Family Studies", the sole major offered by the College of Education, was changed to "Education". Also included was "Medicine or pre-med", which is not a major offered at UW, but which was included because medicine is an example of a field which has achieved near gender parity in spite of being both technical and historically male-dominated (Diekman et al., 2010). Though ideally evaluating all the majors available would be preferred, the number of pairwise comparisons rises exponentially in base number of units compared.

For each pair of majors, participants were asked to rate how different or similar they found that pair of majors, on a 1-7 scale from "Very Dissimilar" to "Very Similar". To facilitate interpreting individual differences, participants also rated their own interest in each major and their skills with and interest in writing, computers, and math, and their communal and agentic goals (Diekman et al., 2010) .

One participant was dropped due to omission of over 100 items, and four requested during debriefing that their data not be used.

Results

Choosing the appropriate number of dimensions

In order to decide how many dimensions to use, I first ran a simple multidimensional scaling model on the aggregate data of all participants. The goal of this is to find a set of dimensions which account for a large percentage of the variance in the data using as few parameters as possible. In this case, a 1-dimensional model accounted for 36% of the variance in similarity ratings; a 2-dimensional model, 52%; 3-dimensional, 61%; and 4-dimensional 66%. A majority of the variance is accounted for by the first two dimensions, which will accordingly be our focus. The third and fourth dimensions account for only a small amount of variance, and will be not be analyzed further. . A map of the majors as they are placed in Dimension 1 and Dimension 2 is presented in Figure 2-1a.

The current study provides us no means of assigning precise labels to these dimensions. In order to understand the meaning of these dimensions, I conducted a followup study. I will report its results below before resuming discussion of the main study.

Figure 2-1a, Multidimensional Scaling Map

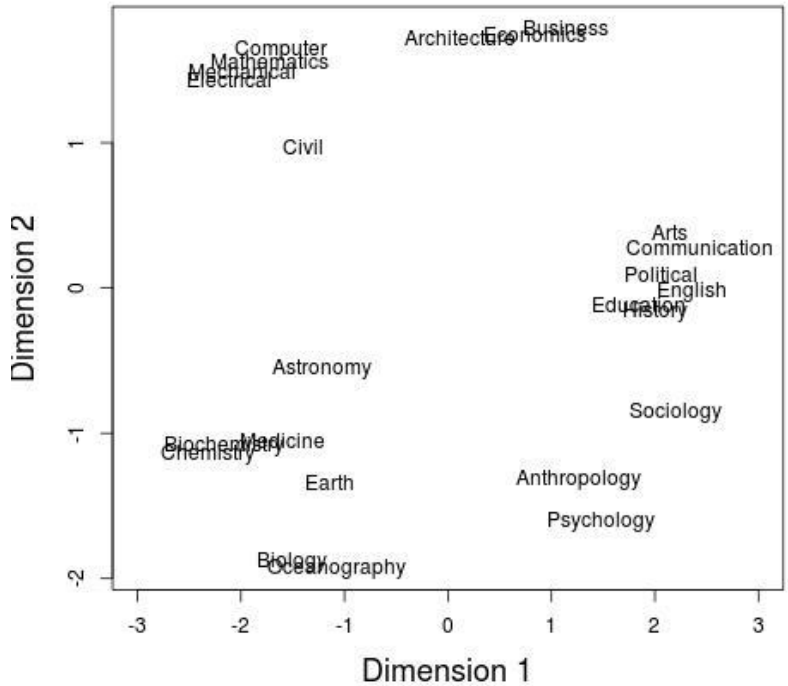
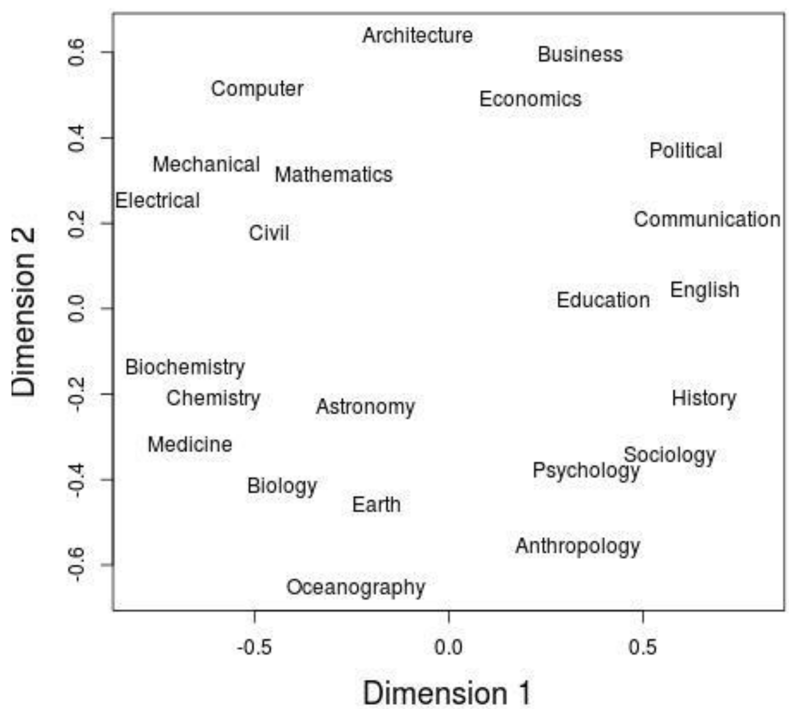


Figure 2-1b, Individual Distance Scaling Map



Labeling Dimensions

Sixty University of Washington undergraduates (43 female) participated in exchange for course credit. Eight participants identified as Black/African American, 27 as Asian/Asian American, 9 as Latino/Hispanic, and 29 as White, and 3 as other. Subjects were shown a list of the seven lowest scoring majors on dimension 1 (Chemistry, Biochemistry, Electrical Engineering, Mechanical Engineering, Mathematics, Biology and Medicine) and the seven lowest scoring (Communications, English, Sociology, Arts, Political, Science, History, and Education). They were asked to describe in their own words how the majors were different, and then to rate them on 39 closed ended items derived from open ended responses in the main study. The procedure was then repeated for the highest scoring (Business, Mathematics, Economics, Computer Science, Mechanical Engineering, Architecture, Electrical Engineering) and lowest scoring (Oceanography, Biology, Earth Science, Anthropology, Chemistry, Biochemistry, Sociology) majors on dimension 2.

Participants saw a very sharp distinction between majors high and low in dimension 1 (Table 2-1). The strongest areas of difference were that high scores in dimension 1 predicted associations with math, science, and high-paying jobs, while low scores in dimension 1 most strongly predicted associations with culture and writing. Out of the 60 subjects, 39 used at least one of the words “science”, “math”, or “technology” in their open-ended responses, and 12 used at least one of “art” and “humanities” (see Appendix B for the full list of open-ended responses). On the basis of this, we may label this a science vs. culture axis. Several of the majors appearing on the “culture” side of the axis are also sciences, but participants evidently did not see them this way.

The distinctions were somewhat less clear for the second dimension, but nevertheless the effects were highly significant (Table 2-2). Majors low in dimension 2 were most strongly associated with the natural world and the environment, while majors high in dimension 2 were associated with the economy, math, and building things. This distinction is not as easy to clearly label with a single English word as in dimension 1. Of the 60 subjects, 20 used at least one of the words “natural” and “environment” in their responses, and 24 used at least one of the words “technology” and “math”. This axis seems to join studies of the natural world and of human behavior on the one hand, and of technology, economics, and politics on the other. I will label this the natural vs. artificial axis for brevity, but it would be more accurate to call it the natural/behavioral vs. artificial/political/economic axis.

Though the primary purpose of this study was to verify that we have correctly understood the meaning of the dimensions which emerged from our scaling algorithm, it is also worth noting that participants saw a gendered component in both dimensions. They believed that majors low in dimension 1 were more often studied by men and majors high in it more often studied by women; and they believed that majors high in dimension 2 were more often studied by men and majors low in it more often studied by women. As we will see, both of these match the actual average preferences of subjects for those majors in the main study. (It may also be interesting to note as an aside that in both cases the list of majors associated with men was also associated with high-paying jobs by subjects).

Trait	Difference	t(df)	Cohen's d
Technology	-2.63	-10.92(59)*****	-1.91
Science	-3.59	-14.64(58)*****	-2.6
Life	0.22	1.94(59)	0.2
The natural world	-1.93	-6.98(59)*****	-1.34
Culture	1.91	6.87(57)*****	1.35
Things created by humans	-0.03	-0.17(59)	-0.03
Large scale phenomena	-0.39	-2.28(58)*	-0.3
Small scale phenomena	0.38	2.05(59)*	0.26
The environment	-1.13	-5.42(59)*****	-0.84
Industry	-1.36	-6.24(58)*****	-1.1
The economy	-0.12	-0.65(57)	-0.1
High-paying jobs	-2.27	-10.41(59)*****	-1.96
Jobs in people-oriented fields	1.23	5.62(59)*****	0.95
Jobs that help society	-0.07	-0.41(59)	-0.06
Math	-3.37	-13.7(59)*****	-2.74
Writing	1.55	6.64(59)*****	1.16
Creativity	0.79	3.83(57)***	0.6
Building things	-1.77	-8.34(56)*****	-1.31
Studied by men	-1.37	-7.62(59)*****	-1.05
Studied by women	0.98	5.66(59)*****	0.71

Table 2-1 Differences between high and low majors on Dimension 1

Trait	Difference	t(df)	Cohen's d
Technology	0.68	3.82(59)***	0.54
Science	-0.8	-3.94(58)***	-0.67
Life	-0.33	-2.57(59)*	-0.3
The natural world	-1.68	-6.59(59)*****	-1.18
Culture	0.27	1.23(58)	0.18
Things created by humans	0.97	4.07(59)***	0.71
Large scale phenomena	-0.37	-2.35(59)*	-0.27
Small scale phenomena	0.17	1.12(59)	0.11
The environment	-1.22	-5.35(58)*****	-0.95
Industry	0.92	3.91(59)***	0.69
The Economy	1.29	5.61(58)*****	0.96
High-paying jobs	0.72	4.18(59)****	0.58
Jobs in people-oriented fields	-0.05	-0.24(59)	-0.04
Jobs that help society	-0.22	-1.56(59)	-0.18
Math	0.98	5.23(59)*****	0.81
Writing	-0.27	-1.53(59)	-0.21
Creativity	0.15	0.98(59)	0.11
Building things	1.15	5.06(58)*****	0.78
Studied by men	0.68	4.22(59)****	0.49
Studied by women	-0.37	-2.57(59)*	-0.28

Table 2-2 Differences between high and low majors on Dimension 2

Individual Difference Scaling

Having chosen a number of dimensions, I can apply the Indscal algorithm to address individual differences in the model. As described above, this results in a consensus map and a list of weights for each participant, indicating the degree of importance that each dimension has for that participant. The consensus map is presented in Figure 2-1b; as you can see, it closely resembles Figure 2-1a.

There were no gender differences in the weight assigned to Dimension 1 by women (*M*=.94, *SD*=.06) and men (*M*=.93, *SD*=.06), *t(37.46)*=.47, *p*=.64. Likewise, there were no differences in the weight assigned to Dimension 2 by women (*M*=1.12, *SD*=.07) and men (M=1.13, *SD*=.06), *t(34.99)*=.73, *p*=.47. Thus it appears that men and women were in agreement as to the dimensions they use for dividing majors, though they may still differ in how they feel about those dimensions.

Relationship between MDS dimensions and interest. Since our goal is to understand what drives interest in various fields, I will next look at the dimensions as predictors of interest in each field. I must note at the outset of this process that our participants are recruited from introductory psychology classes, and thus are likely familiar with and interested in psychology. Indeed, psychology was significantly more interesting (*M*=5.60, *SD*=1.12) than the next most interesting major, pre-medicine (*M*=4.37, *SD*=2.01), *t*=3.37, *p*<.01. The mean levels of interest are presented in Table 2-3. The majority fall between 3 and 4 on the 7 point scale, suggesting that psychology students did not have strong consistent preferences for or against any majors besides psychology.[2]

2 In order to rule out the possibility that the presence of psychology skewed the map created in some way, the analyses in this chapter were repeated with psychology omitted. No substantive changes in the other analyses resulted.

For the most part men and women did not differ in their level of interest in the various majors. As depicted in Table 2-4, men rated Computer Science, Astronomy, Architecture, and Civil, Electrical, and Mechanical Engineering as more interesting than did women, while women rated Education as more interesting. Also included is the percentage of bachelors degrees in each major awarded to women in 2011-2012 (NCES, 2013). The fact that our sample did not show significant gender differences on some majors which are in fact disproportionately male or

Field	Mean	SD
Biochemistry	3.92	1.84
Biology	4.08	2.02
Chemistry	3.69	2.10
Communication	4.01	1.54
Education	4.02	1.62
Economics	3.99	1.95
English	2.85	1.39
History	3.16	1.83
Political	3.50	1.85
Psychology	5.61	1.12
Sociology	4.36	1.80
Business	4.13	1.77
Earth	3.53	1.87
Oceanography	3.60	1.56
Civil	3.23	1.60
Astronomy	3.72	1.69
Computer	3.27	1.97
Medicine	4.38	2.02
Mathematics	3.77	1.85
Mechanical	3.13	1.76
Anthropology	3.77	1.67
Architecture	3.17	1.60
Electrical	2.95	1.70
Arts	3.40	1.92

Table 2-3 Mean Levels of Interest

Field	M	F	t(df)	%female
Biochemistry	4.08	3.72	0.59(28.81)	47.88
Biology	4.12	4.04	0.13(34.17)	58.74
Chemistry	3.85	3.5	0.5(32.45)	48.44
Communication	3.83	4.22	-0.77(31.95)	62.23
Education	3.55	4.59	-2.08(32.82)*	79.43
Economics	4.52	3.33	1.96(34.27)	29.24
English	3.05	2.61	0.99(34.26)	68.43
History	3.61	2.61	1.73(31.82)	40.29
Political	3.98	2.93	1.84(36.4)	43.83
Psychology	5.45	5.8	-0.96(37.2)	75.59
Sociology	4.11	4.67	-0.93(29.36)	69.15
Business	4.17	4.09	0.13(31.93)	48.18
Earth	4	2.94	1.77(29.62)	39.06
Oceanography	3.92	3.22	1.41(35.19)	46.90
Civil	4.06	2.22	4.42(37.46)****	20.85
Astronomy	4.52	2.74	3.78(32.15)***	38.16
Computer	3.98	2.41	2.71(36.84)*	13.05
Medicine	4.31	4.46	-0.22(27.45)	55.11
Mathematics	3.92	3.58	0.56(32.59)	43.09
Mechanical	3.93	2.15	3.73(37.98)***	12.09
Anthropology	3.84	3.69	0.28(29.45)	71.22
Architecture	3.77	2.43	2.81(31.84)**	42.77
Electrical	3.85	1.85	4.64(37.82)****	11.90
Arts	3.28	3.54	-0.41(35.98)	61.21

Table 2-4 Gender Differences in Interest

female should not be over-interpreted, since our sample size here is relatively small.

More central to our current investigation than the baseline level of interest in each field is the degree to which the majors in different parts of the map appeal differently to women and to men. To test this I ran a multilevel model using the two individual difference scaling dimensions and gender as predictors of interest. There was a significant main effect of gender B=.97, Wald c^2=6.91, p<.01, with men generally more interested in majors over all. There was a significant main effect of Dimension 2 (natural/artificial) B=.64, Wald c^2=13.25, p<.001, with more "artificial" fields of study predicting more interest; but there was no main effect of Dimension 1 (science/culture) B=.97, Wald c^2=2.11, p=.15.

Both interactions were significant; women were more likely to prefer fields closer to the culture end of the science/culture axis B=.78, Wald c^2=13.72, p<.001. The mean slope for women was .52 (*SD* =1.41), and for men, -.14 (*SD*=1.22). They were also more likely to prefer fields closer to the "natural" end of the axis B=.80, Wald c^2=7.75, p<.01. The mean slope for women was -.80 (*SD*=1.29), and for men, .52, (*SD*=1.41). The two-way interaction between the dimensions B=-1.08, Wald c^2=3.15, p=.08 and the three-way interaction between the dimensions and gender B=1.19, Wald c^2=5.82, p=.05 were both significant at the trend level, suggesting collectively that women may have a slightly higher preference for the cultural and artificial fields (e.g. business, communication) than the additive main effects would suggest.

These interactions suggest that women's interest is likely to be concentrated in the cultural and natural areas of the map, but of course not all women's interests lie there, as we can see from the large standard deviations. Individual differences in the weighting of the science/culture axis were significant c^2=37.10, p<.00001, as were individual differences in the weighting of the natural/artificial axis c^2=140.98, p<.00001. Figures 2-2 and 2-3 below depict the range of slopes for science/culture for women and men respectively, with each line corresponding to one subject. Likewise, Figures 2-4 and 2-5 depict the slopes for natural/artificial for men and women respectively. The range of slopes in these graphs represents the sensitivity of each subject to the dimension depicted; vertical lines indicate indifference and the closer to horizontal the lines get, the stronger the preference for one side rather than the other in the continuum.

