EXPLICT AND IMPLICIT STEREOTYPE CONTENT: THE CASE OF GENDER IN ACADEMIA

A DISSERTATION PROPOSAL IN PARTIAL FULFILLMENT OF THE

UNIVERSITY OF HAWAIʻI REQUIREMENTS FOR THE DEGREE OF

DOCTORATE OF PHILOSOPHY

IN

PSYCHOLOGY

August 2011

By
Danielle M. Young

Doctoral committee

Elaine Hatfield, Chairperson
Justin Levinson
Charles W. Mueller
Scott Sinnett
Lorieana Yancura

Dedication

To my families- both biological and found.

I love you all, and would never be here without you.

Acknowledgements

So many people were vital to the creation of this dissertation that I am afraid of missing someone or failing to give someone every ounce of appreciation they deserve. Any acknowledgement would completely miss the mark of how much I owe to my committee, my family, and my friends. Thank you to all of them for being with me through this entire process.

Abstract

Women's underrepresentation in the sciences has been ascribed to a variety of causes, including their perceived lack of competence and their abundance of warmth. The stereotype content model (SCM) theorizes that warmth and competence are universal aspects of social cognition and that these two factors underlie stereotypes held about outgroups. Three studies were conducted using subgroups of gender in the sciences as a framework to extend the use of SCM, using both explicit and implicit methods. Study 1 established that female professors in the sciences are perceived as less competent and more warm than their male counterparts. Study 2 replicated the findings from Study 1, and attempted to connect explicit SCM to implicit measures. Study 3 attempted to extend findings from the previous studies using an experiment to investigate whether explicit and implicit SCM can predict behavior and behavioral intentions.

Table of Contents

EXPLICT AND IMPLICIT STEREOTYPE CONTENT: THE CASE OF GENDER IN ACADEMIA

Introduction

Though women are generally underrepresented in academic fields (West & Curtis, 2006), this gender gap is particularly large in the science, technology, engineering, and mathematics (STEM) fields (NSF, 2009). Since 1966 women have received more general bachelor's degrees than men, but women and men did not earn a similar number of bachelor's degrees in the sciences until 2006. Since this time the number of women receiving doctoral degrees in the sciences has increased, however women are underrepresented as full-time faculty in the sciences, and are less likely to have obtained tenured full-professor status than men (NSF, 2009).

Why women are underrepresented in the academic STEM community has been under debate for some time (Sommers, 2009). Are social forces at work, or do evolutionary and biological reasons lie behind this underrepresentation? A variety of reasons have been posed to explain this gender gap. Gender differences in career goals (Diekman, Brown, Johnston, & Clark, 2010), stereotype threat (Good, Aronson, & Harder, 2008), and social forces (such as the effect of child rearing on academic careers) have all been forwarded to explain some part of why women are underrepresented in science. Recently, however, two threads of the debate have become prominent sources for discussion: competence and warmth (Ceci, Williams, & Barnett, 2009; Sommers, 2009).

A particularly contentious point in the debate centers on women's potential competence in science (Tierney, 2010). From Harvard President Summer's (2005) suggestion that biology is to blame for women's lack of representation in the sciences, to Baron-Cohen's (Baron-Cohen, Knickmeyer, & Belmonte, 2005) contention that evolution has formed men to be more analytical, several prominent thinkers of the 21st century assert that women are biologically predisposed to be less competent in the sciences. Though meta-analyses suggest that very few demonstrable differences in aptitude exist on batteries of cognitive and verbal tests (Hyde, 2005; Hyde, Fennema, & Lamon, 1990), it is still widely argued that biological forces explain the gender gap in STEM (Baron-Cohen, 2003; Baron-Cohen et al., 2005; Hedges & Nowell, 1995; Kimura, 2007; Levy & Kimura, 2009). However, some attribute existing gender differences to social and cultural factors. Consistent with this argument are cross-cultural studies demonstrating that higher women's achievement on a national level is related to reduced school age gender-math gaps (Guiso, Monte, Sapienza, & Zingales, 2008). Similarly, nation-level implicit stereotypes about gender and science predicted nation-level gender gaps in 8th grade achievement in science and math (Nosek et al., 2009).

Using women's lack of competence to explain the gender gap is paired with the complementary argument that warmth is a barrier to women's careers in the sciences. In fact, Baron-Cohen's (2003) complementary argument to men being more analytical (thus giving them an advantage in science) is that women are more empathetic, simultaneously leading them away from science careers that require analytical skills and towards careers that use their natural gift of empathy. Recent research has suggested that women are underrepresented in STEM because they are drawn to jobs that are seen as communal in

nature (Diekman et al., 2010). In other words, women are too warm and caring to be interested in jobs in the sciences.

Deciding which piece of the human puzzle contributes the most (or at all) to the lack of women in academic science is a complex task (Ceci et al., 2009; Halpern et al., 2007; Sommers, 2009), and falls outside the scope of this study. What is important to note is that competence and warmth are central arguments employed to explain the deficit of women in STEM disciplines. It is then important to ask what effects do these perceptions have on how women are perceived as scientists, professors, and women.

Arguments for women's competence in science can be illuminated further if taken in context of their general perceived competence. Cuddy and her colleagues (2004) demonstrated that if female workers "gained" warmth (by having children), perceptions of their competence declined. Other research has demonstrated that typical women are perceived as high on warmth and low on competence, and female sub-groups in general are either high in warmth, low in competence (i.e. housewife, chick), or low in warmth, high in competence (i.e. women's libber, career woman) clusters (Eckes, 2002). These findings suggest that even if female STEM professors are perceived as competent, they may suffer from being perceived as less warm than other women (Cuddy, Fiske, & Glick, 2004). Alternatively, if female professors are perceived as "warmer" than male professors it may be at the expense of their perceived competence. Stereotype content model (SCM) provides a framework to understand these potential tradeoffs of warmth and competence.