Figure 2-2 Slopes for Science/Culture, female subjects

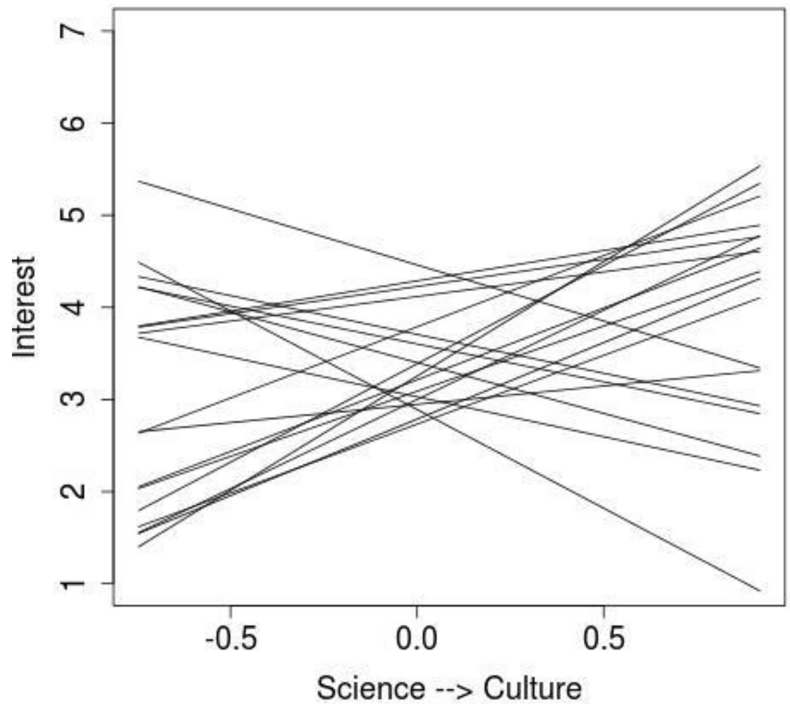


Figure 2-3 Slopes for Science/Culture, male subjects

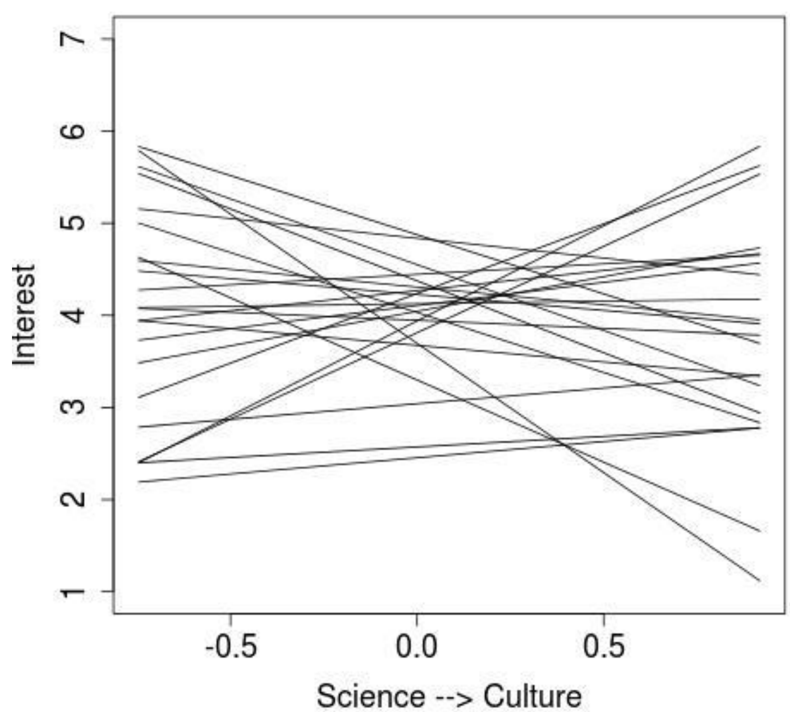


Figure 2-4 Slopes for Natural/Artificial, female subjects

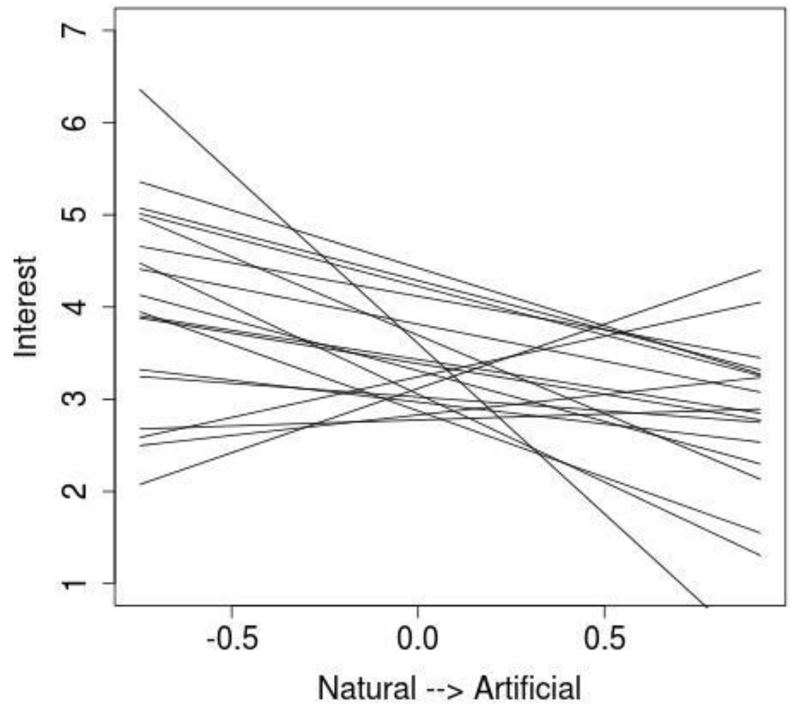
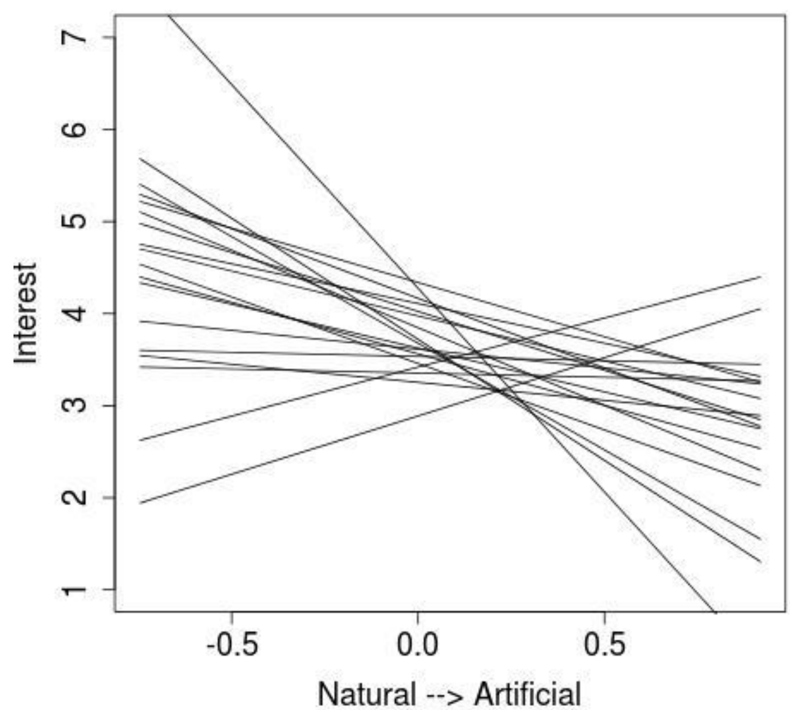


Figure 2-5 Slopes for Natural/Artificial, male subjects



A summary of the relationships between preferences for either the science/culture axis or the natural/artificial axis and available individual difference measures is included in Table 2-5. These included the communal and agentic goals from Diekman et al., (2010), self-ratings of facility with Math [3],Computers[4], Writing[5], and self-rated masculinity and femininity. The correlations, though not strong, support the previously noted relationships between the "science" and "artificial" poles and mathematics, and between the "culture" pole and writing.

	Science/Culture Axis		Natural/Artificial Axis	
Individual Difference	correlation	sig	correlation	sig
Communal Goals	.20	.21	-.30	.06
Agentic Goals	-.06	.71	.10	.53
Math Facility	-.37	.02	.31	.05
Computer Facility	-.22	.17	.12	.44
Writing Facility	.42	.01	-.22	.17
Femininity	.16	.32	-.37	.02
Masculinity	-.13	.40	.25	.11

Table 2-5 Correlations between preferences on axes 1 and 2 and various individual difference measures.

Though self-prototype matching is not the central focus of this chapter (as it will be in the next chapter), it deserves some mention here. We cannot directly evaluate participants' self-prototype match, since we do not know anything about their self-prototype at all. However, the self-prototype matching theory suggests that for any given descriptor of a college majors, participants will have an optimal point (dependent on their self-rating) which corresponds to

3 Mean of "How comfortable are you with math?" "How much would you like to have a career involving math?", Cronbach's alpha=.79

4 Mean of "How comfortable are you with computers" "How much would you like to have a career involving computers?", Cronbach's alpha=.74

5 Mean of "How comfortable are you with writing?" "How much would you like to have a career involving writing?", Cronbach's alpha=.84

maximum interest, and interest will fall off as a parabola on either side of that maximum. In other words, the interest-descriptor relationships should be quadratic. Of course we will only see that relationship appear if we happen to have a given subject's optimal point inside our range, with sufficient data points close to it to resolve a parabola; nevertheless, it may be interesting to attempt the quadratic model and see what we find.

Figures 2-6 through 2-9 show parabolic curves fitted to each participant's data. For ease of reading, I've selected 9 subjects arbitrarily for each plot. For the science/culture axis (Figures 2-6 and 2-7), the quadratic model performed significantly better than the linear one, c^2 (4)=24.13, $p<.0001$. For the natural/artificial axis, there was no significant difference, c^2 (4)=3.57, $p=.47$.

Figure 2-6 Curves for Science/Culture, female subjects

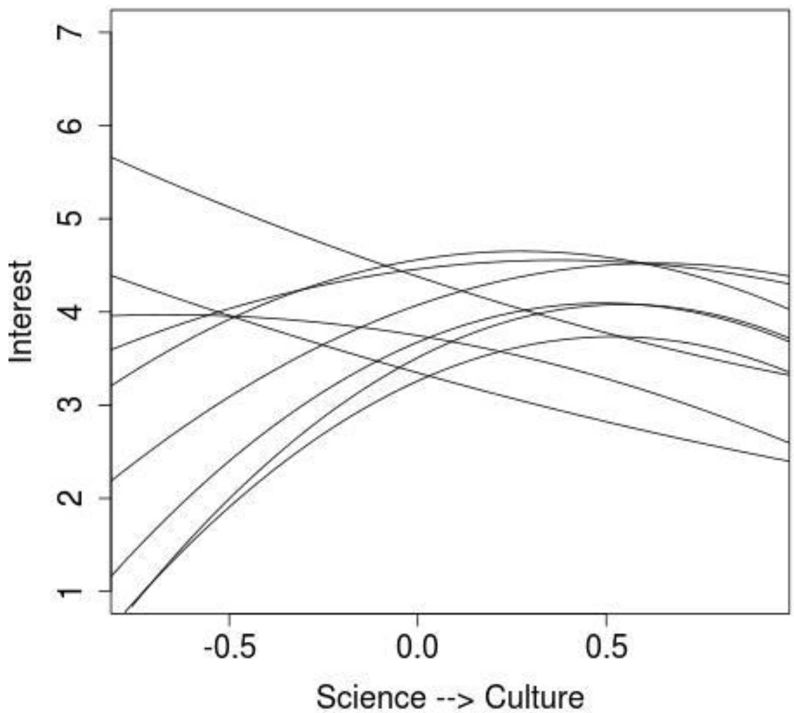


Figure 2-7 Curves for Science/Culture, male subjects

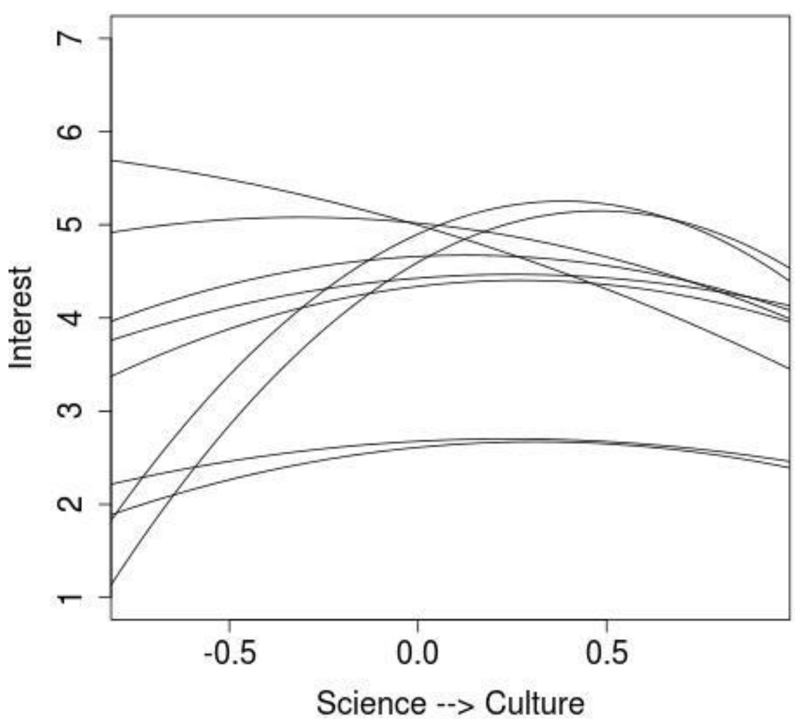


Figure 2-8 Curves for Natural/Artificial, female subjects

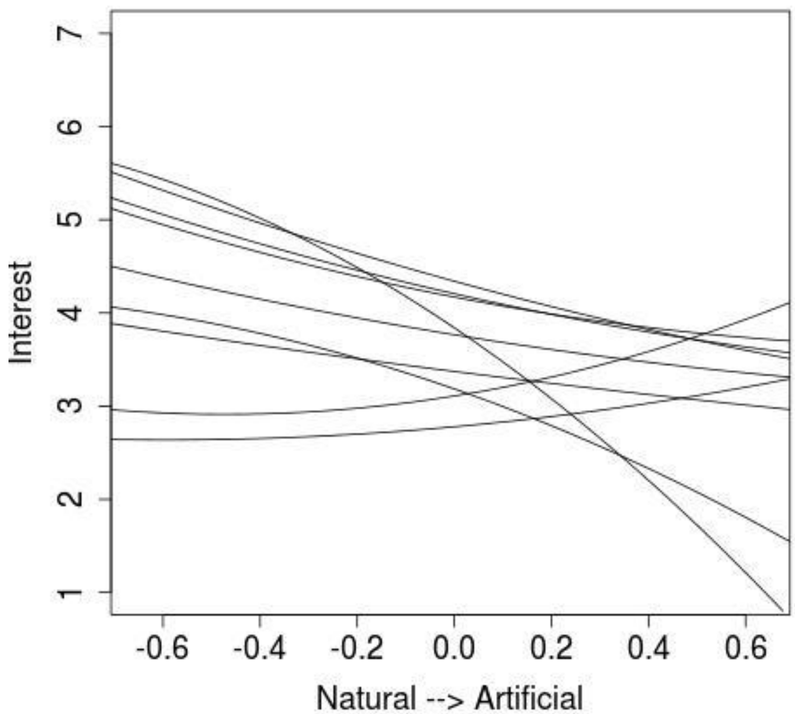
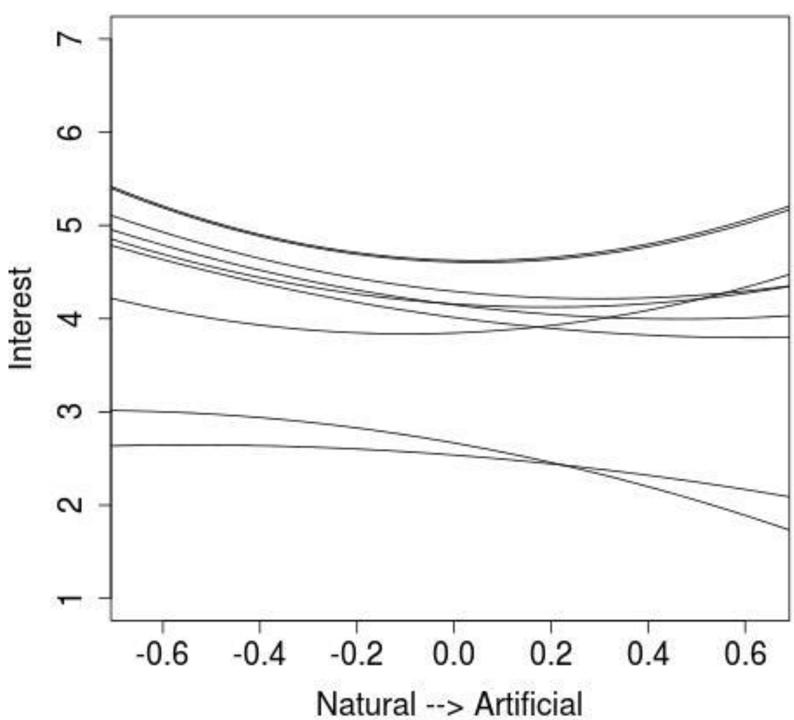


Figure 2-9 Curves for Natural/Artificial, male subjects



Discussion

These studies have given us some insight into the way that students see the fields of study available to them. It appears that they clearly divide their options into STEM and culture-related majors, and also into majors dealing with the natural world and those dealing with artificial kinds (Goal 1). Male and female subjects did not differ in their weighting of the dimensions, and thus did not differ in the maps they created (Goal 2). Women proved to be relatively more interested in the humanities and in natural kinds, meaning that they were least interested in the cluster of majors lying in the science and artificial quadrant of the map (Goal 3).

The contrast between STEM and non-STEM majors is widely acknowledged, as is women's relative absence in STEM fields (e.g. Snyder, Dillow, & Hoffman, 2009), though it is interesting to see which fields participants see as belonging to those categories. The natural/artificial split seems congruent with women's preference for people-related rather than thing-related careers (Lippa, 1998) and with the particular interest of some women in health-related science careers over science in general (Eccles, 2007), but is not identical with either of these. Business, mainly a people oriented field, nevertheless falls on the “artificial” half of the axis, while disciplines studying the environment but not people, such as oceanography, fell on the “natural” half of the axis. So it may be that these people/things dichotomy is narrower than it needs to be. One possible reason for this is that the effect of dimension 1 masks that of dimension 2; women were not more interested in oceanography than men were, but they were more interested in it than in other STEM majors on average. Only by separating the two dimensions does this fact become noticeable.

There are several ways in which these results go beyond simply allowing us to demonstrate the importance of these dimensions. First, since we did not ask participants to directly evaluate category membership, but rather allowed their similarity judgments to guide the results, it's possible for us to discover clusters in subjects' perception that do not match the nominal definitions of the clusters. For instance, my subjects clearly placed chemistry on the natural side of the axis, closest to biology and medicine, though one might have thought a priori that it involved highly artificial products and belonged with the engineering majors. Likewise, subjects placed all of the social sciences on the humanities end of the spectrum, clearly distinct from the “hard” sciences- even psychology students do not appear to regard it as a science in the same way that chemistry and biology are.

Second, rather than just dividing majors into categories, our method assigned numerical values to the extent to which they fit each category. This allows us to say, for instance, that computer science was seen as on the far extreme of the artificial dimension, while mathematics was more moderately artificial and astronomy was neither artificial nor natural in the eyes of subjects. We can therefore predict that whatever gender differences exist in interest in these majors, they are likely to be more extreme for the most extreme exemplars of each dimension.

Third, we have information for each subject about their individual map and preferences for different parts of the map. Though we did not find significantly different maps for men and women, we did find significant individual variation in how much subjects valued high or low majors in each dimension. So even where, overall, women preferred natural over artificial majors, we can point to some women who had the opposite pattern, preferring artificial majors over natural.

The preceding studies have told us how subjects divide the majors available to them, and also how men and women's preferences for those majors fall. In the studies to come, we will investigate the reasons why subjects may prefer some majors (or groups of majors) by directly examining the stereotypes they hold about those majors, and comparing those stereotypes to their self-image. This may allow us to understand more about why we found the preferences we did in this study.

Chapter 3 Stereotypes of Computer Science

The previous study having established that women do seem to have a more negative attitude toward majors in the scientific and artificial quadrant, I would like to focus particularly on one of the majors farthest toward the outer edge of that quadrant- both highly scientific and highly artificial- in order to understand the reasons why it is unpopular with women. Ultimately it will be useful to explore this question for a variety of majors, but our multidimensional map has confirmed that computer science is a useful place to begin.

In the introduction, I have argued for a self-prototype matching model of choice of majors. This approach assumes that, in deciding whether they are interested in computer science, both male and female students are comparing themselves to a stereotype driven image, or prototype, of computer science students. The prediction is that students who see themselves as similar to that prototype will be interested in computer science, while those who see themselves as different from it will not be interested. This does not necessarily argue that similarity causes interest, but it can at least show that interest and similarity relate in a way consistent with that hypothesis.

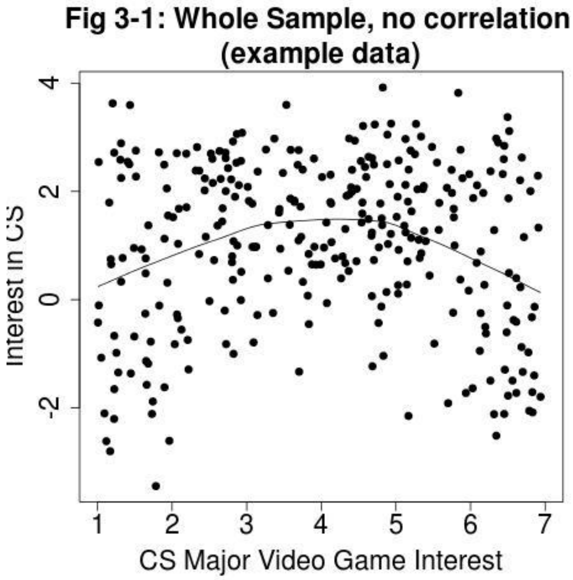
In this chapter, I would like to lay the groundwork for an approach to studying interest in computer science under this model. The advantage of a model which considers self-prototype match, rather than just examining impressions of computer science, is that it does not treat participants as interchangeable. For any given stereotype about computer science, some participants may have identities congruent with that stereotype, in which case it would likely increase their interest in computer science. Other participants may see it as incongruent with their identities, and thus it would decrease their interest. Asking after the match between CS stereotypes and self-perception, rather than just about CS stereotypes, allows us to directly observe such effects.

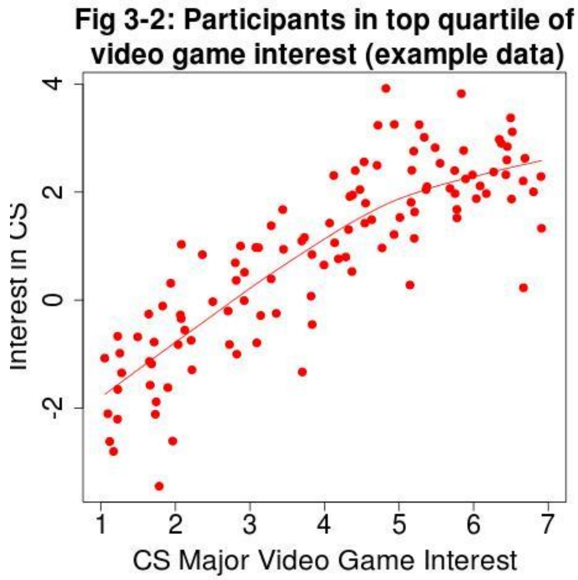
This also allows us to take a nuanced approach to the effect of gender. For instance, if we wished to study the effect of the stereotype that computer science majors work alone and do not collaborate with one another, a study of CS perceptions only might find that women who hold this stereotype are less interested in computer science, while for men there is a smaller effect, and conclude that for women (on average) the perceived solitude of CS work is a problem. But in a self-prototype match model, we do not need to restrict ourselves to considering only the average female participant. We can identify the subset of participants to whom solitary work seems to be a detriment, and those to whom it is a benefit or neutral. If it happens that there are more women who see solitary work as a detriment, then our model can account for gender differences without losing sight of the minority of women who prefer solitary work. But it is not gender per se that we are studying; rather, we are studying the effects of self-perception, and gender has an effect on self-perception. The advantage of this, in the long run, is in the ability to make more detailed predictions about who will be affected by which intervention, rather than relying on gender or other categories as proxies.

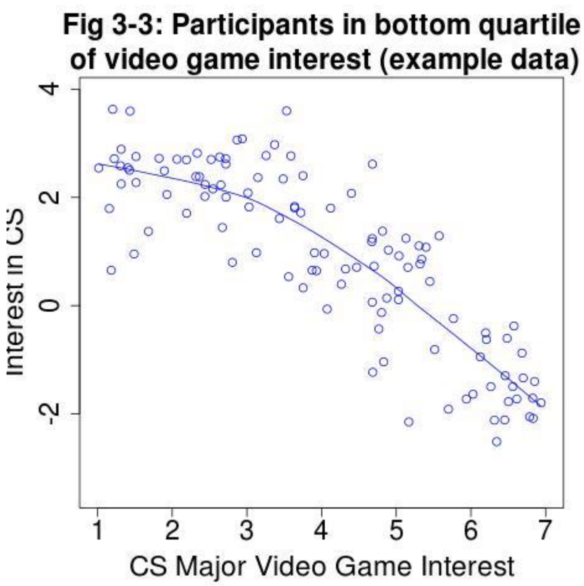
To sum up our example concretely: among participants who themselves prefer solitary work, a belief that computer science students prefer solitary work is likely to increase their interest in the major. On the other hand, for participants who do not enjoy solitary work, that interest is likely to decrease. This will impact the gender parity of the field to the extent that men are more likely than women to match the stereotype of computer scientists when it comes to solitary work.

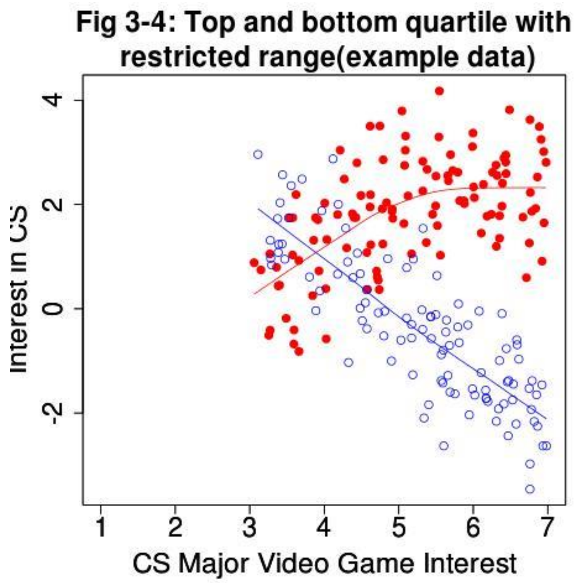
Thus, we expect the effect of stereotypes about computer science to vary based on participants self-description. Increasing perception that computer science majors enjoy solitary work (for example) will have a positive relationship to interest in CS until it reaches the participants self-rated level of solitary work interest, and then have a negative relationship once that level is exceeded. So each subject is expected to have a nonlinear relationship between any given stereotype about CS and interest in CS- a downward-facing parabola. The proposed method differs from standard moderation in that it expects a nonlinear relationship between trait ratings and interest; for example your interest in CS decreases when you believe computer scientists are either much more interested in solitary work or much less interested in solitary work than you are.

Thus, in the present study, I asked participants to rate themselves and a fictitious computer science student on a variety of stereotype dimensions, and I am looking for a nonlinear relationship between interest in computer science and stereotypes which varies based on self-stereotype. Since the purpose of this chapter is to show that this approach to the question of who is interested in computer science is useful and valid, we are focused on showing the interplay between perception of CS and self-perception in predicting interest, rather than focusing on the effect of any particular stereotype. We are interested, though, in the ability of this method to find gender differences, which might appear either in how similar male or female participants felt to computer scientists, or how important similarity was in predicting their interest.









Figures 3-1 through 3-4 indicate the pattern of responses we are looking for, using simulated data (see Appendix B for details). There is no correlation between interest in computer science and the stereotype that computer science students enjoy solitary work (Fig 3-1), so a researcher examining only those variables would find nothing of use. However, significant trends in opposite directions emerge when we consider either only those participants who themselves enjoy solitary work or those who do not (Figs. 3-2 and 3-3; participants divided by tercile). Figure 3-4 overlaps the plots for high and low self-rated enjoyment of solitary work in a sample with restricted range, and thus for high self-rated participants demonstrates the nonlinearity described above. The trendlines in these graphs, and in all subsequent ones in this chapter, were computed using a LOESS local running average model.

Thus, my hypothesis for this study is that participants who see themselves as more similar to computer science students will also be more interested in computer science. I am also interested in determining whether the gender of the target plays a role in this decision, and whether particular stereotypes are more effective as determiners of interest than others, but these goals are exploratory.

Method

Participants

Data were collected from 278 (192 female) University of Washington students enrolled in elementary psychology classes using an online survey. Seventeen participants identified as Black / African American, 121 Asian / Asian American, 23 Latino / Hispanic, 6 Native American / American Indian, 143 as White, and 12 as other. Fifty four participants had taken at least one computer science class.

Materials

Stereotype Items. In order to see the similarity between self-perception and prototype-perception, we need some axes along which participants can be different or similar. In theory, any ways in which participants self-images differs from their CS image would be sufficient. But there may be some ways which are particularly salient to participants, or which are given particular weight. It seems worthwhile to me to cast a relatively wide net in order to capture such constructs, so I have attempted to survey participants on a wide variety of potential CS stereotypes.

First, I generated a list of stereotype items based on literature specifically about computer science. The goal was to sample a wide variety of constructs which various researchers have hypothesized might be components of our stereotypes about computer science, or about science and technology in general. Where the researchers in question provided specific items to measure these constructs, I have adapted them to fit the format of this task. Where they provided more general speculation, I constructed items on my own. 45 items were generated in this way.

- Questions related to agentic goals- that is, that computer scientists are motivated by individual success, the mastery of difficult skills, and the desire to demonstrate individual competence. (Diekman et al., 2010)
- Questions related to communal goals- that is, that computer scientists are not motivated by the desire to care for or work with other people (Diekman et al., 2010).
- Questions related to socialization; either traits describing desire/ability to socialize, lack of success in socializing, or antisocial traits such as arrogance. We expect the first to be negatively related to computer science prototypes, and the second to be positively related (Hannover & Kessels, 2004; Cheryan et al., 2013). The third was not found to be a significant stereotype of science students by Hannover and Kessels, but we include it anyway for the purposes of exploration.

- Questions relating to a singular focus on technology[6] outside the classroom. The stereotype here is that to be a computer science major, one must have a life revolving around technology, including exposure to programming early in life, and technology-related hobbies outside of academic work (Margolis & Fisher, 2002; Cheryan et al., 2013).
- Questions relating to the idea that computer science majors are apart from other students, forming an exclusive club whose members are not able to relate to those outside it (Margolis & Fisher, 2002)
- Questions relating to stereotyped interests like video games, science fiction, and comic books (Cheryan et al., 2009) which are often seen as associated with computer science. These will be labeled as "geeky".

In addition, I added some items on an exploratory basis, relating to the following constructs: that CS majors do repetitive work with little creative input; that CS majors are uninterested in animals and nature; that CS majors dislike physical activity; that CS majors work in basements; that CS majors are unemotional. A total of 8 questions addressed these stereotypes.

Finally, I wanted to examine the effect of similarity on some traits which are not generally seen as particularly stereotypical of computer scientists. Here, we might expect idiosyncratic effects between participants if they disagree strongly on the extent to which CS majors have a certain trait. Capturing this might show us how variations in stereotypes may

---

6 This will be described as "tech-focused" in tables for the sake of brevity, but readers should remember that it indicates a singular degree of focus on technology both inside and outside the classroom, not just facility with technology. It is unlikely that one could succeed in computer science without a high degree of technical aptitude, but there is no reason it needs to pervade leisure time in addition to class and homework time, or why it needs to have been an interest from an early age. The latter two are the construct being measured.

produce variations in interest. But in doing this, it was important to note that since computer science is male dominated, male stereotyped traits might be seen as closer to CS majors than female-stereotyped traits. As such, I opted to use the 20 gender-neutral items from the Bem Sex Role Inventory (Bem, 1974) as a list of common adjectives without gender bias. Including these 20 items brought the total number of traits to 73.

Note that these items do not correspond exactly to the dimensions derived in chapter 2, since those were evaluations of the fields of study themselves and not of the people who study them. Based on that study, we expect participants to see computer science as scientific and artificial; the questions about whether the prototypical student is singularly focused on technology may be seen as a way of asking whether this classification of the field is also applied to the people who study it as a stereotype.

Interest in Computer Science. We gauged participants' interest in computer science using three items: “I could see myself majoring in computer science”, “I would enjoy taking a computer science class”, and “I would do well in a computer science class.” These were related with an alpha of .89, so they were taken to jointly measure interest in computer science. We also directly measured participants’ perceived similarity to the target with the item “I am similar to the computer science major I imagined.”

Procedure

Participants were asked to visualize a fellow University of Washington student, knowing that this person's major was computer science, and then to imagine an interaction with them. In order to encourage more vivid visualization, participants were asked to describe the person they pictured, the kind of work he or she did, what it would be like to interact with him or her, and what his or her attitude toward the participant would likely be. After answering these open ended questions, participants proceeded to rate the target on a variety of traits, and then to rate themselves on the same.

At the outset of this study, I was uncertain whether to expect different stereotypes for male or female computer science students. Thus, I randomly assigned participants to imagine either a male or female computer science student. The hope was that this would allow us to explore the differences if any between the male and female stereotypes, which would inform how we present future stimuli.

Results

Some items were both conceptually and empirically related to each other, and thus were combined into scales. Specifically, scales relating to agency, communality, sociability, social isolation, tech-focus, unusualness (i.e. separation from non-CS students), and geek interests. In other cases, only one or a few items related to a particular construct, in which cases responses to these items were analyzed separately.

Similarity between ratings of self and target was defined using a multidimensional Euclidean distance model, where distance is computed as the sum of the root squared difference between self-rating and target rating for all items, i.e. $\sqrt{\sum (T - S)^2}$. Thus, a large difference between self-rating and major rating produces a large distance, which is implicitly a low similarity, whether the difference itself is positive or negative. This produces a difference rather than a similarity metric; in order to make this more intuitive, when displaying specific values we have subtracted each distance score from 6 (the maximum value) in order to produce a similarity measure.

Gender of target

Generally, the gender of the target appeared to make little difference in participants' responses. A 2x2 ANOVA (target gender by participant gender) predicting interest was computed to examine this effect. Table 3-1 summarizes its results. There was a significant main effect of participant gender, with women less interested in computer science than men, $F(1,274)=65.49$, $p<.00001$. There was no main effect of target gender, $F(1,274)=.002$, $p=.81$ and

	Participant Gender	
Target Gender	M	F
M	3.61(1.72)	2.15(1.39)
F	3.71(1.88)	2.06(1.33)

Table 3-1 Interest in CS by gender

	Participant Gender	
Target Gender	M	F
M	3.87(0.6)	3.77(0.62)
F	3.91(0.66)	3.65(0.58)

Table 3-2 Target similarity by gender

no interaction, $F(1,274)=.46$, $p=.50$. Thus, male and female targets were associated with the same level of interest in CS for male and female subjects. Our power to detect a Cohen's *d* effect size of .5 (a moderate effect) with the current sample is .99, so we can be reasonably confident that if there is truly a difference for this population between interest in CS when a male or female target is presented, it is small.

Likewise, the perceived similarity of the target to the participant did not differ significantly by target gender (based on comparisons of trait ratings). A 2x2 ANOVA was again computed to test this effect, summarized in Table 3-2. There was a significant main effect of participant gender, with women less similar to either target than men, $F(1,270)=5.32$, $p=.02$. There was no main effect of target gender, $F(1,270)=.76$, $p=.38$, and no interaction, $F(1,274)=1.03$, $p=.31$. Thus, gender of target did not play a significant role in determining participants' perceived similarity.

Taken together, the previous findings suggest that target gender was not important in participants' decision process, and I will not examine it in much greater detail. Ratings of female targets differed from ratings of male targets only in a few items: perceived sexism, preference for socializing with men, messiness, and playing of video games. See Table 3-5 for a summary of these results. As such, I will collapse across conditions in the analysis below, as including condition does not make any difference in our results.

	1	2	3
1. interest in CS	1	0.54	0.42
2. self-rated similarity		1	0.43
3. profile similarity			1

Table 3-3. Relationship between interest in CS and similarity to CS majors

Interest and similarity

We will next examine the relationships among interest in computer science, self-rated similarity to computer science, and similarity based on self and target trait ratings. The correlations are presented in Table 3-3. There was an overall effect of similarity between self-rating and target-rating on interest in computer science, with high similarity predicting high interest. Participants' perception of similarity to the target correlated with both the similarity of their self and target ratings and with their interest in computer science. Interest and both measures of similarity are strongly related to each other.

Gender of participant moderated the relationship between perceived similarity and interest (B=-.25,SE B=.11, p=.02), but this moderation was only a trend for similarity of ratings (B=.49,SE B=.29, p<.10). In both cases the moderation indicated that similarity was of slightly more importance in predicting the decisions of male participants.

In order to evaluate which stereotypes about computer scientists may play a particularly weak or strong role in determining participants interests, correlations were computed between interest and the items (and scales as defined above). Table 3-4 gives the correlations. Figures 3-5 through 3-8 depict the distributions of interest in CS and CS stereotyping for participants high and low in self-stereotyping for several of the more significantly predictive categories, to be compared with the example graphs presented earlier. A full set of graphs of this kind can be found in Appendix D.