Stereotype Content Model

In reaction to research that defines stereotypes as antipathy that triggers only negative or hostile feelings and actions, Fiske, Cuddy, Glick and Xu (2002) proposed a stereotype content model (SCM). This model contends that instead of being unilaterally negative, stereotypes incorporate both positive and negative attributes and can elicit a number of different reactions (pity, fear, anger, envy). Furthermore, all SCM can be understood as a combination of differing warmth and competence perceptions. According to the SCM, warmth and competence are universal aspects of social cognition, and thus of stereotype content (Fiske, Cuddy, & Glick, 2007). In other words, there are two major pieces of information that one needs to know about every human one meets: their intentions (warmth), and if they have the means (competence) to carry out these intentions. Traits such as friendliness, good-intentions, and good-nature underlie warmth; while competence is comprised of traits such as confidence, capability, and intelligence. Rules of thumb about a group's warmth and competence form the content of a stereotype (thus, stereotype content model).

The SCM provides a way to conceptualize stereotypes as multifaceted and ambivalent, as opposed to out-and-out negative evaluations. The two dimensions of SCM, warmth and competence, provide four general stereotype clusters: in-group, paternalistic stereotype, envious stereotype, and contempt (Fiske et al., 2007). In-group and reference-group members (i.e. middle-class, students) are generally perceived as both warm and competent (Cuddy, Fiske, & Glick, 2008). Paternalistic stereotypes are applied to out-groups who are seen as warm, but not competent (i.e housewives, elderly people). Envious stereotypes are applied to groups who are seen as competent, but not warm (i.e

feminists, Black professionals). Groups who are viewed as neither warm nor competent (i.e. homeless people, welfare recipients) elicit *contempt* and pity (Cuddy, et al., 2008). Thus, understanding the SCM for a group can be used to understand and predict people's attitudes and behaviors towards various groups (Cuddy et al., 2008).

Stereotype Content, Gender, and STEM

To my knowledge, scholars have not investigated the stereotype content of gendered professors in different disciplines. Though Cuddy and her colleagues (2004) found that professors are rated as relatively high in both warmth and competence, professors were rated in a male subgroup (so professors were assumed to be male), limiting the findings interpretability with regards to gender. Furthermore, the group "professor" was broad, and did not differentiate between disciplines.

Previous educational psychology research provides some clues regarding stereotype content. Male and female professors are generally held to different standards (Abel & Meltzer, 2007; Bennett, 1982; Miller & Chamberlin, 2000; Sprauge & Massoni, 2005). For example, female professors are expected to posses both warmth and competence, while male professors' are only expected to be competent (Kierstead, D'Agnostino, & Dill, 1988). This suggests that perceptions of female professors will be negative if they do not fit into stereotypical femininity (which is characterized by warmth).

Student ratings differ not only based on the gender of a professor (which are often used in tenure and promotional decisions), they also vary by class topic (Basow, 1995; Basow & Silberg, 1987; Centra & Gaubatz, 2000). For example, in a four-year

longitudinal study, male professors in the natural sciences, but not the humanities, were given significantly higher ratings than their female counterparts (Basow, 1995). These demonstrated interactions between gender and academic discipline, coupled with the scientific discourses about warmth and competence suggest academia is an ripe field in which to explore SCM.

The Current Research

This investigation will explore three aspects of stereotypes: explicit (Study 1) and implicit (Study 2) SCM and behavioral intention (Study 3) while extending stereotype content theory to subgroups of academia. Study 1 will use previously established stereotype content research methods to take a closer look at what stereotypes are held generally for professors of different genders in different disciplines. Study 2 will establish implicit measures of individual SCM (as measured by the ScIAT), and explore if a relationship exists between these two constructs. Study 3 will integrate stereotype content into an experimental paradigm that will explore implicit SCM, explicit SCM and behavioral intention.

Study 1

Though current scientific discourse deploys competence and warmth as barriers to women in the sciences, it is unclear if these beliefs are more widely held. This study will explore what perceptions about competence and warmth college undergraduates hold about male and female professors in the sciences and humanities. Though research activities can be the focus of an academic career, an academic career generally includes

teaching. Furthermore, undergraduates have a great deal of experience evaluating professors, making professor perception an excellent tool to begin exploring beliefs about male and female scientists. In keeping with the discourse that women are less competent than men in STEM disciplines, and that women's warmth further impedes their progress in STEM, this study hypothesizes:

1) There will be an interaction effect between gender of professor and discipline on ratings of warmth and competence. Specifically, female STEM professors will be rated as warmer and less competent than male professors. This pattern will differ in the humanities.

This study also seeks to address a concern about the current measurement of stereotype content. Generally SCM is measured by having participants rate specified groups (e.g. Germans or the mentally ill) on traits comprising warmth (e.g. friendly, well-intentioned, trustworthy, warm, good-natured, sincere) and competence (e.g. competent, confident, capable, efficient, intelligent, skillful). They are instructed to indicate what society thinks of these groups, not their own opinion. This procedure is designed to reduce social desirability effects. Fiske and colleagues (2007) imply that this measurement system, in circumventing social pressures, is actually a form of implicit measure. Because being aware of a societal attitude or belief may differ from individually possessing this attitude or belief, this study further hypothesizes that:

2) Asking about societal views will produce different results than asking about personal views.

Methods

Participants

106 students from the University of Hawaii at Manoa participated for course credit. Six participants were dropped for not correctly completing the survey, and a further 14 were dropped for having one or more missing values on key variables, leaving 86 participants[1]. The mean age of these remaining 86 participants (society= 39; you =47) the mean age was 21.95 years. Forty-sex were females, and ethnically they were quite diverse (27.9 % identified as Caucasian, 30.3 % Asian-American, 11.6% Asian, 11.6% Other, 4.7% Hawaiian/Part Hawaiian, 4.7% Hispanic, 2.3% African American/Black, 2.3% Native American, 2.3% Pacific Islander, 2.3% did not respond). 57% were mainly raised in Hawaii, and 76.7% were social science majors.

Materials

Two versions of a SCM survey was randomly distributed to participants. Both surveys asked participants to think about groups of male or female professors who belong to either humanities (literature, music, and women's studies) or science (biology, chemistry, physics) disciplines. Pre-testers rated these disciplines as highly characteristic of their respective fields.

Half of participants received a survey asking how *society* viewed the academic groups, while the other half responded with how *they* viewed the academic groups. For the society version of the survey, participants were prompted:

---

[1] Whether to impute missing values or simply drop cases is a difficult decision. Because the proposed analysis is multivariate in nature, the decision to delete listwise was made. Analysis of the difference of the average of scores (e.g. average of male humanities warmth) was performed, and the addition of these scores did not alter the current findings.

> *We are interested in how society views the following social groups. When responding, please think only about <gender> members of these groups. On a scale of 1-7 (1=not at all, 7=extremely) please indicate what society believes about the listed groups. Remember, we are not interested in your personal beliefs, but in how you think these groups are viewed by society.*

For personal opinion versions of the survey, participants were prompted:

> *We are interested in how you view the following social groups. When responding, please think only about <gender> members of these groups. On a scale of 1-7 (1=not at all, 7=extremely) please indicate what you believe about the listed groups. Remember, we are interested in your personal beliefs.*

Participants used a seven-point Likert scale to rate the specified group on a series of three warmth and three competence traits. A sample item for competence is "In your opinion, how intelligent are [female] biology professors?" A sample item for warmth is "In your opinion, how sincere are [male] biology professors?' These materials, methods and instructions are modified from previous studies (Eckes, 2002; Fiske, et al., 2002).

Data Reduction

All warmth items and competence items were found to have high reliability (both Cronbach's $\alpha = .96$), and were collapsed into separate composite scores for each gender and discipline (e.g. male humanities competence, female science warmth). These scores ranged from 9 to 63.

Results

Means for the composite scores are represented in Table 1.1. Composite scores were submitted to a 2 (Version: Society/You, between subjects) X2 (Gender: Male/Female) X 2 (Field: Humanities/Science) X 2 (SCM: Competence/Warmth) mixed ANOVA. There were no significant effects for the survey version. In order to ease

interpretation the survey version was dropped from the model, and a 2 (Gender) X 2 (Field) X 2 (SCM) repeated measures ANOVA was run. This did not alter the effects from the previous model. In this second ANOVA, all main effects (Gender, Field, SCM) and interactions (Gender*Field, Gender*SCM, Field*SCM, Field*Gender*SCM) were significant (see Table 1.2).

Table 1.1
Means for Composite Scores

	Humanities		Sciences	
Gender	Warmth	Competence	Warmth	Competence
Male				
Mean	45.03	48.31	41	51.84
SD	8.25	7.93	9.51	7.36
Female				
Mean	47.52	50.05	43.84	49.55
SD	8.70	8.19	8.22	7.99

The three-way interaction of Gender*Discipline*SCM tells us that the interaction of SCM and gender is not the same between disciplines ($F(1,85)=16.41$, $p=.00$, $r=.40$). Follow-up analysis[2] demonstrates that the interaction between SCM and gender is significant for ratings of science ($F(1,85)=45.14$, $p=.00$, $r=.36$), but not humanities ($F<1$), professors (Figure 1). In the humanities male professors are perceived as both less competent ($M=48.31$) and less warm ($M=45.04$) than female professors ($M= 50.05$, $t(85)=-2.7$, $p=.01$, $r=.18$; $M=47.52$, $t(85)=-3.49$, $p<.001$, $r=.20$). This is in contrast to the sciences where males are rated as more competent ($M=51.84$) than females ($M=49.55$,

[2] Follow-up analyses for the three-way interaction were performed across humanities scores (level 1) and science scores (level 2). These analyses were corrected with the original model's corresponding MSEs. The significance level for interactions at level 1 and level 2 were determined using an adjusted family wise error rate.

t(85)=3.33,p=.00, r=.19), but females are warmer (M=43.84) than males (M=41.00, t(85)=-3.83, p=.00, r=.21).

Table 1.2
2 (Gender) X 2 (Field) X 2 SCM) Repeated Measures ANOVA

Source	SS	df	MS	F	Sig.
Gender	244.33	1.00	244.33	12.50	0.00
Error	1661.67	85.00	19.55		
Field	237.23	1.00	237.23	4.13	0.05
Error	4882.27	85.00	57.44		
SCM	5369.31	1.00	5369.31	104.46	0.00
Error	4369.19	85.00	51.40		
Gender*Field	145.14	1.00	145.14	4.51	0.04
Error	2735.86	85.00	32.19		
Gender*SCM	372.15	1.00	372.15	19.23	0.00
Error	1644.86	85.00	19.35		
Field*Gender*SCM	1240.95	1.00	1240.95	41.92	0.00
Error	2516.55	85.00	29.61		
Gender	205.49	1.00	205.49	16.41	0.00
Error	1064.51	85.00	12.52		

Though the three-way interaction subsumes interpretation of other significant results, it is useful to consider the some of the two-way interactions in light of the three-way interaction (Kirk, 1995). A two-way interaction between gender and SCM (F(1,85)=19.231, p=.000, r=.43) reveals that though male and females have similar overall ratings on competence (M=50.08, 49.8, p>.05) female professors have a higher overall rating of warmth (M=45.68) than male professors (M=43.02, t(85)=-4.629, p=.000). A second interaction between discipline and SCM is such that humanities professors are rated as less competent (M=49.18) than science professors (M= 50.69,

t(85)=2.30, p=.02[3], r=.22). However, humanities professors are rated as generally warmer (M=46.28) than science professors (M=42.42, t(85)=-5.06, p=.00).