Stereotype Class	r	sig
Communal	-0.33	<.0000001
Isolated	-0.07	ns
Sociable	-0.34	<.0000001
Tech-Focused	-0.50	<.0000001
Unusual	-0.29	<.000001
Geek	-0.37	<.0000001
Agentic	-0.32	<.0000001

Table 3-4 Relationship between similarity interest in CS on each stereotype scale

	Rating of computer scientists				Rating of oneself...			
	Male	Female	t(df)	p	Male	Female	t(df)	p
Communal	3.91	3.74	1.43(276)		5.25	5.62	-2.50(276)	<.05
Isolated	3.49	3.67	1.13(276)		3.38	3.37	0.03(276)	
Sociable	3.43	3.67	0.68(276)		4.89	5.02	1.06(276)	
Tech-Focused	5.52	5.39	1.27(276)		3.85	2.94	6.95(276)	<.0000001
Unusual	4.66	4.56	.069(276)		3.10	2.70	2.68(276)	<.01
Geek	5.28	4.77	3.66(276)	<.001	4.76	4.43	3.57(276)	<.001
Agentic	5.47	5.54	0.76(276)		4.76	2.8	6.25(276)	<.0000001

Table 3-5 Differences in stereotype ratings by subject gender

Figure 3-5 Effect of CS stereotypes on interest for participants with high or low self-rating in communality diamonds=low, circles=high

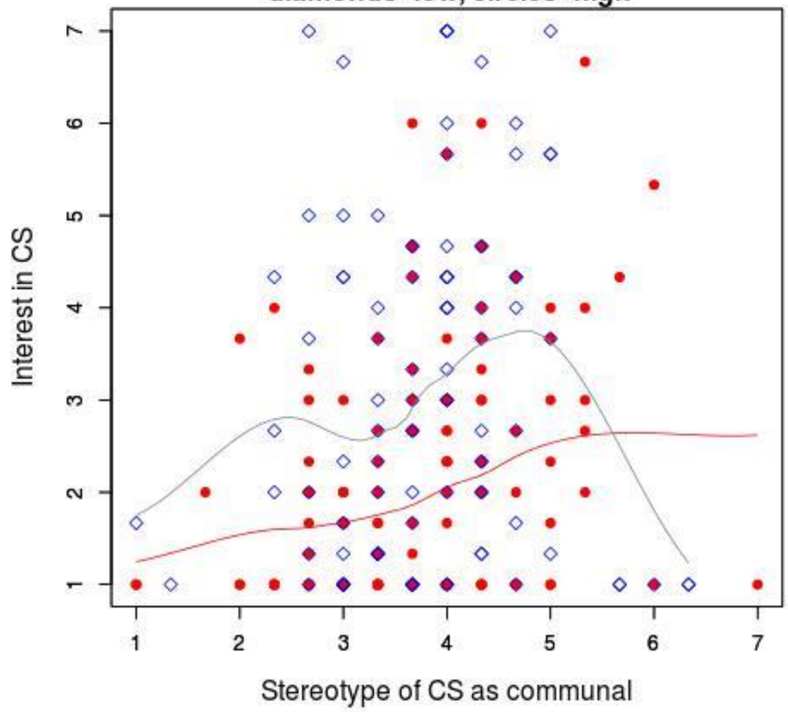


Figure 3-6 Effect of CS stereotypes on interest for participants with high or low self-rating in sociability diamonds=low, circles=high

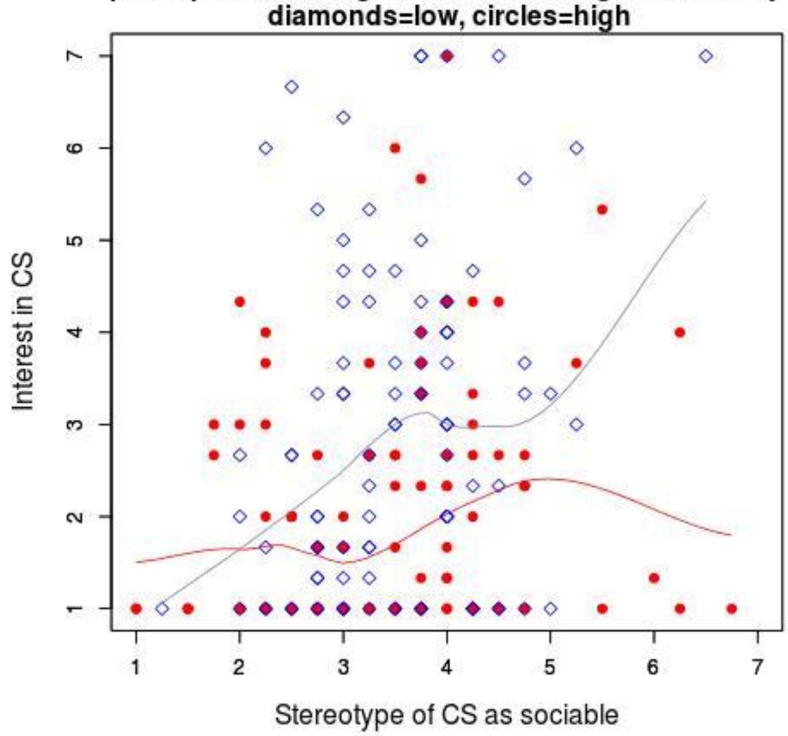


Figure 3-7 Effect of CS stereotypes on interest for participants with high or low self-rating in technology-focus diamonds=low, circles=high

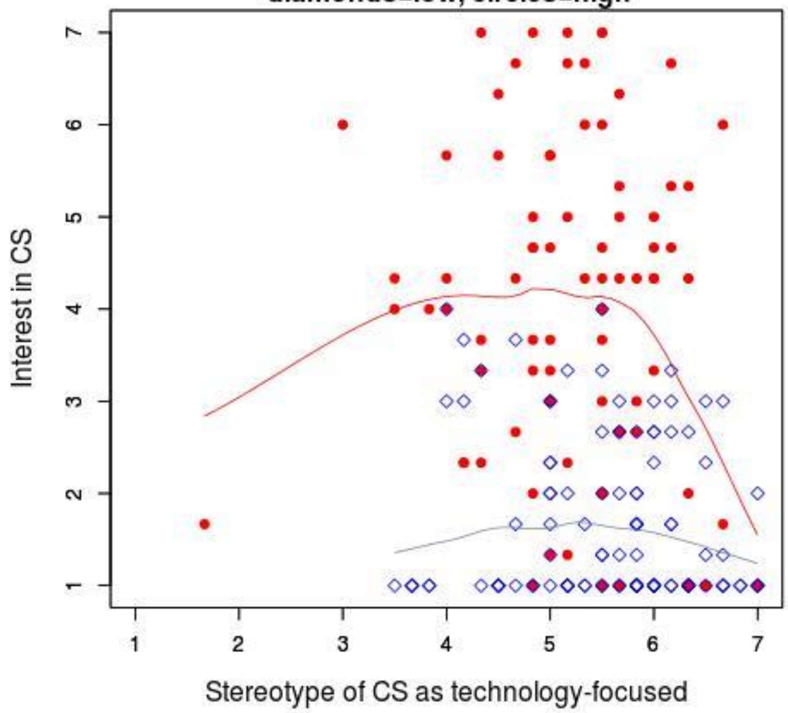
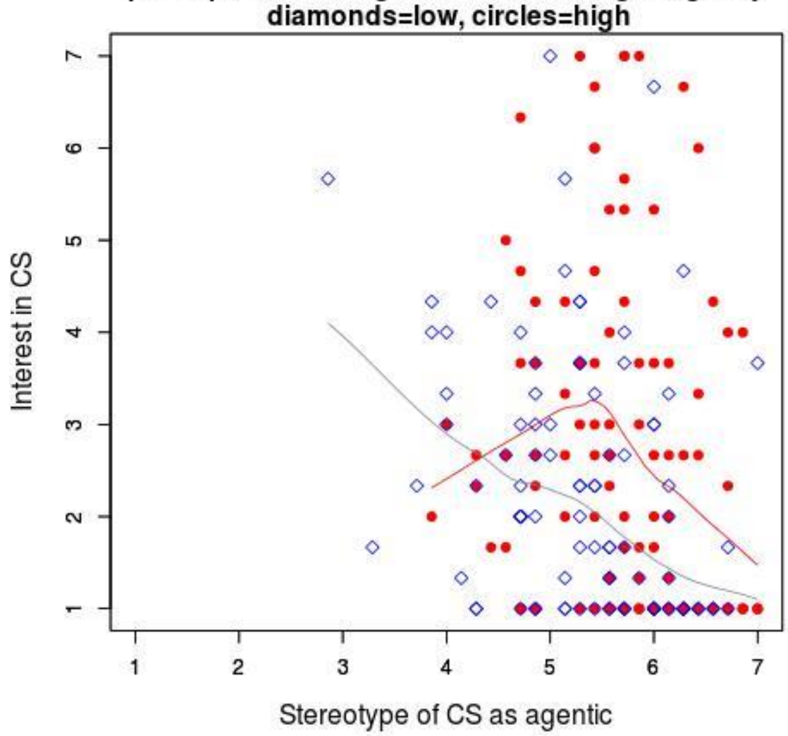


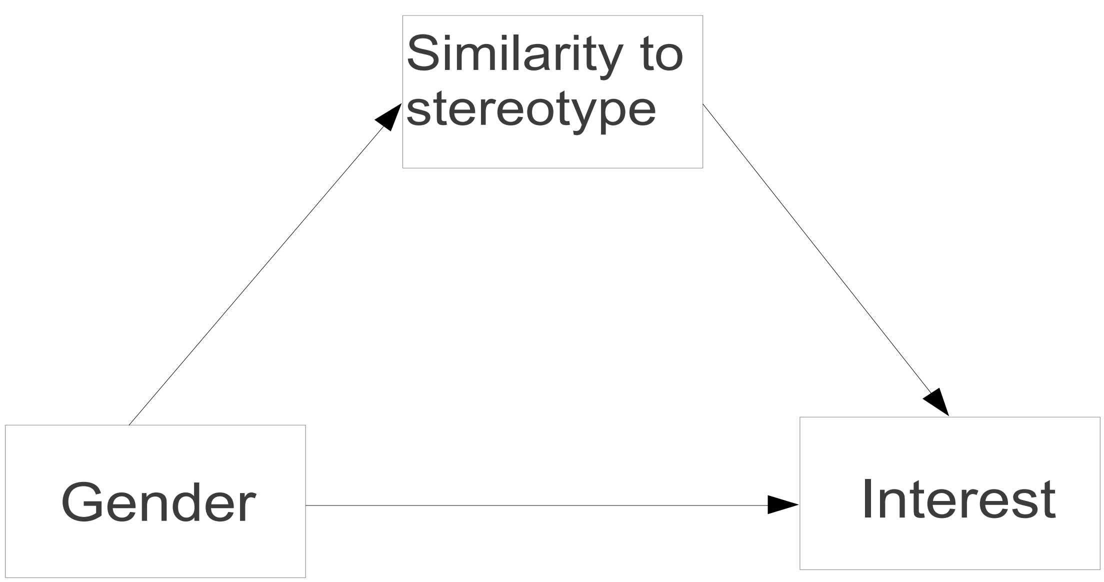
Figure 3-8 Effect of CS stereotypes on interest for participants with high or low self-rating in agency diamonds=low, circles=high



Stereotype	Rating of computer scientists Male	Female	t(df)	p	Rating of oneself... Male	Female	t(df)	p
spends a lot of time playing video games	5.01	4.69	1.53(276)		4.12	2.16	8.29(276)	<.0001
dislikes feminine things	3.73	3.81	-0.41(276)		3.23	2.11	5.86(276)	<.0001
is moody	3.26	2.81	2.35(276)	<.05	2.76	2.09	4.13(276)	<.0001
is tactful	2.59	2.16	2.59(276)	<.05	3.47	2.75	3.97(276)	<.0001
is content to do repetitive tasks	5.02	5.19	-0.86(276)		4.23	3.37	3.72(276)	<.001
is unusual	4.83	4.47	1.83(276)		4.65	3.82	3.94(276)	<.001
doesn't mind messy environments	4.13	4	0.79(276)		3.94	3.2	3.88(276)	<.001
is interested in science fiction	5.3	5.23	0.39(276)		4.17	3.44	2.99(276)	<.01
likes outdoor activities	2.9	2.79	0.65(276)		4.76	4.15	2.81(276)	<.01
is arrogant	3.84	3.8	0.17(276)		3.17	2.55	3.14(276)	<.01
is secretive	3.47	3.11	1.98(275)	<.05	4.09	3.55	2.9(276)	<.01
often works in enclosed environments	4.71	4.86	-0.72(276)		2.73	2.21	2.48(276)	<.05
pays less attention to appearance	4.26	4.61	-1.63(276)		3.55	3.09	2.18(276)	<.05
is reliable	4.78	5.24	-3.08(276)	<.01	5.36	5.65	-2.03(275)	<.05
is friendly	4.72	4.68	0.23(276)		5.53	5.87	-2.26(276)	<.05
more socially comfortable with men	2.65	2.56	0.52(276)		3.12	3.65	-2.13(276)	<.05
is conscientious	3.01	2.78	1.17(276)		2.31	1.92	2.18(276)	<.05
is unpredictable	3.67	3.47	1.03(276)		3.2	2.72	2.43(276)	<.05
is inefficient	4.55	4.68	-0.86(276)		5.49	5.81	-2.17(275)	<.05

Table 3-5b Differences in stereotype ratings by subject gender (significant comparisons only- see appendix)

Figure 3-9 Mediation Diagram; Gender predicts interest in a major via similarity to stereotypes of that major.



Predicting gender differences in interest

Male and female subjects rated themselves differently in a number of traits, which are given in Table 3-5. Differences in self-ratings suggest that men and women might differ in interest in computer science because the different self-ratings correspond to different degrees of interest in computer science. Table 3-6 shows Sobel mediation tests for each of the scales; similarity mediated the relationship between gender and interest in computer science when that similarity related to agency, communality, geekiness, and focus on technology, but not otherwise. Figure 3-9 depicts the mediation model used in these tests.

Stereotype	c	c'	Sobel Z
Communal	-0.44	-0.39	-2.25*
Isolated	-0.44	-0.44	0.20
Sociable	-0.44	-0.4	-1.93^{t}
Tech-Focused	-0.44	-0.32	-3.79****
Unusual	-0.44	-0.41	-1.46
Geek	-0.44	-0.37	-3.39****
Agentic	-0.44	-0.39	-2.26*

Table 3-6 When does similarity mediate the relationship between gender and interest in CS?

Since our purpose in this chapter is in large part to explore this new method, it is useful to compare it to simpler methods and show where it gains. The simplest and most common approach to studying the effect of a stereotype would be to examine the effect of holding that stereotype on a dependent variable. My current approach adds two things to this, mathematically: it takes into account the participants' self-description on the same stereotype items as a predictor, and it introduces a nonlinear component by taking the difference between self and target ratings. Table 3-7 shows the increase in R2, first when adding self-ratings, and then when adding the nonlinear component, on each of our measures.[7]

	R^2 for target rating	ΔR^2 from self rating	ΔR^2 from distance model
Communal	.03	.11*****	.01
Isolated	.00	.02*	.08***
Sociable	.10	.15***	.15
Tech-Focused	.06	.23*****	.01
Unusual	.04	.06****	.01
Geek	.00	.17****	.19**
Agentic	.04	.02**	.05***

Table 3-7 Additional variance accounted for by considering self-ratings and nonlinear distance effects in predicting CS interest

Discussion

Thus, the self-prototype matching hypothesis was essentially supported. Across categories of item, there was a significant relationship between similarity to a prototypical CS student and interest in CS, and similarity mediated the gender differences in CS interest. Moreover, the current methodology was able to distinguish specific sorts of similarity which appear to have been especially influential in approaching this decision, for instance communality, agency, technology-focus, and “geeky” interests such as video games. Since this

7 The communality stereotype cluster included one item of particular relevance to Lippa's (1998) “people” dimension, “aspires to help or care for others”. This item was endorsed more strongly by women (M=5.73, SD=1.26) than by men (M=5.38, SD=1.52), $t(139.35)=-1.86$, p=.06. This question significantly mediated gender differences in interest, c=-.44, c'=-.38, Z=4.85, p<.0001. Thus, Lippa's theory that interest in person-related careers is particularly high in women is borne out by this data.

model has been relatively successful, we will refine it further in future studies.

Before moving on from these results, I wish to pause for a moment and discuss the traits (comic books, science fiction, and video games) labeled by this study as “geeky”, since the use of this slang term may seem an arbitrary label to some readers. Our use of the label is justified by its use by subjects; the word “geek” and the closely related “nerd” appear 71 times among our 278 open-ended responses, so the word is clearly on our subjects' minds. We do not at this point know why participants are inclined to connect these particular interests and this slang term to computer science, or whether some deeper construct underlies this stereotype, but both the current study and previous research (Cheryan et al., 2009) support that it is a relevant category to consider.

It appears to have made very little difference to our participants whether the student they were imagining was male or female. It's possible, of course, that this is just a failure of manipulation, and that the stereotypes of male and female computer science students are quite different. But it's also a reasonable conclusion that for the most part, computer science stereotypes overwhelmed gender stereotypes on these topics. One might have expected, all else being equal, that women would consider themselves more similar to a female than to a male target, but this did not happen. This outcome is in line with some research, suggesting that the gender of a stereotyped STEM student is irrelevant from the point of view of identification (Cheryan et al., 2011). Accordingly, we will leave aside the question of target gender in future versions of this study.

The main importance of this finding is to show that we should expect the effects of stereotypes to depend on the self-image of the person perceiving them. So a manipulation changing stereotypes about a major would be likely to produce opposite effects in subjects with high and low self-ratings in those stereotypes. Participants similar to the altered stereotypes would be more interested, but those who had been similar to the prior stereotypes would be less interested. Likewise, a measurement of the effect of a stereotype will capture the effect of that stereotype on people of average self-rating in that stereotype, but may not be reflective at all of those whose self-rating is either above or below average. Thus, researchers would benefit from explicitly measuring participants self-perception, so that groups of participants who see themselves differently can be analyzed separately.

Though this study was useful as a first approach, it left some matters unsolved. First, in the name of casting a wide net, it often asked relatively few questions in assessing constructs of interest, which runs the risk of producing unreliable answers, and makes it difficult for us to determine what aspect of each question was relevant. Our labels for the question categories are at this point quite speculative. We know that the answers to some sets of questions were predictive of interest, but we don't yet know what the meaning behind those sets is.

Second, we have presumed that every participant gives the same weight to similarity in every category. That is, though people may differ in how close to CS they are, everyone is treated as giving the same weight to similarity. Unfortunately, since each person only gave us a single outcome variable, interest in computer science. So we don't have enough information to examine per person effects. My next study will ask participants to rate multiple majors so we can see the degree to which different people value different types of similarity differently.

Chapter 4 A Multilevel Model of Academic Choice

So far, we have discussed the structure of stereotypes about computer science, and the role that self-perception plays in students' reactions to them. In this chapter, we examine how stereotypes and student reactions to them vary across majors, in order to derive a sense of the general trends. We will look at the within-person correlations between stereotypes of majors and interest in them to derive a response pattern for each subject.

Preference for majors is fundamentally an intra-personal process, even though we've been studying it interpersonally so far. That is, when we talked about a correlation between similarity and interest, what we meant is that people with higher similarity to computer scientists had higher interest. But the real phenomenon of interest is whether any given person is more likely to be interested in fields to whose members he or she feels similar. The extent to which this is true might well vary between people. In essence, in the previous study we have aggregate across subjects in order to get an approximation for how individual subjects would behave if we varied their level of similarity to computer scientists; here, we are looking within each person to see what their interest is at varying levels of similarity to various STEM majors.

In doing this, we are still interested in the same question as when we focused on computer science. There are three reasons to broaden our focus to include a greater variety of majors. First, as noted, this approach allows us to compare participants to one another directly. It seems likely that there are multiple reasons why students might be interested or deterred from interest in computer science; one reason might be powerful for one group of students, while being a non-factor for another group. And this has the potential to sabotage efforts to increase diversity, since efforts that appeal to one group may be neutral or even negative for other groups.

If we know the spectrum of things that students want in a STEM major, we can make a better effort to show (or create) that spectrum in our classrooms. The ideal message is not that CS conforms to the priorities of one particular group of students over another, it is that CS is diverse enough to be valuable to students with a wide range of priorities.

Second, we have so far discussed CS as anomalous, but presumably it is not entirely unique; we expect that analogous trends should be present in some other STEM fields. But we also know that not all STEM fields show such a gender bias to the same extent. Thus, looking at fields which do have problematic gender dynamics as well as those which do not allows us to study both the presence and the absence of the problem. It is possible that some of what makes biology (for instance) attractive to women could be adapted to less balanced STEM fields, or comparison to biology could make the problems present in computer science clearer.

Thirdly, this dataset affords us the opportunity to look at how stereotypes about STEM majors vary according to the personal characteristics of students. In the previous study, we restricted ourselves to concluding that men and women did not differ in their stereotypes of computer science. But in this study, we can examine which majors have relatively consistent or relatively variable stereotypes, since we can compare their stereotype variation to one another. We can also compare the stereotypes of different groups of students, as we did in the previous study, with considerably more precision since we have more datapoints per subject.

Statistically speaking, if we want to consider how subjects differ from each other in response pattern, we need more than enough data points per subject to constitute a pattern. Conceivably we could ask students to compare themselves to multiple kinds of CS student (e.g. different sub-specialties) in order to get more than one data point, but it seems likely those prototypes would have high overlap. But by combining this goal with the goal of learning how CS relates to other majors, we can solve both problems at once.

Method

The core of this study is identical to that of the previous study, except that here, each participant evaluates a series of imagined students, each of whom is in a different major. (In this study, the gender of the target was not specified). Each target is rated on a variety of traits, and participants rate themselves on the same traits, as well as indicating their interest in each major. So for each participant, we have a profile describing the self, a profile describing a prototypical member of each major, and an interest rating in the major itself.

This study was run separately on two populations, in order to reduce the risk of finding effects applicable only to a particular sample. A population of 89 users (52 female) on Amazon Mechanical Turk (MTurk) were asked to rate four majors (CS, Physics, Biology, and Chemistry) in exchange for monetary compensation, and a population of 172 University of Washington psychology undergraduates (83 female) rated nine majors (CS, Physics, Biology, Chemistry, Medicine, Digital Art, Oceanography, Environmental and Electrical Engineering) in exchange for course credit. The Mturk sample included 7 participants who identified as Black/African American, 8 as Asian/Asian American, 4 as Latino/Hispanic, 2 as Native American/American Indian, and 70 as White. Their mean age was 34.5 years old. The UW sample included 6 participants who identified as Black/African American, 91 as Asian/Asian American, 11 as Latino/Hispanic, 1 as Native American / American Indian, 68 as White, and 4 as other. Their mean age was 19.2 years old.

The majors selected were STEM majors in accordance with the science/culture axis described in the previous chapter, and were chosen to span the natural/artificial axis. “Digital” art was included in order to have a representative of the cultural half of the axis for the sake of comparison, with the word digital intended to justify its inclusion in a list of technical majors in the eyes of subjects. I will present the results of these side-by-side and refer to them as Sample A and Sample B where necessary.

Hypotheses

1) As before, self-rated similarity to a prototypical student will predict interest in that student's major. I expect this to be true for each of the majors studied.

2) Interest in each major will also be predicted by what we might call stereotype similarity- the degree of overlap between stereotypes about that major and one's self-perception. (Hereafter I will just refer to this as similarity, without prefix, as it is the main kind of similarity currently being studied.)

3) For some majors (but not for all) there should exist a significant gender difference in interest.

4) Where it is present, similarity should mediate this gender difference in interest, at least partially.

5) For at least some traits, there will be variation between subjects in how important that trait is in determining interest. In other words, we expect to find traits X such that for some subjects, interest in a major strongly depends on its level of X , while for others X makes less or no difference.

6) This approach, in which self-ratings are taken into account as well as stereotypes, will produce a better-fitting model of participants' choices than examining stereotypes alone would.

Exploratory Goals

1) I want to examine whether majors with very different gender ratios differ in terms of the stereotypes I am measuring here.

2) I want to identify which traits are important predictors of interest for subjects in general, and which are important for some subjects but not others. In particular, it would be interesting to find clusters of traits whose importance goes up or down together.

3) Does similarity in some stereotypes more effectively mediate gender differences in interest than in others?

4) Is similarity more important in some majors than others?

Another open question is raised by my 4th hypothesis above, which proposes that gender differences in interest may be mediated by similarity. As noted in the previous chapter, if this is the case, gender differences might emerge in several ways. Since similarity here is a comparison between self-rating and various stereotype ratings, then if similarity is a mediator 1) women and men might have different self-ratings, 2) women and men might have different stereotypes, and 3) women and men might give weight to different similarity on different traits. The current study will be in a position to test all three of these separately.

Many of the preceding questions will be answered using multilevel modeling (MLM). I will generally avoid too much technical discourse about this method. Like a regression, a multilevel model uses several predictor variables to estimate an outcome variable, assigning each a coefficient. A predictor has a significant effect when its coefficient is different from zero. Unlike normal regression, a multilevel model takes into account that its data is organized in groups (in our case, each subject is a "group"). In the current context, we will typically have a coefficient for similarity for each subject for each major, and we'll track how the coefficients vary with major. We can also look for individual differences, in that the pattern of variation of the similarity coefficient will be different for different subjects. We will occasionally look directly at these coefficients, but more often we will compare alternative models and see which fits the data better overall.

Results

Hypothesis 1 and 2

Similarity should correlate with interest between subjects. Table 4-1 indicates the correlations among interest, self-rated similarity, and similarity of stereotype profiles. In all of the selected majors in both samples, self-rated similarity to students in a given major was significantly correlated with interest in that major. For all majors in the Mturk sample and for all majors except Medicine, Biology, and Chemistry in the UW sample, rating oneself as similar along the stereotype dimensions to a target student was also significantly correlated with interest in that student's major. This supports the central idea that participants prefer majors where they see the students as more similar to themselves. The correlations between these two types of similarity were moderate (*r* between .14 and .47).

Partial correlations between profile similarity and interest, controlling for self-rated similarity, are presented in Table 4-1. The fact that these correlations are small suggests that participant’s stereotype ratings are not giving us additional information about their similarity which was concealed by their self-rated similarity judgments. Instead, it is giving us the opportunity to clarify the ways in which subjects see themselves as similar.

Hypothesis 3

There should be gender differences in interest between men and women. A series of independent samples t tests were conducted to determine the overall differences between male and female subjects in the sample in terms of interest in the various majors. Their results are summarized in Table 4-2. In the UW sample men were significantly more interested than women in Physics, Computer Science, and Electrical Engineering, and significantly less interested in Digital Art. As predicted, men were more interested in some but not all STEM majors, and Computer Science in particular was a point of difference. Mturk participants showed significant differences only in interest in Physics. Mturk participants were also more interested in STEM majors in general (M=3.77) than UW participants were (M=3.63, t=2.95, p<.01).

UW Sample

Major	profile similarity	self-rated similarity	profile/self-rating correlation
Physics	0.21 **	0.56 *****	0.22 **
Biology	0.14	0.58 *****	0.17 *
Computer Science	0.3 ****	0.65 *****	0.35 *****
Chemistry	0.13	0.53 *****	0.14
Environment Eng.	0.2 **	0.57 *****	0.18 *
Electrical Eng.	0.34 ****	0.63 *****	0.31 ****
Oceanography	0.18 *	0.49 *****	0.12
Digital Art	0.22 **	0.54 *****	0.39 *****
Medicine	0.1	0.56 *****	0.14

Mturk Sample

Major	profile similarity	self-rated similarity	profile/self-rating correlation
Physics	0.51 *****	0.67 *****	0.45 ****
Biology	0.28 **	0.57 *****	0.23 *
Computer Science	0.39 ***	0.69 *****	0.37 ***
Chemistry	0.38 **	0.55 *****	0.37 **

Table 4-1: Correlations between interest and similarity

UW Sample

Major	Male	Female	t(df)
Physics	3.73	2.8	4.08(165.85)****
Biology	3.97	4.31	-1.5(167.89)
Computer Science	3.82	2.78	4.63(163.88)*****
Chemistry	3.89	3.65	0.97(167.41)
Environment Eng.	3.28	2.98	1.65(168.94)
Electrical Eng.	3.61	2.75	3.63(165.94)***
Oceanography	3.32	3.51	-0.98(164.99)
Digital Art	3.03	3.77	-3.49(160.81)***
Medicine	4.18	4.22	-0.14(166.29)
Mturk Sample			
Major	Male	Female	t(df)
Physics	4.03	3.09	2.22(70.45)*
Biology	4.3	3.58	1.95(78.03)
Computer Science	4.5	3.75	1.95(78.44)
Chemistry	3.5	3.81	-0.7(46.14)

Table 4-2: Gender differences in interest in each field

Hypothesis 4

Differences in similarity should mediate gender differences in interest. As depicted in Table 4-3, similarity to members of a target major did not mediate the effects of gender on interest in either sample. In the UW sample, mediation effects were computed separately for each category of traits for each major. The significant results are summarized in Table 4-3b. Gender differences in interest were mediated by perceived similarity regarding geekiness and unusualness in physics, computer science, and electrical engineering, but not in other majors.

We turn now to our multilevel model analysis. Since the Mturk sample considered only four majors, I will present its results in Appendix C, and will focus on the more powerful nine

UW Sample			
Major	p	lower bound	upper bound
Physics	0.38	-0.14	0.05
Biology	0.77	-0.05	0.07
Computer Science	0.23	-0.22	0.05
Chemistry	0.76	-0.07	0.05
Environment Eng.	0.75	-0.06	0.09
Electrical Eng.	0.22	-0.26	0.06
Oceanography	0.22	-0.03	0.12
Digital Art	0.52	-0.06	0.12
Medicine	0.95	-0.05	0.05
Mturk Sample			
Major	p	lower	upper
Physics	0.48	-0.58	0.27
Biology	0.62	-0.17	0.28
Computer Science	0.95	-0.30	0.32
Chemistry	0.34	-0.18	0.51

Table 4-3: 95% CI of Sobel mediation test of the effect of gender on interest via similarity.

Major/Stereotype	p	lower bound	upper bound
Physics/Unusual	<.01	-0.4096	-0.0565
Computer Science/Unusual	<.01	-0.5248	-0.0919
Electrical Engineering/Unusual	<.01	-0.5428	-0.0963
Oceanography/Unusual	<.10	-0.0085	0.2034
Physics/Geek	<.10	-0.2418	0.0133
Computer Science/Geek	<.01	-0.4524	-0.0704
Electrical Engineering/Geek	<.05	-0.3483	-0.0198

Table 4-3b Mediations of gender effect on interest by stereotypes

major comparison in the UW sample. In this chapter I will present the results for the items collected in scales as described above. The results for the items not connected to a scale are also given in Appendix C. The purpose of the multilevel analysis is to explore the variation both between and within subjects in a single model. Each participant has his or her own idiosyncratic pattern of stereotypes and attitudes to each major, and we would like to be able to capture both the average response and the degree of variation among responses.

Hypothesis 5

There will be variation between subjects in how important each trait is in determining interest. For each participant, for each trait, our multilevel model will produce a coefficient determining the relationship between that trait and interest for that participant. We may think of this as a sensitivity coefficient, since it determines how sensitive the participant is to being different from someone in regards to that trait. To assess the degree of variation in the predictive power of various scales on interest, we can look at the distribution of sensitivity coefficients among participants. A coefficient of high mean absolute value indicates a strong average sensitivity to that trait, and a high standard deviation in the mean coefficient indicating that participants differ considerably in sensitivity. Table 4-4 shows these means and standard deviations. All the means depicted indicate average relationships significant at the $p<.00001$ level.

To determine whether the individual differences depicted are significant, I compared two models, one in which individuals vary in the degree to which similarity predicts interest, and the other in which all participants are assumed to have the same coefficient. The difference in likelihood between these models gives a test of the individual differences in sensitivity, the significance of which is given in the p(sd) column of Table 4-4. There were significant individual differences in all the scales except Agency, Unusualness, Geekiness. Thus, in most cases there is significant individual variation in sensitivity, but subjects are relatively uniform in considering those three traits to be important.

To help in visualizing this, Figures 4-1 through 4-9 depict the quadratic models for 40 arbitrarily selected subjects (20 male, 20 female). Each curve is recentered with the subjects' self-rating corresponding to 0; thus, in an idealized self-prototype matching universe, the plots would look like a set of parabolas of different heights, all centered at 0. In the case of scales with low individual variation in sensitivity (e.g. agency, unusualness, geekiness) this is precisely what we see. Among the high individual variation in sensitivity majors (e.g. communality, sociability) most subjects still showed the downward parabolic curve, but some showed no relationship, linear relationships, or even an inverted parabola in some cases. So for those scales, the extent to which participants use similarity between themselves and a major on that scale to determine interest has a variable pattern.

Stereotype	Mean	SD	p(SD)
Communal	-0.49	0.21	0.01
Agentic	-0.66	0.12	0.47
Sociable	-0.46	0.24	0.00
Isolated	-0.39	0.16	0.02
Antisocial	-0.28	0.17	0.00
Unusual	-0.55	0.03	0.87
Tech-focus	-0.21	0.17	0.00
Geek	-0.37	0.02	0.91
Neutral	-0.69	0.26	0.00

Table 4-4: Means and individual differences in the effect of similarity on interest, UW sample.

Figure 4-1 Curves for communality, female subjects

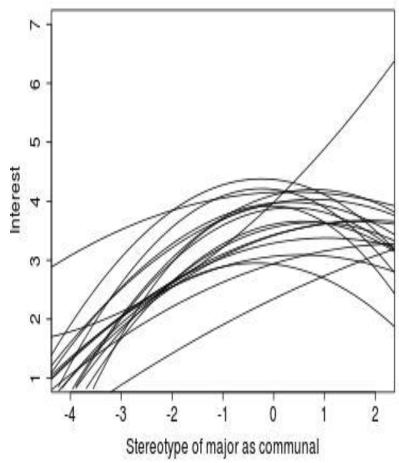


Figure 4-2 Curves for agency, female subjects

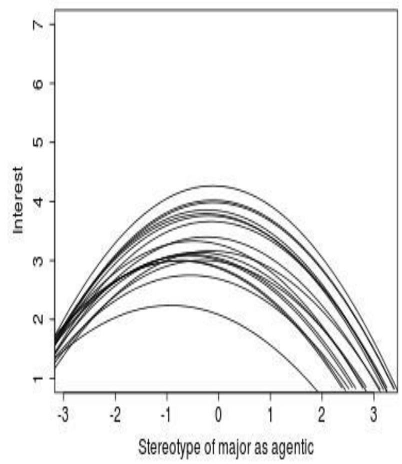


Figure 4-3 Curves for sociability, female subjects

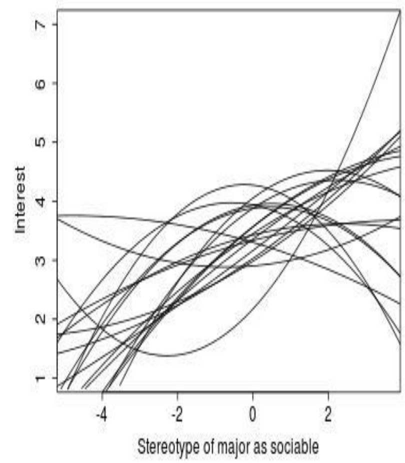


Figure 4-4 Curves for isolation, female subjects

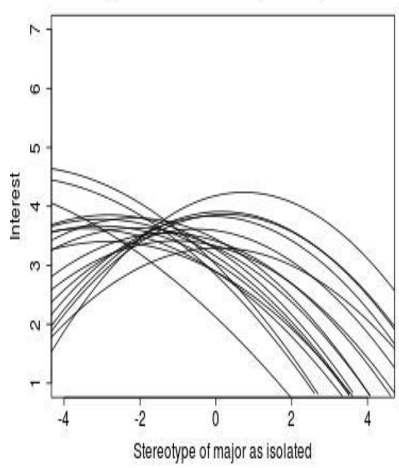


Figure 4-5 Curves for antisociality, female subjects

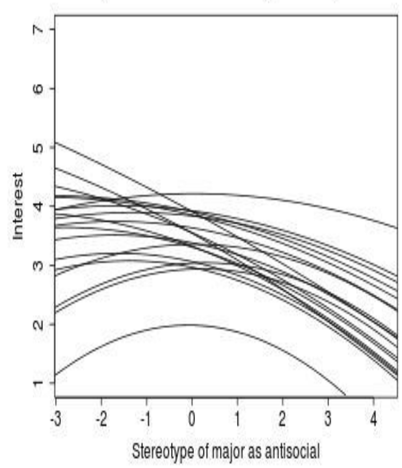


Figure 4-6 Curves for unusualness, female subjects

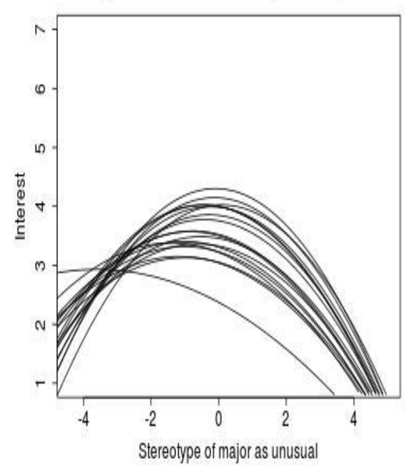


Figure 4-7 Curves for technology-focus, female subjects

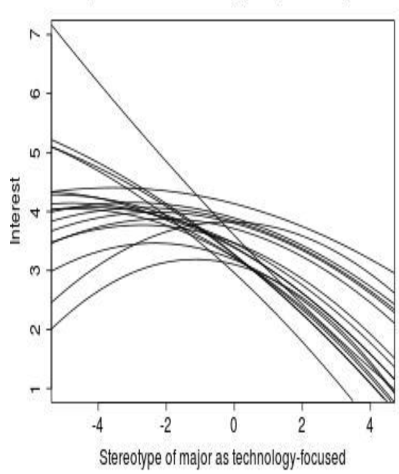
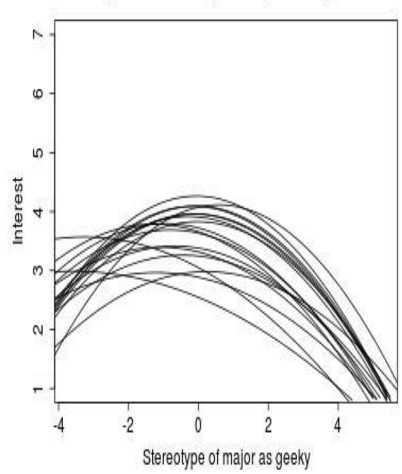