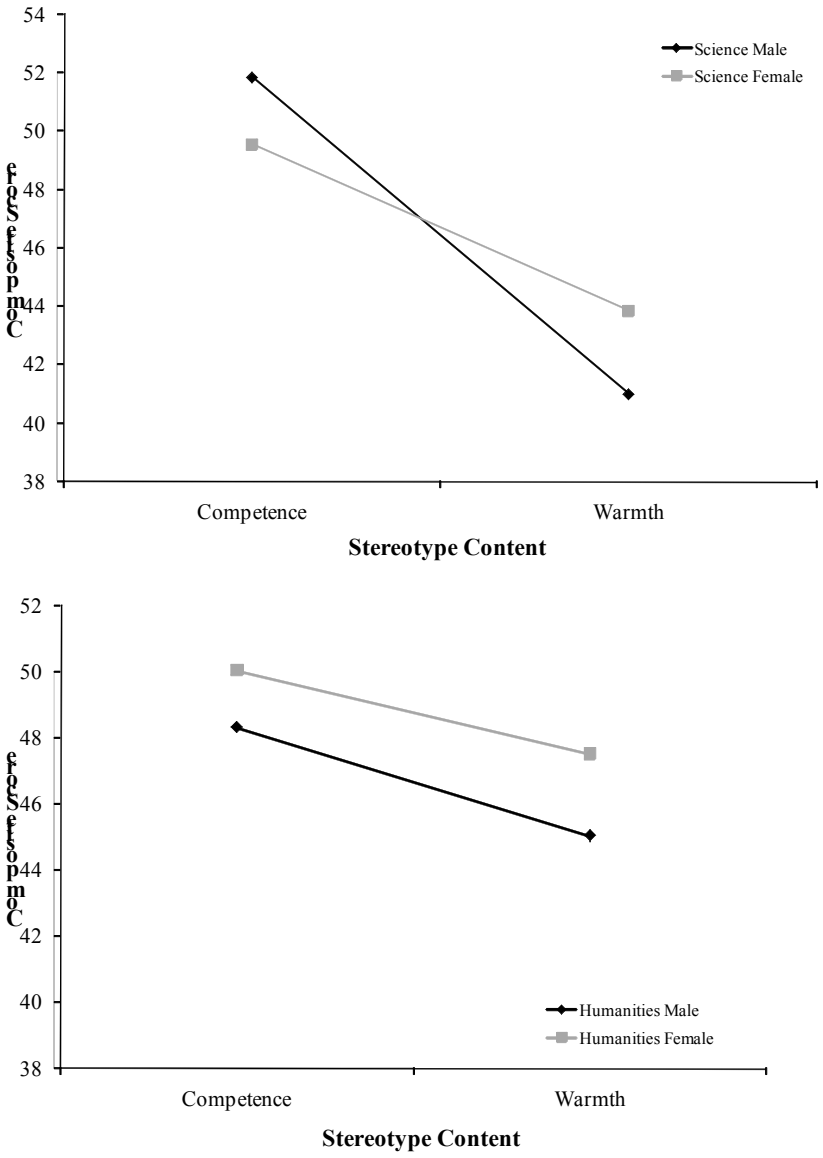

Figure 1. *Field x Gender X SCM*

[3] While there are no hard and fast rules about alpha corrections to follow-up simple effects analyses, it is common for them to be tested at the .05 level. Despite this, higher p values (that would not hold under a Bonferroni correction) should be treated with caution.

Discussion

Hypothesis 1 was supported by the data. Female professors in the sciences were rated as warmer and less competent than male professors in the sciences. This is in spite participants generally rating male and female professors as equally competent, and humanities professors as being less competent than science professors. These results support observations about the discourse surrounding barriers to women in the sciences. Assertions of women's lack of competence and abundance of warmth are not only forwarded in the scientific literature, these perceptions are also held by university students.[4]

These perceptions take on new meaning in the context of the data failing to support Hypothesis 2. Ratings received for the "society" and the "you' versions were not statistically different. Generally, stereotype content research has used "society" as a method to circumvent potential demand characteristics, however, at least in the case of professors, personal opinion and projections of what society thinks do not differ. This is surprising given that access to knowledge of social stereotypes and beliefs does not determine that an individual holds these beliefs.

There are two possible ways to interpret this finding. One is that, in fact, the current stereotype content measures *are* implicit in that participants are unaware of what is being tested. Due to the explicit nature of the questions, it seems unlikely that students would be unaware of what is being measured, though of course they were unaware of the

[4] Who are, at least in theory, the most likely to go on to have careers in STEM.

hypotheses of the study. While participants did not directly compare male and female professors in the humanities and sciences, participants did rate these groups sequentially. The more likely explanation is that participants were not motivated to hide their impressions of professor's warmth and competence. Perhaps the story about male and female scientists' warmth and competence is so strong that it is not seen as socially undesirable to report perceiving men as more competent and women as more warm. Future research should test this hypothesis by comparing explicit SCM measures to implicit measures.

It is interesting to consider relative warmth and competence of science professors in light of the humanities professions. Female humanities professors were perceived as both more competent and warmer than male humanities professors, however women are also underrepresented in the upper echelons academia of many humanities disciplines (such as history). Though history is a fairly solitary profession, with scholars generally working on individual, rather than collaborative projects, female warmth is not generally put forth as a barrier to women's successes in history. Similarly, it would seem that men are present in the humanities *despite* their perceived lack of competence. This study was not designed to investigate these particular questions, however the current data does raise some interesting questions about ways that stereotype content model may be related to role congruity theory (Eagly & Karau, 2002).