Figure 4-8 Curves for geekiness, female subjects



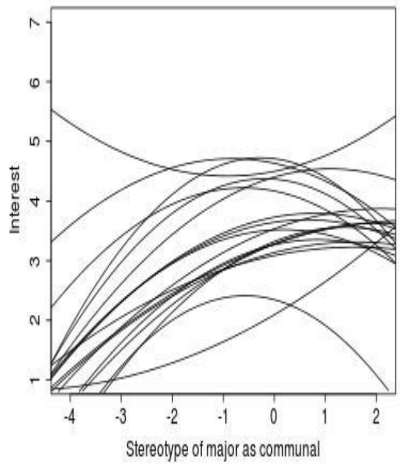


Figure 4-9 Curves for communality, male subjects

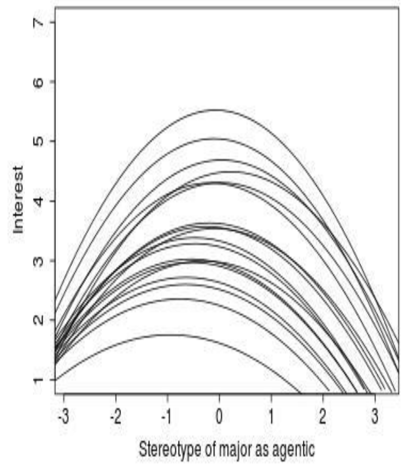


Figure 4-10 Curves for agency, male subjects

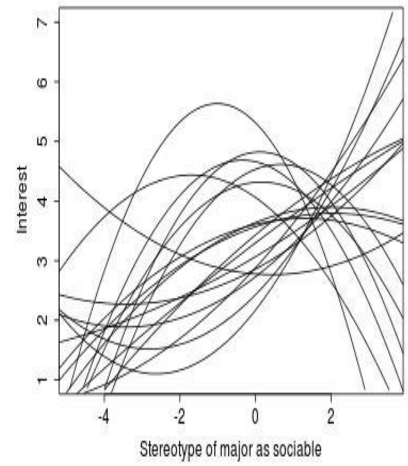


Figure 4-11 Curves for sociability, male subjects

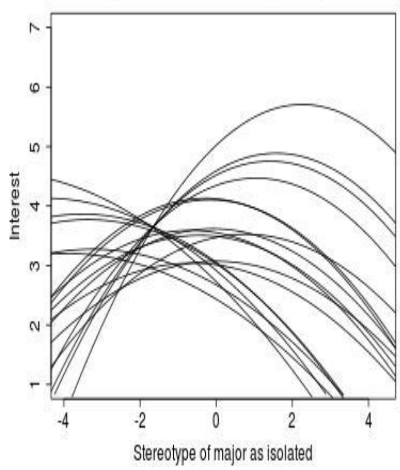


Figure 4-12 Curves for isolation, male subjects

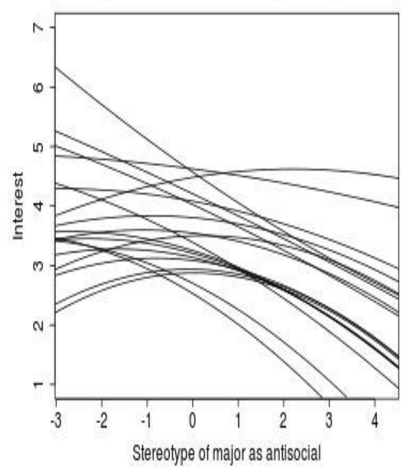


Figure 4-13 Curves for antisociality, male subjects

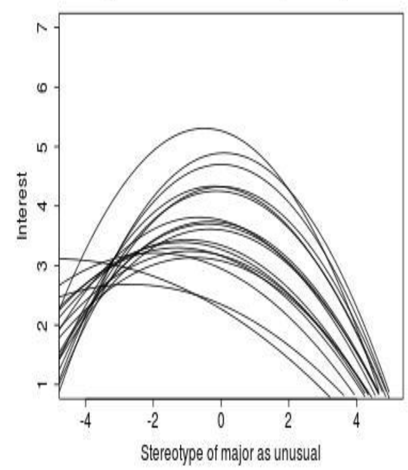


Figure 4-14 Curves for unusualness, male subjects

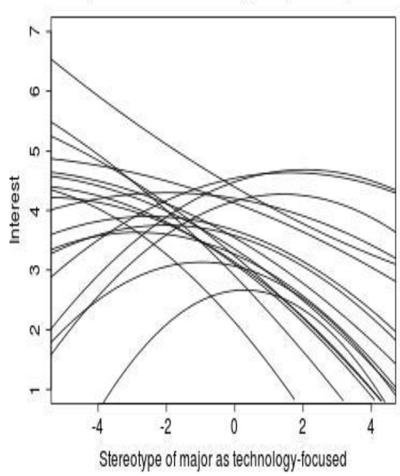


Figure 4-15 Curves for technology-focus, male subjects

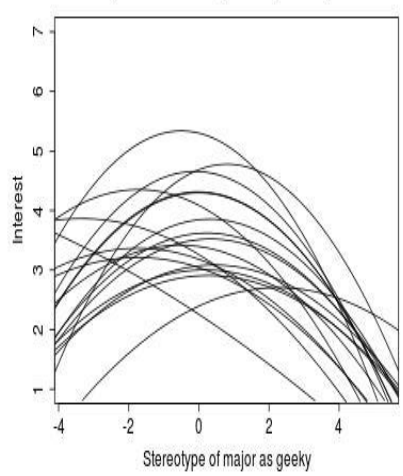


Figure 4-16 Curves for geekiness, male subjects

Having generated sensitivity values for each stereotype, it's natural to wonder if these sensitivities cluster together. However, since these values are regression coefficients rather than measured variables, the validity of doing statistical tests on them is questionable. A correlation table of the sensitivity coefficients is included in Table 4-5 for illustrative purposes, but this should not be considered statistically rigorous. The results do suggest, however, that sensitivities to communality, sociability, and technology tend to relate highly to each other, as do sensitivities to agency, isolation, unusualness, and geekiness.

	Communal	Agentic	Sociable	Isolated	Unusual	Tech-focus	Geek	Neutral
Communal	1.00							
Agentic	-0.18*	1.00						
Sociable	0.6**	-0.19*	1.00					
Isolated	0.38**	-0.71**	0.53**	1.00				
Unusual	-0.01	0.59**	-0.16*	-0.62**	1.00			
Tech-focus	0.5**	-0.36**	0.59**	0.7**	-0.38**	1.00		
Geek	0.41**	-0.68**	0.54**	0.95**	-0.59**	0.73**	1.00	
Neutral	0.53**	-0.59**	0.68**	0.93**	-0.5**	.70**	0.91**	1.00

Table 4-5 Correlations among sensitivity coefficients

Hypothesis 6

The similarity based approach will produce a better model than considering stereotypes alone. In order to assess the validity of the similarity model, in which the square of differences between self and target predict interest, I would like to compare it to a model in which perceptions of the target linearly predict interest (i.e. a model which uses only stereotypes without relating them to self-perception). Again, the two models are computed and we compare the likelihood of each to generate a test of the hypothesis that the similarity model better explains the data than the linear model. The log likelihoods of each model, and the p value for the difference between them, are reported on Table 4-6. For all of the scales except Tech-focus, the similarity model was a significantly better fit. Thus, we support our decision to consider self-perceptions rather than focusing on stereotypes alone.

	LogLik(Linear)	LogLik(Simil)	p(diff)
Communal	-2264.21	-2254.96	<.0001
Agentic	-2154.97	-2135.44	<.0001
Sociable	-2213.17	-2200.38	<.0001
Isolated	-2277.71	-2263.74	<.0001
Antisocial	-2319.36	-2314.2	<.05
Unusual	-2313.84	-2274.98	<.0001
Tech-focus	-2328.53	-2324.83	ns
Geek	-2291.23	-2274.71	<.0001
Neutral	-2086.8	-2063.44	<.0001

Table 4-6: Relative likelihood of linear and similarity models for predicting interest, UW sample.

Discussion

My hypotheses were broadly supported by this data. The first set of hypotheses, concerned with verifying that our approach was sound, were well supported. Self-rated similarity significantly predicted interest for all majors (Hypothesis 1), and profile based similarity did for most majors as well (Hypothesis 2). Gender differences appeared in majors approximately where we would expect them (Hypothesis 3): a preference by men in the majors found in the scientific and artificial quadrant of the map in Chapter 2, and a preference by women for the single cultural major represented (digital art), but no effects in the other fields. We did not find a general mediation of interest by stereotype similarity, which is perhaps not surprising given that the theme of the study is the distinctness of different majors and different stereotypes; however, we did find mediation for some majors and some stereotypes (Hypothesis 4). The multi-level model allows us to form a more finely grained picture than do multiple mediation analyses, however.

Hypothesis 5 was supported: I found that across majors, similarity to the prototypical student in each major in each of the stereotype categories predicted interest in each major, but that how true this was differed between subjects. There were individual differences in the strength of the similarity effect for communality, sociability, isolation, anti-sociability, singular focus on technology, and for the "neutral" items, but there were no significant individual differences for agency, unusualness, or geek stereotypes; thus, some stereotypes appear to be high in consensus and some low. I was also able to confirm that this focus on similarity was beneficial (Hypothesis 6) in terms of accurately modeling interest, since it achieved superior model fit compared to a model only considering stereotypes eight of the nine stereotype categories considered.

It is interesting to note that the "neutral" category- the average similarity on items not attached to any of the cited stereotypes- was nevertheless a strong predictor of interest. This suggests that the prototype matching effect is not confined to stereotypes most often cited as drivers of inequality, but rather applies to a broad range of traits about which participants happened to have stereotypes.

As far as the exploratory goals are concerned, we did in fact find that majors found to have strong gender differences in Chapter 2 had them here as well (Goal 1). Though we were able to identify high and low sensitivity traits, they did not emerge into readily apparent clusters in the correlation table (Goal 2) in an unambiguous way. We did in fact find that stereotypes about geekiness and unusualness were better mediators than the others, but this is not a particularly high powered comparison (Goal 3). And similarity was a predictor of interest for majors across the board, with strong correlations in all cases (Goal 4).

The results of our multi-level model reaffirm and expand on the results of the previous chapter, as we found once again that similarity was a predictor of interest in STEM fields, and that different stereotypes behave differently. The additional contribution of this study is that it enables us to examine within-subjects patterns. This means in turn that we can see response patterns for each subject, not just a singular response, so we can tell how much response patterns vary between subjects. As it turns out, subjects vary considerably in their patterns for some stereotypes, but not for others. This simple statement has broad ranging implications. It suggests that research which treats men and women as monoliths is losing information about their responses in several ways. As we showed in both here and in Chapter 3, such research will fail to find the effects of individual differences in self-perception among men and women, which may mean that the same intervention could have opposite effects on different subjects. Furthermore, as we found in this chapter, if the constructs being studied happen to be low-consensus stereotypes, some participants' interest will depend strongly on the construct as it relates to their self-perception, and some participants will much less affected or even unaffected. Averaging a strongly affected and an unaffected population together to get a moderately affected population is misleading, and loses the opportunity to target interventions at the strongly affected population.

In the next section, I will discuss the broader implications of this research, as well as highlighting some practical ways that the conclusions we reached here can be applied, and some limitations of the current project.

Chapter 5: General Discussion

I began this dissertation in search of the causes of female underrepresentation in STEM fields. This problem has persisted in the face of years of attempts to increase parity, and it seems to affect some fields much more strongly than others. So in order to address the problem, we may need to step back and thoroughly examine the reasons women have for choosing some STEM fields over others in as much detail as possible. In doing this, I have found support for the idea that a clash between women's self-image and their image of STEM fields may be at fault, and that the importance of different aspects of that image varies from person to person.

We began here from the intuition that people seek out situations where they believe that they will fit in, or where people like them are successful. This self-prototype matching approach predicts that in general people will be attracted to fields of study when they think the prototypical member of the field is similar to themselves. And equally, if there is a clash between women's self-image and their image of a field, they are likely to avoid it. Thus, if women are avoiding some fields, it may be because their self-images don't match their prototypes of those fields as well as men's self-images do. This might be because women's self-images tend to differ from men's, or because women conceptualize the fields in different terms, but in either case we would expect interest in a field to track with perceived similarity to people in that field.

This approach makes interesting predictions when we consider that there are individual differences in self-image within the genders. Unless the prototype of a given field's students is very extreme in a certain trait, some students likely associate it with more of a given trait than themselves, while others associate it with less. For example, for students who see themselves as high in communality, high communality correlates positively with interest, while for those who see themselves as low in communality, high communality correlates negatively. A student's peak interest is likely to be in fields whose perceived traits happen to lie wherever their own is, and to fall off nonlinearly on either side.

Therefore, we must pay careful attention to individual differences. What is appealing to one student may be repellent to another, and it may not in the end be helpful to speak of “women” as an aggregate. Even where rates of female participation are low, some women might have different reasons for avoiding STEM fields than others, and it might even be the case that the same factors that are repelling one group of women are attracting another. It's also important to study a wide range of stereotypes together in one study, so that we have a better chance of capturing the ones that are important to each subject and of seeing which tend to vary together.

Accordingly, my studies here have pursued a deeper understanding of the stereotypes people hold about STEM fields, how those stereotypes relate to their self-image, and how these facts predict their interest. The first two studies lay the groundwork for this analysis, by examining the relationships among fields and the stereotypes of a particular field, respectively, and the third study examines stereotypes and interest across many fields. The multi-field approach is crucial here because it allows us to study individual differences; in Study 2 we can only approach our core question crudely by grouping participants according to their self-image, while in Study 3 we can treat each participant as unique and examine how participants differ. We are looking for several specific individual differences: differences in men's and women's self-image, differences in their perception of computer science, and difference in the weight that they give various stereotypes about computer science. Findings in any of these areas would serve to explain gender-discrepant outcomes.

Part 1 – Gender Differences in Interest in Science

First I want to explore the results as they touch specifically on gender differences in STEM interest. The expected differences in interest by gender did appear, and in this section I will discuss the ways in which our self-prototype model was useful in explaining them.

I found that gender differences in interest in fields were predicted on both axes of the multilevel map, science vs. culture and natural vs. artificial (Chapter 2, Goal 3). Since these two dimensions emerged naturally from the multidimensional scaling model, we can conclude that they are relevant to students' decision making. And since there were only weak differences in the weight given to them in the individual difference scaling model, it appears that there is a high degree of consensus as to which majors fall in the highly gender-different quadrant (Chapter 2, Goal 2).

I found that various stereotypes discussed in the literature as explanations for gender differences in STEM fields were actually seen as more characteristic of majors in this quadrant than outside it, confirming that they are at least potential explanations for why women might avoid STEM fields. I also found that in the case of most of these stereotypes, women rated themselves as less similar to STEM majors than men did. For instance, both men and women rated themselves as less technology-focused and agentic than their rating of computer science students in chapter 3, but this discrepancy was significantly larger for women than for men.

Gender differences in perceived similarity to STEM majors did in fact account for differences in interest in STEM, though it did not account for all the variance (Chapter 3, Hypothesis 1; Chapter 4, Hypotheses 3 and 4). There were significant individual differences in which stereotypes drove this similarity effect, meaning that some women were put off from

STEM fields by different stereotypes than others.

Several potential sources of gender difference in outcomes were not supported by this data, in spite of having a high degree of power to detect them. It made no difference whether participants imagined a male or a female prototypical computer science student. Men and women did not differ in their multidimensional scaling maps, nor did they differ substantively in their stereotype profiles for the STEM majors. Gender differences entered into the results primarily via men's and women's different self-perception, and their different weighting of various stereotypes.

Part 2 – Prototype Matching Model

Second, I want to discuss the ways in which this study has provided an interesting exploration of the self-prototype model itself, as compared to models that don't take self-perception into account. The highly repeated design of Study 3 in particular provides a very strong environment to explore the relationship between perceived similarity and interest with a focus on individual differences.

We largely supported the self-prototype matching hypothesis. Participants' own traits did strongly predict their level of interest in majors that matched or didn't match those traits, and the expected nonlinear patterns manifested (Chapter 3, Hypothesis 1; Chapter 4, Hypotheses 1 and 2). The prototype matching model accounted for the data better than alternative models, so my starting intuitions appear to be justified (Chapter 4, Hypothesis 6).

This naturally means that there is considerable difference in individual patterns of response, since individuals differ both in their own traits and their stereotypes about the field. Increasing the perceived agentic traits of computer science, for instance, would be predicted to increase interest for some people and decrease it for others.

Moreover, participants differed in how strongly they weighted difference from their prototype along various axes (Chapter 4, Hypothesis 5). To some participants, differences in communality were quite important, whereas to others they didn't matter very much. Some stereotype categories had a great deal of variation in how important they were, while others were seen consistently as important or unimportant.

In some cases, the self-prototype model was not supported, because participants characteristically preferred either more or less of a trait than they themselves had. This may be because, for instance, even participants who see themselves as socially isolated might prefer a less socially isolated major. Or it might simply be a consequence of the noise created by participants weighing each major on multiple characteristics simultaneously.

This extends the state of knowledge about self-prototype matching in STEM fields (as represented by Hannover & Kessels, 2004) in several ways. First, it extends the model to a broader base of stereotypes (where they used only a set of trait adjectives), a broader set of majors (rather than only students preferring or not preferring science or humanities) and to older students who are in more of a position to choose their own educational direction. Second, their account was concerned only with the average difference between girls' and boys' ratings, and not with individual differences within those gender groups. And third, they focused only on the question of whether or not overall self-prototype match predicted interest in each area of study, rather than considering match separately on different stereotypes in order to determine which were most important and to what extent that importance varied.

Implications

Taken together, what these points tell us is that a particular subset of STEM majors has difficulty attracting women because our stereotypes of those majors clash with women's self-image. We may not be able to say definitively whether this is the real mechanism used to make this choice, because participants' ability to introspect about the reasons for their choices is limited. However, their preferences do appear to be *consistent* with the self-prototype matching model, and thus lend support to that theory.

The highly repeated measures design used in Study 3 makes it possible to examine the effects of various characteristics within subjects. We can see how subjects interest in majors tracks with their perceptions of those majors, and we can see how much weight each subject gives to each stereotype about computer science. This lets us get beyond the question of whether something is important to subjects in general, and allows us to identify the subjects for whom it is important. I hope that this will be beneficial for researchers exploring STEM stereotyping in the future, as this has created a fairly broad taxonomy of stereotypes about STEM fields along with information about how those stereotypes compare to the average student. A large amount of literature has discussed the image of STEM fields, but previously we have not had a large scale examination of many stereotypes together in a way that allows them to be compared.

I have also found evidence that there exist groups of students who are more or less sensitive to various stereotype categories, and I hope these will also be useful in future research. Ideally, if participants can be divided into types based on sensitivity, then we can research those types as separate entities, both in order to understand why they are different and to understand how they might respond differently to changes in policy.

We can also make several more concrete statements about particular STEM stereotypes. Based on the current data, we can see that students appear to be very much in agreement that agency, unusualness, and geekiness are features for which being similar to people in your major matters. Interventions relating to these stereotypes are likely to have consistent effects across subjects, once self-ratings are taken into account. Agency and geekiness also showed strong gender differences in self-rating (while unusualness showed more moderate differences), suggesting that they may be broadly important predictors of gender differences. On the other hand, properties like sociability, technology focus, and communality showed considerable variations in sensitivity levels across subjects, suggesting that there are subpopulations of women for whom they are important, and other subpopulations for whom they are not. Thus, researchers focusing on these would do well to attempt to identify the subpopulations who are sensitive or insensitive to them.