This study raises many questions for further research. One is how these perceptions affect judgments and behavior. Do these perceptions affect women's desire to participate in the sciences? How do they affect professor's ratings? Can these perceptions be altered? What if we manipulate perceptions of female scientists as warm?

Furthermore, since some SCM content might be more subject to demand characteristics, perceptions of professors set themselves up to be an excellent way to establish implicit measures of SCM.

This study has several limitations. The first is the population is confined to undergraduates at the University of Hawaii at Manoa, limiting generalizability of the results. The second is that even though the predicted effects were found, it is possible that implicit measures may illuminate how automatic these perceptions are. Finally, since these measures are self-reported, it is impossible to determine how these explicit judgments may affect actual behaviors.

Despite these limitations, the confirmation that female scientists are perceived as less competent and warmer than male scientists suggests that these beliefs are not merely forwarded by people actively engaged in the women in STEM debates, but are held by individuals. It is for future research to determine if and how these beliefs may themselves be barriers to women in STEM.

Study 2

Since SCM is a relatively new theory, the majority of research has been focused on validating it (Oldmeadow & Fiske, 2010). Researchers have established that SCM is a valid construct to describe stereotypes for a range of outgroups, and that it is applicable cross-culturally. Explicit societal SCM has been studied in several countries in Europe, East Asia, and Latin America (Cuddy, et al., 2008; Fiske, et al., 2007). Generally SCM is measured by asking what society, not individuals, thinks about social groups in order to

decrease demand characteristics. The ability to access and report societal beliefs can be very different, however, from individually holding those same beliefs.

Study 1 demonstrated that individuals can explicitly express SCM, at least in regards to professors. However it is unknown what role implicit biases may play in SCM. Because warmth and competence impressions are formed with only 100ms viewing time of a face (Willis & Todorov, 2006), there is good reason to believe that stereotype content may have implicit components. However, to my knowledge only one publication has addressed the implicit nature of stereotype content. Wade and Brewer (2006) investigated whether the female subgroups of businesswomen and housewives were implicitly associated with warmth and competence. Using a lexical decision task to test implicit activation of stereotype content, these researchers found that female participants responded more quickly than males to positive words (regardless of if they addressed warmth or competence) following a businesswoman prime, while male participants responded more quickly than females to positive words (regardless of their association with warmth or competence) following a housewife prime. They concluded that a general attitude towards groups was automatically activated, not the ambivalence present in stereotype content.

Wade and Brewer's (2006) initial look into implicit aspects of stereotype content could be improved in several ways. First, it does not directly compare individual's explicit SCM with their performance on the implicit task. It is possible that their sample was simply low in individual endorsement of SCM. Second, lexical decision-making tasks also may not be the best tool to investigate implicit aspects of SCM. These tasks are generally used to investigate semantic priming, which may not be sensitive to

associations between groups and their perceived warmth and competence. The Implicit Associations Task (IAT) may be a more appropriate tool.

The IAT rests on the simple assumption that people co-categorize objects and ideas faster if they have close associations (Greenwald, McGhee, & Schwartz, 1998). The standard IAT asks participants to sort two categories (e.g. *Male/Female*) pairs with two attributions (e.g. *Competence/Warmth*), and measures the reaction time difference between these associations. The automatic nature of the measure allows researchers to reach beyond self-report (which can be influenced by many factors) to assess spontaneous responses.

The IAT can only measure associations between *Male* and *Competence* compared to associations between *Female* and *Warmth* (and vice versa). The SCM conceptualization of perceptions varying along a continuum of warmth and competence allows social groups to possess any combination of warmth and competency. This multi-axied measure does not logically align to a single-dimensional IAT. What is needed is a way to measure if gender is associated with warmth.

In recent years the Single Category IAT (ScIAT) has been developed to test preferences and attitudes that may vary along several dimensions (Karpinski & Steinman, 2006). The ScIAT measures the association between two categories (*Male/Female)* and a single attribution (*Competence).* This allows, for example, males to be associated with competence without effecting their association with warmth. This means that ScIAT scores should be able to map onto explicit SCM scores.

This initial study will attempt to replicate the explicit SCM findings from Study 1 and investigate the implicit nature of these SCM. Using a version of previously

established paradigms (Fiske, et al., 2002). Implicit individual level SCM data will be collected using ScIATs to explore the relationship between implicit associations between male/female, science/humanities, and warmth/competence.

Hypotheses:

1) Explicit scores will replicate findings from Study 1: There will be an interaction effect between gender of professor and discipline on ratings of warmth and competence. Specifically, the ratings of female and male professors in the STEM disciplines will differ on warmth and competence, and this pattern will be different for professors in the humanities.
2) There will be significantly higher implicit associations (as measured by the ScIAT) between a) science and competence b) humanities and warmth c) male and competence and d) female and warmth.
3) Explicit competency and warmth scores for gendered professors in different disciplines will correlate to corresponding implicit associations.

Methods

Participants

Participants were 79 ethnically diverse (49.4% Asian, 19% White/EuroAmerican, 11.4% Multi-racial, 20.2% Other), university students recruited for an online study. 54.4% identified as female, with a mean age of 22.12. Over half of the sample (63.3%) had been raised in the continuous United States. Participants were recruited using targeted online strategies that included message board postings as well as more traditional classroom recruitment.

Materials

*Stereotype Content***.** Participants rated both male and female professors[5] belonging to STEM (biology, chemistry, physics) and the humanities (literature, history, art history)[6] disciplines. Participants rated these groups on three traits relating to warmth and three traits relating to competence using a 7-point scale [1=not at all, 7=extremely; (Fiske, et al., 2002)]. Participants were asked about their opinion, as opposed to societies', as the previous study demonstrated that society and individual viewpoints about professors did not differ, and to ensure that explicit responses reflected an individual's opinion. A sample item for competence is "In your opinion, how intelligent are [female] biology professors?" A sample item for warmth is "In your opinion, how sincere are [male] biology professors?" These materials, methods, and instructions replicate Study 1 with two exceptions: instead of rating professors in a counterbalanced order the computer protocol randomized the order of professor groups and questions within a professor group were also randomized.