Limitations

The most obvious limitation of these studies is that they focused on students who were not studying gender-biased STEM fields. There is no obvious right choice of sample here, since focusing on students who are studying gender-biased STEM fields would give us a sample self-selected as resistant to the forces we are studying. Since my purpose is to understand the factors that influence female undergraduates early in their careers to avoid STEM fields, using participants who have in fact avoided those fields seems the best compromise. The Mechanical Turk sample in Study 3 can at least confirm that the stereotypes of STEM we are dealing with are not confined to students taking psychology courses. Ideally, this work could be supplemented in the future by work focused on reasons for leaving STEM fields, which would then investigate the responses of women currently in those fields. An implication of this difficulty is that we do not know what sort of contact participants have with the majors in question, so we don't know the basis for their knowledge. Presumably in some cases they have a greater degree of exposure to majors they are not studying (e.g. through friends who are studying them) and in other cases they rely more strongly on stereotypes.

A second limitation, as alluded to earlier, is that we cannot know whether self-prototype matching is actually the mechanism used to make decisions about majors, only whether subjects' preferences are consistent with self-prototype matching. It may be that the causal direction here is reversed, with subjects perceiving a field as more similar to themselves when they like it more. It may also be that the convergence between similarity and preference is coincidental, with some other process entirely determining what majors are actually chosen.

This limitation appears to be inherent to the topic, for two reasons: first, because the decision of what to study is not made at a single discrete moment, so observing that decision would require an extensive longitudinal study; and second, because subjects' introspection about their decision making process is not necessarily reliable. Future research might address the direction of causality by manipulating either similarity or liking and seeing if the other follows suit. However, the question of whether the similarity-liking relationship is the driver of choice or simply correlated with that choice would remain unresolved. A longitudinal field study in which the effect of similarity on actual academic choices would seem to be necessary to resolve it finally.

Applications

Given the focus on individual differences, the question of how to apply our results may seem paradoxical. If increasing perceived agency would make physics seem more interesting to some women and less interesting to other women, it would seem that any intervention would hurt as well as help. The solution lies in making the case that a diverse set of personalities are all able to succeed in physics.

For example, rather than exchanging the idea that physicists have agentic goals for the idea that physicists have communal goals, we would wish to communicate the idea that people with communal goals as well as people with agentic goals can fulfill those goals by studying physics, because there are a variety of ways to study physics. That some physicists work alone and some work in teams, and that some like comic books while others don't. Rather than changing the prototype of who fits into STEM, we must focus on changing the idea that only one kind of person fits.

For the researcher trying to address the problem of stereotypes that effect women's interest in some STEM field, then, both our finding that the effect of stereotypes depends on self-concept and that the particular stereotypes that matter vary from person to person are important considerations. Researchers may wish to focus on high-consensus stereotypes (like agency and geekiness) where subjects generally agree that the stereotype is important. Alternatively, if they wish to focus on low-consensus stereotypes, they might screen subjects for the importance of a given stereotype in determining their interest, in order to target subjects who are responsive to those stereotypes.

Regardless, they will want to be wary of assuming that reducing a stereotype will have a uniform effect, since for students whose self-perception is higher in that stereotype than their perception of the target major, reducing the stereotype will reduce interest, while for the rest it will increase interest. Researchers may wish to target their interventions at students with self-ratings in a particular range. Alternatively, it may be possible to “split” the stereotype of a field, by implying that there are multiple types of people who can succeed there, and that people who resemble any of those types should consider the field a good fit. Further research is needed to determine whether this is practical.

How we intervene to counter the effects of stereotypes also depends in part on the reality of the fields in question. In these studies I have focused on the perception of those fields rather than the actuality, but in designing interventions based on this data it will be crucial to understand what really goes on in the field. If there is in fact a wide range of people who are able to fit in and succeed in a given field, then the task we have is one of education- showing women in particular that, contrary to what they might have heard, there are people like them in STEM. This suggests more visibility for female role models, perhaps, but more importantly, more visibility for role models of either gender with a wide range of personality characteristics. On the other hand, if it really is the case that only some kinds of people are represented in STEM fields, the task becomes one of reform. We would need to examine and intervene against the forces that push people with particular personality traits out of the field. This might mean creating alternative networks so that students with less-common traits are aware of one another and able to offer support. It may also mean changing structures that require students to work in particular ways that advantage some over others, or eliminating bias against less typical students in mentors or peers. Here, the main advantage of my research is to tell us which classes of students are in danger of being marginalized on the basis of which traits, which still leaves us the task of correcting that marginalization.

In this study, I have examined the effect of similarity directly. That is, we determine whether subjects are sensitive to similarity in a given stereotype by measuring their interest, self-stereotype, and target-stereotype. This approach is not extremely cumbersome, but future research may wish to speed up the process of determining what sort of manipulations are likely to work for a given subject. If subjects are able to consciously report their sensitivity to similarity, that would of course be faster; but we do not yet know whether they are conscious of that sensitivity, or whether self-presentation effects would interfere with their reporting it. Future research might also profitably look for individual difference variables or demographic measures which correlate with sensitivity to their constructs of interest.

To sum up the recommendations this paper makes to future researchers studying the effect of stereotypes on gender inequality, there are several concrete steps which should be taken. To study some stereotype, one must first determine how important it is to subjects in one's target population, and what degree of variability to expect in that level of importance. Our research finds that some stereotypes will have high variability in interest and some low, and researchers looking at highly variable (low consensus) stereotypes should probably use research designs that account for individual variation in response pattern, such as multi-level modeling.

To study gender differences, one must also look at how women and men see the stereotype as applying to them, for two reasons. First, because of our finding that gender differences in self-perception mediate gender differences in interest, only the stereotypes showing gender self-perception differences are likely to be fruitful avenues for intervention. It's conceivable for gender differences in interest to appear without gender differences in self-perception (for instance because of gender differences in stereotypes about the target), but at least in my studies this did not happen. Second, because whatever intervention the researcher attempts is likely to have opposite results for participants who perceive themselves as very high and very low in the stereotype; thus, even when there is a large difference in men's and women's self-perceptions, researchers should expect exceptional behavior from women with exceptional self-ratings. If self-ratings are measured, this potential problem is easily controlled for.

For researchers studying STEM fields, we have provided a taxonomy of stereotypes about computer science, and determined their level of consensus and their degrees of gender differences. For researchers interested in other areas, the methods used here should be able to be adapted without great difficulty.

It may seem paradoxical to design a broadly beneficial intervention when the effects of an intervention vary from person to person, but it is not my intent to say this problem is too complex to be tackled. The methods outlined in this study allow researchers to evaluate where individual differences are likely to appear, which is allows one to minimize their impact- either by controlling for them, by choosing to focus on relatively uniform stereotypes, or by targeting interventions at those for whom they will work best. It may also be possible to show students that, even if they would not fit in with the average STEM student, that there are still some students like them who do well. Though further study is needed on the subject, my research suggests that the most broadly successful interventions should convey the message that a variety of kinds of people are welcome in STEM fields, rather than trying to change the current image to some competing image which may repel as many as it attracts.

Conclusions

There are two things I believe we can take away from this study most strongly. First, that people tend to be interested in majors in which they imagine they would find people similar to themselves, and women tend to see a divide between their own identity and the kind of identity to be found in male-dominated STEM fields. Thus, lowering the barrier to women's participation in STEM fields requires us to identify the ways in which women feel different from STEM students, and to correct that difference by encouraging women to believe that people like them can be found succeeding in STEM.

And second, there is no single stereotype or set of stereotypes which could be corrected in order to address the gender inequality problems in STEM, because women differ from one another both in what their ideal environment would look like and in which aspects of that environment matter most to them. Interventions aimed at attracting one group of women may alienate a second and leave a third unmoved. A complete approach to equality will need to take into account the diversity of women's selves, by taking into account the self-perceptions and priorities of its subjects. If we want to make multiple different groups of women feel comfortable in STEM fields, perhaps we should not try to change which type of person feels comfortable. Instead, we should change the notion that only a single type of person can comfortably study STEM.

References

Adams, G., Garcia, D. M., Purdie-Vaughns, V., & Steele, C. M. (2006). The detrimental effects of a suggestion of sexism in an instruction situation. *Journal Of Experimental Social Psychology*, *42*(5), 602-615. doi:10.1016/j.jesp.2005.10.004

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal Of Consulting And Clinical Psychology*, *42*(2), 155-162. doi:10.1037/h0036215

Benbow, C. P., & Stanley, J. C. (1982). Consequences in high school and college of sex differences in mathematical reasoning ability: A longitudinal perspective. *American Educational Research Journal, 19,* 598-622.

Biglan, A. (1973). The characteristics of subject matter in different academic areas. *Journal Of Applied Psychology*, *57*(3), 195-203. doi:10.1037/h0034701

Burke, P. J., & Reitzes, D. C. (1991). An identity theory approach to commitment. *Social Psychology Quarterly*, *54*(3), 239-251. doi:10.2307/2786653

Cantor, N., Mischel, W., & Schwartz, J. C. (1982). A prototype analysis of psychological situations. *Cognitive Psychology*, *14*(1), 45-77. doi:10.1016/0010-0285(82)90004-4

Ceci, S. J., & Williams, W. M. (2010). Sex differences in math-intensive fields. *Current Directions In Psychological Science*, *19*(5), 275-279. doi:10.1177/0963721410383241

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin*, *135*, 218–261. doi:10.1037/a0014412.

Cejka, M., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality And Social Psychology Bulletin*, *25*(4), 413-423.

Chassin, L., Corty, E., Presson, C. C., Olshavsky, R. W., Bensenberg, M., & Sherman, S. J. (1981). Predicting adolescents' intentions to smoke cigarettes. *Journal Of Health And S social Behavior*, *22*(4), 445-455. doi:10.2307/2136684

Cheryan, S., Plaut, V. C., Davies, P. G., & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology*, *97*, 1045–1060. doi:10.1037/a0016239.

Cheryan, S., Plaut, V. C., Handron, C., & Hudson, L. (2013). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles*, *69*(1-2), 58-71. doi:10.1007/s11199-013-0296-x

Cheryan, S., Siy, J. O., Vichayapai, M., Drury, B., & Kim, S. (2011b). Do female and male role models who embody STEM stereotypes hinder women's anticipated success in STEM? *Social Psychological and Personality Science*, *2*, 656–664. doi:10.1177/1948550611405218.

Correll, S. J. (2001). Gender and the career choice process: The role of biased self-assessments. *American Journal Of Sociology*, *106*(6), 1691-1730. doi:10.1086/321299

Correll, S. J. (2004). Constraints into Preferences: Gender, Status, and Emerging Career Aspirations. *American Sociological Review*, *69*(1), 93-113. doi:10.1177/000312240406900106

Crowley, K., Callanan, M. A., Jipson, J. L., Galco, J., Topping, K., & Shrager, J. (2001). Shared scientific thinking in everyday parent–child activity. *Science Education*, *85*(6), 712-732. doi:10.1002/sce.1035

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. Journal Of Personality And Social Psychology,*46*(5), 991-1004. doi:10.1037/0022-3514.46.5.991

Diekman, A. B., Brown, E., Johnston, A., & Clark, E. (2010). Seeking congruity between goals and roles: A new look at why women opt out of STEM careers. *Psychological Science*, *21*, 1051–1057.doi:10.1177/0956797610377342.

Eagly, A. H. (2007). Female leadership advantage and disadvantage: Resolving the contradictions. *Psychology Of Women Quarterly*, *31*(1), 1-12. doi:10.1111/j.1471-6402.2007.00326.x

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735–754. doi:10.1037/0022-3514.46.4.735.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology Of Women Quarterly*, *11*(2), 135-171. doi:10.1111/j.1471-6402.1987.tb00781.x

Eccles, J. S. (1994). Understanding women's educational and occupational choices: Applying the Eccles et al. model of achievement-related choices. *Psychology Of Women Quarterly*, *18*(4), 585-609. doi:10.1111/j.1471-6402.1994.tb01049.x

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal Of Personality And Social Psychology*, *102*(4), 700-717. doi:10.1037/a0026659

Good, C., Aronson, J., & Harder, J. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal Of Applied Developmental Psychology*, *29*(1), 17-28. doi:10.1016/j.appdev.2007.10.004

Hannover, B., & Kessels, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high school students do not like math and science. *Learning and Instruction*, *14*, 51–67. doi:10.1080/01973530802375003.

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science*, *269*, 41−45.

Hill, C., Corbett, C., & St. Rose, A. (2010). *Why so few? Women in science, technology, engineering, and mathematics*. Washington, DC: AAUW

Jones, M., Howe, A., & Rua, M. J. (2000). Gender differences in students' experiences, interests, and attitudes toward science and scientists. *Science Education*, *84*(2), 180-192. doi:10.1002/SICI) 1098-237X(200003)84:2180::AID-SCE3>3.0.CO;2-X

Jones, M.G., & Wheatley, J. (1989). Gender influences in classroom displays and student-teacher behavior. *Science Education*, *73*(5), 525-545.

Jones, M.G., & Wheatley, J. (1990). Gender differences in teacher-student interactions in science classrooms. *Journal of Research in Science Teaching*, *27*(9), 861-874.

Kahle, J. B., & Lakes, M. K. (1983). The myth of equality in science classrooms. *Journal Of Research In Science Teaching*, *20*(2), 131-140. doi:10.1002/tea.3660200205

Kihlstrom, J.F., & Cantor, N. (1984). Mental representations of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology*. Vol. 17. New York: Academic Press.

Koch, S. C., Müller, S. M., & Sieverding, M. (2008). Women and computers. Effects of stereotype threat on attribution of failure. *Computers & Education*, *51*(4), 1795-1803. doi:10.1016/j.compedu.2008.05.007

Lippa, R. (1998). Gender-related individual differences and the structure of vocational interests: The importance of the people–things dimension. *Journal of Personality and Social Psychology*, *74*, 996–1009. doi:10.1037//0022-3514.74.4.996.

Lundeberg, M. A., Fox, P. W., & Punćcohaŕ, J. (1994). Highly confident but wrong: Gender differences and similarities in confidence judgments. *Journal Of Educational Psychology, 86*(1), 114-121. doi:10.1037/0022-0663.86.1.114

Margolis, J., & Fisher, A. (2002). *Unlocking the clubhouse: Women in computing.* Cambridge, MA, US: The MIT Press.

Marsh, H. W., & Yeung, A. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal, 35*(4), 705-738. doi:10.2307/1163464

Mitchell, Lesa, Overcoming the Gender Gap: Women Entrepreneurs as Economic Drivers (September 1, 2011). Available at SSRN: http://ssrn.com/abstract=1934906

Morgan, C., Isaac, J. D., & Sansone, C. (2001). The role of interest in understanding the career choices of female and male college students. *Sex Roles, 44*(5-6), 295-320. doi:10.1023/A:1010929600004

National Center for Education Statistics, U.S. Department of Education. Institute of Education Sciences. (1997) *Women in Mathematics and Science.* (NCES 97-982)

National Center for Education Statistics, U.S. Department of Education. Institute of Education Sciences. (2012) *Integrated Postsecondary Education Data System. Completions Component,* retrieved from http://nces.ed.gov/programs/digest/d13/tables/dt13_318.30.asp

Niedenthal, P.M., Cantor, N., & Kihlstrom, J.F. (1985). Prototype matching: A strategy for social decision making. *Journal of Personality and Social Psychology, 48*, 575-584.

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math = male, me = female, therefore math ≠ me. *Journal Of Personality And Social Psychology, 83*(1), 44-59. doi:10.1037/0022-3514.83.1.44

Park, L. E., Young, A. F., Troisi, J. D., & Pinkus, R. T. (2011). Effects of everyday romantic goal pursuit on women's attitudes toward math and science. *Personality And Social Psychology Bulletin, 37*(9), 1259-1273. doi:10.1177/0146167211408436

Rudman, L. A., & Fairchild, K. (2004). Reactions to Counterstereotypic Behavior: The Role of Backlash in Cultural Stereotype Maintenance. *Journal Of Personality And Social Psychology, 87*(2), 157-176. doi:10.1037/0022-3514.87.2.157

Sekaquaptewa, D., & Thompson, M. (2002). The differential effects of solo status on members of high- and low-status groups. *Personality And Social Psychology Bulletin, 28*(5), 694-707. doi:10.1177/0146167202288013

Sekaquaptewa, D., & Thompson, M. (2003). Solo status, stereotype threat, and performance expectancies: Their effects on women's performance. *Journal Of Experimental Social Psychology, 39*(1), 68-74. doi:10.1016/S0022-1031(02)00508-5

Setterlund, M.B. & Niedenthal, P.M. (1993). "Who am I? Why am I here?": Self-esteem, self-clarity, and prototype-matching. *Journal of Personality and Social Psychology, 4*, 769-780.

Snyder, T.D., Dillow, S.A., and Hoffman, C.M. (2009). *Digest of Education Statistics 2008* (NCES 2009-020). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal Of Experimental Social Psychology, 35*(1), 4-28. doi:10.1006/jesp.1998.1373

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52*(6), 613-629. doi:10.1037/0003-066X.52.6.613

Steinpreis, R.E., Anders, K.A., & Ritzke, D. (1999). The impact of gender on the review of the curricula vitae of job applicants and tenure candidates: A national empirical study. *Sex Roles, 41*(7/8), 509-528.

Strenta, A. G, Elliott, R., Adair, R., Matier, M., and Scott, J. (1994). Choosing and leaving science in highly selective institutions. *Research in Higher Education 35*, 513-547.

Trix, F., & Psenka, C. (2003). Exploring the color of glass: Letters of recommendation for female and male medical faculty. *Discourse & Society, 14*(2), 191-220. doi:10.1177/0957926503014002277

Van de Gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2008). Mathematics participation and mathematics achievement across secondary school: The role of gender. *Sex Roles, 59*(7-8), 568-585. doi:10.1007/s11199-008-9455-x

Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes of minority students. *Science, 331*(6023), 1447-1451. doi:10.1126/science.1198364

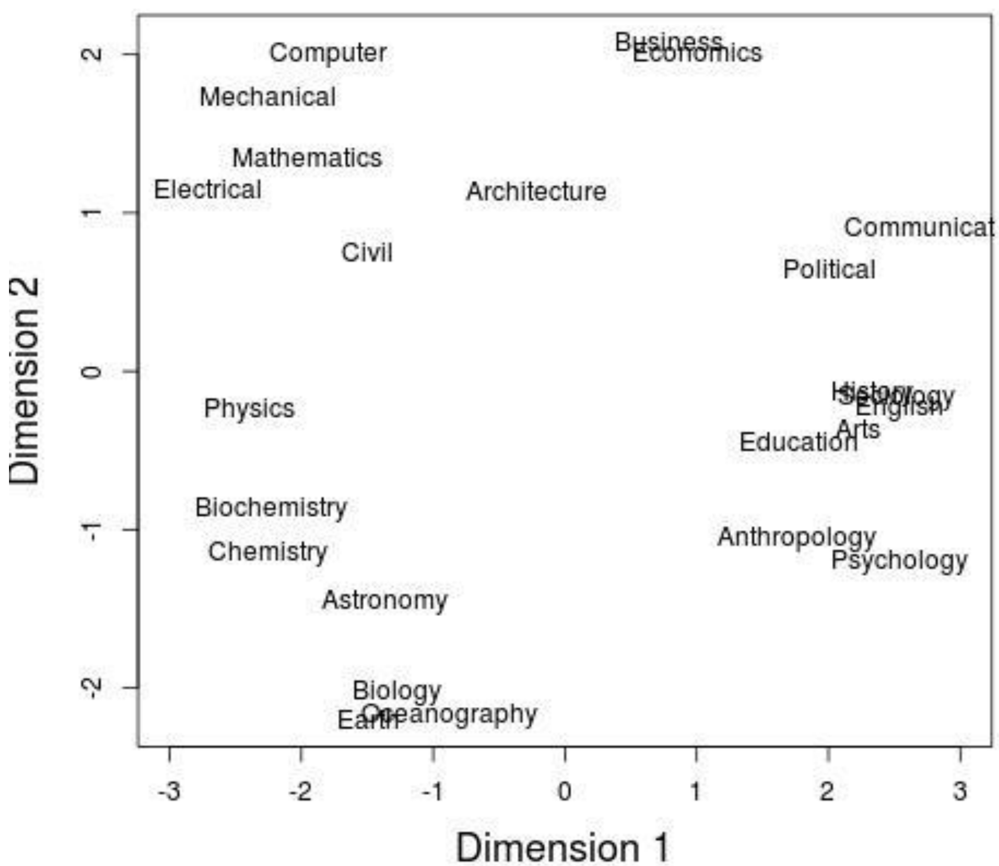
Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal Of Personality And Social Psychology, 102*(3), 513-532. doi:10.1037/a0025731

Wenneras C. and Wold, A. (1997) Nepotism and sexism in peer-review. *Nature, 387*, 341–343.

Appendix A
Multidimensional Scaling Map adjusted to include Physics

Note that medicine was removed from the list of majors to be rated in order to keep the number of comparisons constant.

Figure A-1, Multidimensional Scaling Map



Appendix B

Collection of open-ended responses describing multidimensional scaling dimensions, from Chapter 2

Participants were asked the following: “Below are two lists of majors offered at UW. We'd like to know how you think the two lists are DIFFERENT. That is, what do the majors on List A have in common that the majors on List B don't have, or vice versa.”

Dimension 1 (tentatively labeled science vs. culture)

List A Chemistry Biochemistry Electrical Engineering Mechanical Engineering Mathematics Biology Medicine

List B Communications English Sociology Arts Political Science History Education

[1] "List A is for the people those who are good at science-related subjects and I think more men are in those (list A) majors"

[2] "List A are STEM based, List B is humanities based"

[3] "all of the majors in list A all take paths toward medical/science job fields, and list b are paths toward political/government/business type of jobs"

[4] "List A is science and Math and List B is Visual Arts"

[5] "List a is more science based"

[6] "List A is more science based"

[7] "List A has more problem solving majors"

[8] "List A is more humanities-based, while List B is more science-based."

[9] "The top set are harder; more math heavy, have more intelligent people in them."

[10] "Math and Science based courses versus Humanities and Social Sciences courses"

[11] "Aside from Political Science and Education, it will be more difficult to find a stable job after attaining a major in List B than List A majors."

[12] "list a is science and math related, list b is more social science, I&s vs. NW"

[13] "List A: Hard List B: Easy"

[14] "List A contains the hard sciences and math, while list B contains liberal arts majors"

[15] "List A is primarily STEM majors"

[16] "List A is a much more rigorous and challenging set of majors."

[17] "List A is science based."

[18] "List A consists of majors that involve science and scientific methods at their core. List B consists of majors that deal with the arts and explore the way cultures interact with each other."

[19] "List A contains the \"harder\" majors offered at UW, while List B contains the easier, fluff majors. List A also contains more weed out classes."

[20] "List A contains science related fields, while B is more liberal arts"

[21] "List A is all about natural world, on the other hand List B is general science"

[22] "List A are stem majors you'd get a BS for, list b are ones you get a BA for"

[23] "Science vs. social science"

[24] "list a has more science majors that require intense studying"

[25] "List A are a list of sciences, list B is a list of social sciences"

[26] "List A is science, List B is not"

[27] ""

[28] "List A is more STEM majors focusing more on math, science, and engineering."

[29] "List A deals with math and sciences while list B is more vocational"

[30] "List B consists of Humanity majors"

[31] "List A is more related to natural world but list B is more about culture"

[32] ""

[33] "List A is more about science while list B is more about human art."

[34] "All majors in list A seem to apply to a more scientific/mathematic field while those in list B seem to be more social and people centered majors."

[35] "List A is science majors, List B is Art majors."

[36] "List A relates to the math and sciences while list B are more vocational"

[37] "math and science concentrated"

[38] "List A is more problem solving sciences"

[39] ""

[40] "List A is more science and math based and list B is more liberalarts based"

[41] "List A are technical degrees while List B is not."

[42] "List A: Science, List B: Liberal Arts"

[43] "List A is more technical than list B."

[44] "Nursing"

[45] "List A has options of majors that are more rigorous"

[46] "Math Related"

[47] "list A is associated with math or science, but list B does not."

[48] "List A is physical/natural science focused while list B centers around human/social sciences."

[49] "List A is much more science based, whereas list B are more liberal arts majors."

[50] "A: science vs B: humanities"

[51] "List A is mostly math and science based. List B on the other hand is more reading and writing based."

[52] "list a is science list b is arts"

[53] "List A is science based while List B is Humanities based"

[54] "List A more science based"

[55] "List A is more math and science based, whereas list B is more arts"

[56] "List A is natural/physical sciences. List B is social sciences."

[57] "A is more science based and B is liberal arts"

[58] "List A contains science while list B contains social skills"

[59] "List A is science and list B is not"

[60] "List A's majors all include hands on lab while List B majors include lots of writing and reading."

Dimension 2 (tentatively labeled natural vs artificial)

List A Business Mathematics Economics Computer Science Mechanical Engineering Architecture Electrical Engineering

List B Oceanography Biology Earth Science Anthropology Chemistry Biochemistry Sociology

[1] "List A is for building/creating system and List B is study of environments"

[2] "List B is more natural world based"

[3] "list a is based off of a lot of math and list b is more science oriented"

[4] "List B is all Science and List A is all Math related"

[5] "list B is more natural world based science, list A more mathmatical"

[6] "List B is more related to earth-y stuff"

[7] "List A is more Math and problem solving"

[8] "List A is more human-based, and List B is more animal/environment/natural world-based."

[9] "List A studies more human phenomena than list B, with the exception of medical applications in list B."

[10] "Built environment versus natural world applied courses"

[11] "List A is more related to the general structure of a society than List B. List B consists of majors that will help many societies (AKA big picture stuff)."

[12] ""

[13] "Natural world vs. Human creation"

[14] ""

[15] "List A seems to have more practical, and higher paying jobs ultimately"

[16] "List A is more math based whereas List B is more natural science based."

[17] "List A is mathematical based and B is environmental"

[18] "List A involve knowing how to use math and applying it to real world situations. List B deals with natural world."

[19] "Both lists are comparable. They both contain majors that are relatively difficult to succeed and do well in for the average student."

[20] "List A is math based while B is science based."

[21] "Environment vs. technology"

[22] "List A majors are more hard core, both are STEM"

[23] "Math science vs. natural and human sciences"

[24] ""

[25] "First list involves mathematical reasoning, second list involves earth sciences"

[26] "List A is more like technology, List B is more natural world"

[27] ""

[28] "List A is the majors that are related to industries and making a lot of money."

[29] "List A is more math based and list B is more science based"

[30] "List B relates to Science and the outside world"

[31] "List A is more about math but List B is more about nature and life."

[32] ""

[33] "List A is more about human society, but List B is more about nature."

[34] "List A is centered more on technology and economics while B is more science/ environmental majors."

[35] "List A is more about human society, but List B is more about nature."

[36] "List A is math based while list B is science based"

[37] "List A is math oriented"

[38] "List A is more math"

[39] ""

[40] "List A is more math based and list B is more science based"

[41] "List A generally requires more math than B."

[42] "A: equations, B: Memorization"

[43] ""

[44] ""

[45] "Created versus natural world"

[46] ""

[47] "List A is more related to machines and money but list B is basically related to natural world."

[48] "List A focuses on social/human sciences while list B focuses on natural sciences"

[49] "List B is much more environmentally oreinted, whereas List A is very industrial and technical and mathematical."

[50] ""

[51] "List A is very engineering and applied math based while List B is less math based."

[52] "the first is math and the second is science"

[53] "List A is math based and List B is science based"

[54] "List B is more related to earth-y stuff"

[55] "List A is more math focused, whereas List B is science"

[56] "A: technology B: lab"

[57] "B is more nature based"

[58] "List A is the business world and math while List B is science"

[59] ""

[60] "List B is more about natural world."

Appendix C
Details of Simulated Data, Chapter 3

Three hundred simulated subjects were created, with ratings of interest in computer science and both self and target ratings (i.e. ratings of a computer science student) in an unspecified stereotype. For each participant, each variable was determined as follows:

Target Rating: A random number was drawn uniformly from the range of 1-7 inclusive.
Self Rating: A random number was drawn uniformly from the range of 1-7 inclusive, and Gaussian noise was added (mean 0, standard deviation .25) to produce a more realistic plot.
Interest: was defined as the 3 minus the root square difference of self rating and target rating, plus a Gaussian error term (mean 0, standard deviation .5).

In the case of the fourth plot, where an off-center distribution was required, this setup was altered so that target ratings ranged uniformly from 3-7 instead of 1-7, thus ensuring that target ratings would on average be 1 point higher than self ratings.

Appendix D
Additional plots showing the relationship between stereotypes of computer science and interest in computer science, for participants high or low in self-rating in those stereotypes.

Figure D-1 Effect of CS stereotypes on interest for participants with high or low self-rating in isolation
diamonds=low, circles=high

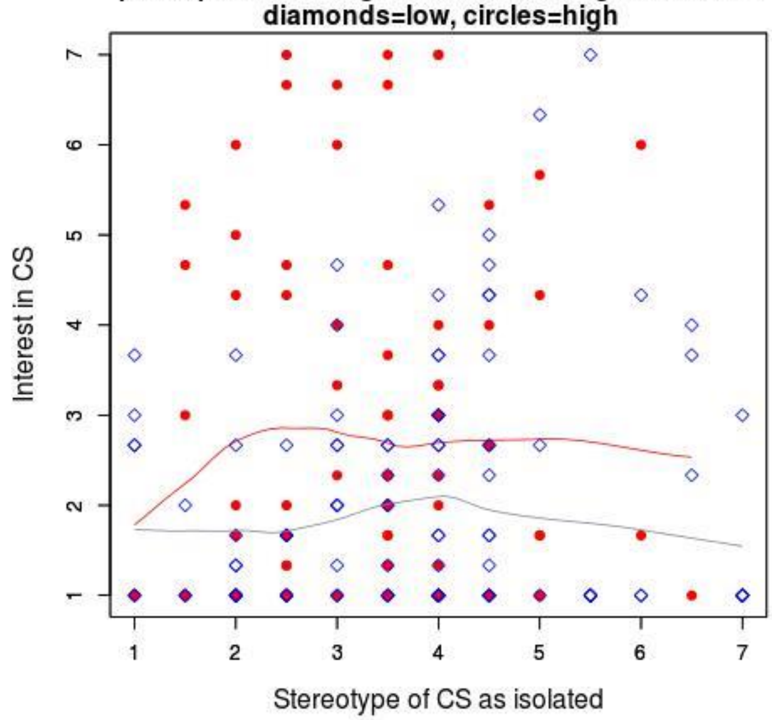


Figure D-2 Effect of CS stereotypes on interest for participants with high or low self-rating in unusualness
diamonds=low, circles=high

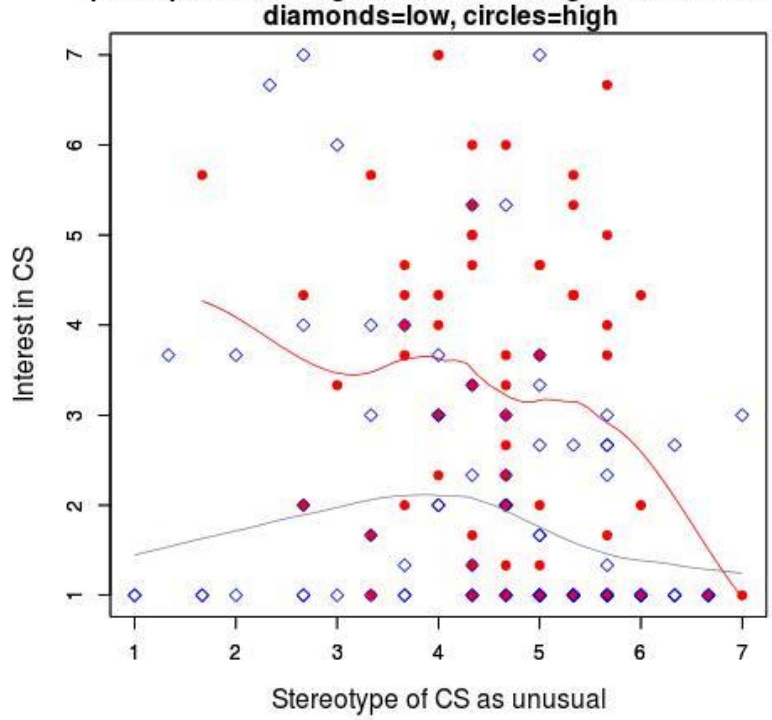
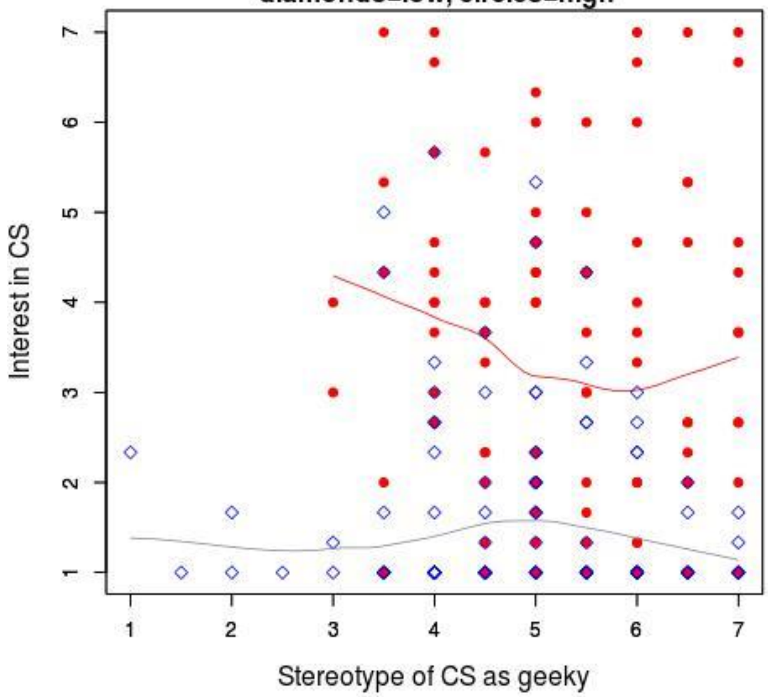


Figure D-3 Effect of CS stereotypes on interest for participants with high or low self-rating in geekiness diamonds=low, circles=high



Appendix E

Miscellaneous Tables for Chapter 4

Trait	Mean	SD	p(SD)
Seeks excitement	-0.19	0.07	0.16
Likes animals	-0.09	0.1	0.01
Good dancer	-0.11	0.1	0.04
Athletic	-0.11	0.09	0.02
Witty	-0.12	0.11	0.05
Creative	-0.19	0.06	0.25
Likes variety	0.04	0.03	0.58
Emotional	-0.13	0.08	0.30
Passionate	-0.19	0.17	0.00
Polite	-0.15	0.06	0.27
Abstract	-0.08	0.07	0.12
Outdoor activities	-0.03	0.06	0.09
Calm	-0.05	0.07	0.28
Easy to talk to	-0.18	0.14	0.00
Wears black	-0.07	0.08	0.07
Expressive	-0.12	0.07	0.21
Politically Opinionated	-0.1	0.06	0.08
Messy	-0.08	0.03	0.61
Good at explaining	-0.06	0.08	0.20
Gives advice	-0.13	0.19	0.00
Shallow	-0.08	0.13	0.01

Table C-1 Means and individual differences in the effect of similarity on interest, UW sample (individual items

Trait	Mean	SD	p(SD)
Communal	-0.82	0.18	0.410
Agentic	-0.55	0.11	0.814
Sociable	-0.57	0.06	0.821
Isolated	-0.37	0.04	0.932
Antisocial	-0.18	0.24	0.072
Unusual	-0.58	0.06	0.715
Tech-focus	-0.42	0.12	0.606
Geek	-0.43	0.19	0.109
Neutral	-1.07	0.23	0.366

Table C-2: Means and individual differences in the effect of similarity

Trait	Mean	SD	p(SD)
Seeks excitement	-0.18	0.19	0.17
Likes animals	-0.27	0.06	0.60
Good dancer	-0.09	0.01	0.98
Athletic	-0.1	0.05	0.75
Witty	-0.02	0.01	0.99
Creative	-0.14	0.04	0.87
Likes variety	-0.23	0.01	1.00
Emotional	-0.19	0.06	0.72
Passionate	-0.29	0.02	0.96
Polite	-0.15	0	1.00
Abstract	-0.07	0.04	0.86
Outdoor activities	-0.15	0.07	0.50
Calm	0.01	0.03	0.92
Easy to talk to	-0.27	0.06	0.65
Wears black	-0.05	0.03	0.86
Expressive	-0.01	0.08	0.45
Politically Opinionated	-0.11	0.18	0.07
Messy	0	0.04	0.83
Good at explaining	-0.25	0.11	0.46
Gives advice	-0.25	0.02	0.95
Shallow	-0.24	0.06	0.70

Table C-3 Means and individual differences in the effect of similarity on interest, Mturk sample (individual items not collected in Table C-2)

Trait	LogLik(Linear)	LogLik(Quad)	p diff
Communal	-590.5	-580.74	<.0001
Agentic	-541.75	-539.28	0.29
Sociable	-582.04	-574.44	<.01
Isolated	-594.05	-589.83	0.08
Antisocial	-578.6	-577.93	0.86
Unusual	-592.04	-581.49	<.0001
Tech-focus	-580.85	-579.87	0.74
Geek	-586.81	-581.03	0.02
Neutral	-569.92	-556.45	<.0001

Table C-4: Relative likelihood of linear and similarity models for predicting interest, Mturk sample.