ScIATs. For this study, two sets of single-category IATs (ScIAT) were created. The first was designed to test how general academic disciplines are associated with warmth and competence, and the second was designed to test the association between gender and warmth and competence. The ScIAT, as opposed to the IAT, measures the

---

[5] Study 2 was also performed in an experimental format where participants were randomly assigned to rate either male or female professor groups. The results of this experimental study mirror the current study. They are not the main focus of analysis, as the sample size for this study is larger.

[6] Inspection of the previous study suggested that male women's studies professors were seen as less competent than other male humanities professors. Thus art history, which was rated as similarly representative of humanities in a pretest, replaced women's studies. Why male women's studies professors should be rated as less competent is an extremely interesting question that is not addressed in this study.

association between one concept (such as "warmth") and two categories. Thus, two ScIATs were created to test the relationship between the target category (male/female or science/humanities) and competence and warmth (e.g. a discipline-competence ScIAT, and a discipline-warmth ScIAT), resulting in four ScIATs.

The ScIATs were created and run using Inquisit software. Inquisit software allows millisecond accurate reaction times to be collected using an online sample. These ScIAts used established Science/Humanities words[7] and established gender words[8] (Nosek, et al., 2009). The Sc-IATs consisted of five blocks each, and were a modified version of the ScIAT presented in Karpinski (Karpinski & Steinman, 2006). In the first block, two words (e.g. "Science" & "Humanities") appeared the opposite top corners of the screen. A series of words related to the sciences and humanities flashed randomly across the screen, and participants categorized the words as quickly as possible by pressing appropriate pre-assigned keys (such as "E" for "Science" and "I" for "Humanities"). If participants made an error, a red X appeared on the screen. Participants then corrected their error, and this total response time was recorded. In the second block, a "practice" block consisting of 19 trials, the category of "Science" is paired with "warmth". A series of words related to the sciences, humanities, and warmth flashed randomly across the screen, and participants categorized the words as quickly as possible by pressing appropriate pre-assigned keys (such as "E" for "Science" OR "warmth" and "I" for "Humanities"). The third block, a test block, is identical to the second block, except it is

[7] Science: Biology, Physics, Chemistry, Math, Geology, Astrology, Engineering; Liberal Arts: Philosophy, Humanities, Arts, Literature, English, Music, History.
[8] Male: Man, Boy, Father, Male, Grandpa, Husband, Son, Uncle; Female: Girl, Female, Aunt, Daughter, Wife, Woman, Mother, Grandma.

40 trials long. The fourth (practice) and fifth (test) block repeat the second and third blocks, except "warmth" is paired with "humanities".

Procedure

After clicking on a link provided in a recruitment email, participants arrived at the experiment website. They were given informed consent form, and instructed to click "START" if they wished to participate. They then took the four ScIATs and responded to the explicit warmth/competence rating questions. The order of the ScIATs and explicit ratings were counterbalanced to diminish any order effects, and the rating items were presented in randomized order. Participants also filled out a brief demographics form at the end. When they had completed the task, a thank you screen debriefed them with the purpose of the study and thanked them for their participation.

Data Analyses

*Stereotype Content***.** All warmth items and competence items were found to have high reliability (Cronbach's α = .921, .945, respectively), and were collapsed into scores for each gender and discipline (e.g. male humanities competence, female science warmth).

ScIATs. Reaction times to the ScIATs were transformed into D-scores using an adaptation of the improved D score algorithm (Greenwald, Nosek, & Banaji, 2003)[9]. Only 51 participants had D scores that were not eliminated due to poor reaction times or

[9] Trials with response times longer than 1000ms were deleted, and participants who had more than 10% of trials faster than 300ms or 10% of categorization errors were removed from analysis. Compatible blocks were subtracted from their incompatible blocks (such that larger D scores correspond with a stronger association between the "compatible" categories). These blocks were then divided by their pooled standard deviation, and the related practice and test blocks were averaged.

excess errors, and analysis will use this lower n-size only when taking into account these implicit measures.

Results

Explicit Ratings

Means for the composite scores are represented in Table 2.1. The explicit composite scores were submitted to a full factorial 2 (gender of professor) X 2 (science/humanities) X 2 (warmth/competence) repeated-measures ANOVA. All main effects (Male/Female, Humanities/Science, Warmth/Competence) and interactions except for the gender*field interaction were significant.

Table 2.1
Means for Composite Scores

	Humanities		Sciences	
Gender	Warmth	Competence	Warmth	Competence
Male				
Mean	44.58	48.14	40.23	49.63
SD	8.08	7.66	9.25	7.78
Female				
Mean	46.36	48.10	42.90	48.37
SD	8.59	7.93	8.87	8.01

The predicted three-way interaction of gender, discipline, and competence/warmth was obtained (See Table 2.2, $F(1,77)=6.09$, $p=.02$, $r=.27$). Follow-up analyses[10] demonstrate that the interaction between SCM and gender was significant for ratings of science ($F(1,77)= 42.41$, $p<.05$, $r=.6$) and humanities ($F(1,77)=9.13$, $p<.05$, $r=.33$) professors (Figure 2).

[10] Follow-up analyses for the three-way interaction were performed across humanities scores (level 1) and science scores (level 2). These analyses were corrected with the original model's corresponding MSEs. The significance level for interactions at level 1 and level 2 were determined using an adjusted family wise error rate.

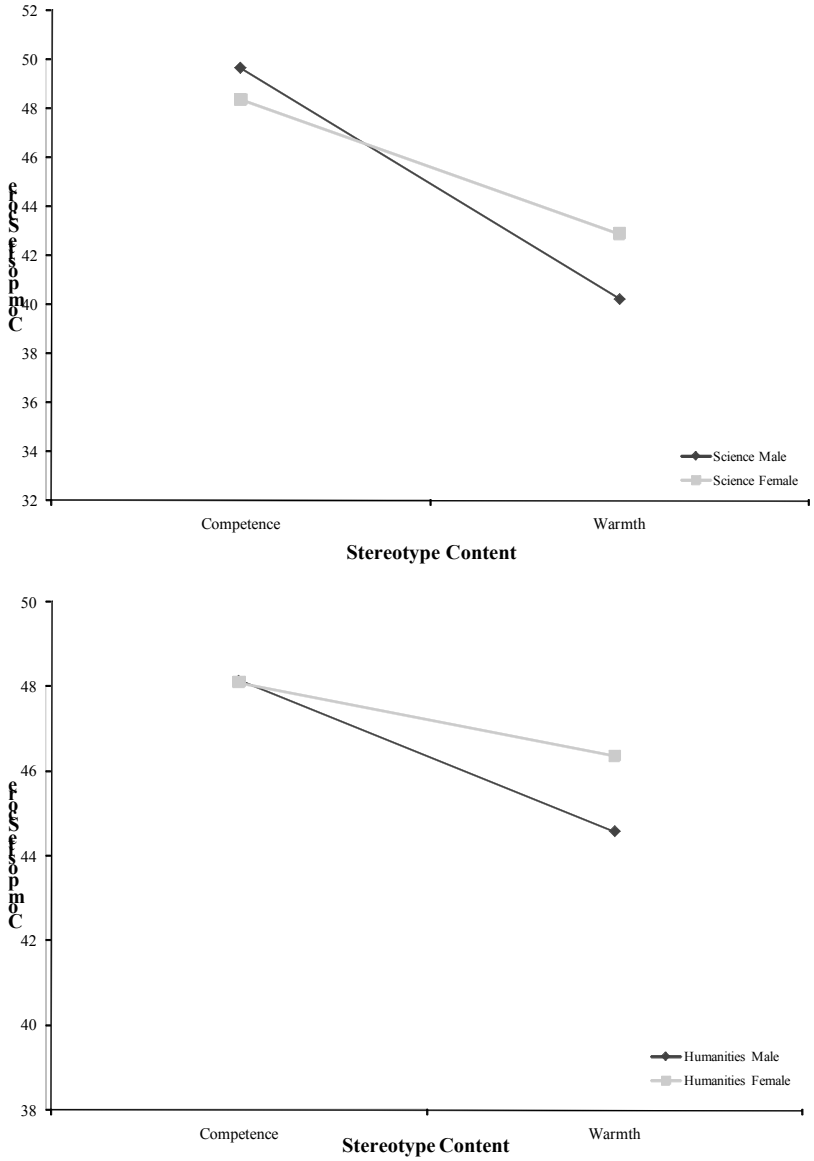

Figure 2. *Field X Gender X SCM*

In the sciences males are rated as more competent (M=49.63) than females (M=48.37, t(77)=2.53, p=.01, r=.18), but females are seen as warmer (M= 42.9) than males (M=40.23, t(77)=-4.03, p=.00, r=.22). This pattern did not hold for the humanities. There was no effect of gender on competence (t(77)=.08, p=.94), however female

professors were perceived as warmer (M=46.36) than male professors (M=44.58, t(77)=-2.44, p=.02, r=.18).

Table 2.2
2 (Gender) X 2 (Field) X 2 SCM) Repeated Measures ANOVA

Source	SS	df	MS	F	p
Gender	96.98	1	96.98	5.13	0.03
Error	1456.52	77	18.92		
Field	357.03	1	357.03	10.55	0.00
Error	2605.47	77	33.84		
SCM	3970.31	1	3970.31	71.65	0.00
Error	4266.69	77	55.41		
Gender*Field	1.08	1	1.08	0.09	0.77
Error	969.42	77	12.59		
Gender*SCM	321.64	1	321.64	17.05	0.00
Error	1452.36	77	18.86		
Field*SCM	891.85	1.00	891.85	30.57	0.00
Error	2246.15	77.00	29.17		
Field*Gender*SCM	43.10	1	43.10	6.09	0.02
Error	544.90	77	7.08		

Though this three-way interaction subsumes all other results, and makes interpretation of other effects difficult, unnecessary, or misleading (Kirk, 1995), it is useful to examine the other effects in light of the three-way interaction. An interaction effect was found between gender and SCM (F (1,77)=17.053, p=.00, r=.43). Females were perceived as being warmer (M=44.63) than males (M=42.4, t(77)=-3.76, p=.00, r=.22). The difference between overall ratings of female (M=48.24) and male (M=48.89) competence only approached significance (t(77)= 1.77, p=.08). Similarly, and interaction between field and SCM (F(1,77)=30.57, p=.00, r=.53) reveals that the humanities were perceived as warmer (M=45.47) than the sciences (M=41.56, t(77)=-5.174, p=.00, r=.25), but only marginally less competent (t(77)=1.8, p=.08[11]).

[11] Due to the large number of analyses, this marginal result should be considered carefully.

Sc*IAT Analysis.*

The means for the ScIATS are presented in Table 2.3. The computed D scores were submitted to one-sample t-tests. As predicted warmth was implicitly associated with female (t(51)=3.529, p=.00, r=.25), though competence was not implicitly associated with either gender (t(51)=.743, p=.46, r=.12). As predicted, humanities was significantly associated with warmth (t(51)=8.25, p=.00, r=.37). Counter to prediction, however, humanities was also significantly associated with competence (t(51)= -6.34, p=.00, r=.33).

Table 2.3
ScIAT One-Sample T-Test

	d_gc	d_gw	d_fc	d_fw
Mean	0.03	0.15	-0.25	0.33
SD	0.33	0.31	0.28	0.29
T	.74	3.53***	-6.34***	8.25***

Note: Higher d_gc scores indicate an association between male and competence, higher d_gw scores indicate an association between female and warmth. Higher d_fc scores indicate an association between science and competence, higher d_fw scores indicate an association between humanities and warmth. ***p<.001

Correlational analysis revealed a significant negative relationship between field warmth and field competence (r=-.407, p=.00), such that as people had higher associations between science and competence, their evaluations of humanities and warmth decreased. There were no other significant correlations between implicit measures (Table 2.4). There was only one significant correlation between implicit and explicit measures. Implicit associations between field and warmth were significantly and positively related to explicit scores of females' warmth in the humanities (r=.34, p=.01).

Table 2.4
Pearson Correlation Matrix among ScIAT Measures

	d_gw	d_fc	d_fw
d_gc	.01	-.08	-.08
d_gw		-.09	.177
d_fc			-.407**

**p < 0.01

Discussion

The explicit results from this study replicated findings from Study 1. In the sciences, participants explicitly evaluate females as less competent than, and males as less warm than females. Because male and female professors in the humanities are generally rated as equally competent, this suggests that there is something special about the relationship between STEM disciplines and gender that is affecting competence ratings as opposed to women in general being viewed as less competent then men. This adds strength to the findings in Study 1 by replicating the previous results with a different sample.

This study also attempted to establish individual implicit measures of warmth and competence. Gender and warmth, and field and warmth were significantly associated--that is that female and humanities had a stronger association with warmth than did male and science. Competence, however, did not have a significant association with either male or female, and it had a stronger association with humanities than science.

Two of the implicit scores showed associations in the expected direction. Warmth and female were significantly associated, which is consistent with research that demonstrates a connection between women and perceptions of warmth (Eckes, 2002). This association does not come as a surprise, and indeed reflects results that women are also strongly associated with home and family (as opposed to career) (Nosek, Banaji, &

Greenwald, 2002). A second significant association was demonstrated between warmth and the humanities. This result is also consistent with the explicit results of this study. A third significant association between humanities and competence was not in the expected direction. In the previous study professors in science disciplines were generally rated as more competent than humanities professors. This directly conflicts with previous explicit studies that report that the sciences are rated as higher in competence. This calls into question the validity of this implicit measure, though of course it is also reasonable to call the explicit measure into question. Let us briefly consider the relationship between the implicit and explicit measures before coming back to this important topic.

There was only one significant relationship between the explicit and implicit measures. As implicit scores associating humanities and warmth increased, so did explicit female humanities warmth scores. Though the lack of correlation to explicit measures paired with some counter-intuitive association directions raises serious questions about the validity of the implicit measures, I will still briefly consider some of the issues raised by these findings.

It is difficult to interpret the implicit measures in light of the explicit measures. Why do the implicit scores fail to predict explicit ratings? Such findings are common for sensitive evaluations, however it is generally in the direction of finding evidence of implicit bias when there is no explicit bias. Why then do people readily express explicit evaluations of the sciences and men as more competent, but fail to produce these results in implicit measures?

Perhaps participants are espousing explicit preferences that do not reflect their implicit beliefs. Yet, the majority of research on implicit and explicit bias does not

support this possibility. Instead, studies show that people's explicit bias is the result of their implicit bias not being filtered (due to motivation, processing capacity, etc) (Correll, 2008; Gonsalkorale, Sherman, & Klauer, 2009; Parks-Stamm, Heilman, & Hearns, 2008). The results of this paper do not support this proposal either. In fact, two of the ScIATs demonstrated implicit associations in the predicted directions, suggesting that at the very least there are some strong associations between warmth, women, and humanities disciplines. Still, even the predicted associations are not related to the explicit measures.

An alternative explanation is that the ScIAT measure is not sensitive enough to fully capture the complexity of evaluating a gendered professor. A female scientist may require a different activation than either a female or science. The ScIAT developed for this study does not have the ability to co-evaluate the categories of "gender" and "discipline," so the failure for these to map onto explicit evaluations supports this interpretation. Further evidence in support of this idea is the significant correlation between implicit scores of humanities and warmth and explicit ratings of female warmth in the humanities. Previous research has demonstrated that females are implicitly associated with the humanities (Nosek, et al., 2002; Nosek, et al., 2009), so it is possible that associations between female, humanties, and warmth are entwined in a way that the other categories are not.

This study has several limitations. The ScIAT, which traditionally has a greater number of errors and higher score drop out, lead to a large amount of rejected data. While 78 participants produced useable data for the explicit measures, only 51 of those produced valid enough d_scores to warrant interpretation. This 35% drop rate raises further concerns about the validity of interpreting the ScIAT scores. While the explicit

scores demonstrate a general belief that males in the sciences are more competent and less warm than females in the sciences, this study does not address how these beliefs about broad groups translate into judgments, intention or behavior about specific individuals. Future research should investigate how these explicit beliefs about female and male scientists effect evaluations about individuals in the field.

Study 3

Beyond knowledge that stereotypes exist, it is important to understand how (and if) people use them. In recent years, researchers have started exploring how SCM may inform behavioral intent. SCM has been shown to predict behavioral intent towards fictitious immigrant groups (Cuddy, Fiske, & Glick, 2007), and group-project partners (Cuddy, et al., 2008). Still, exploring the relationship between SCM behavioral intent is relatively new, and leaves many avenues to be explored. This question holds particular relevance in the academic world, where perceived competence may alter how other professors, administrators, and students react to an individual. It may influence student's ratings of professors which may be used to make tenure and other promotional decisions.

To my knowledge, no research has explicitly merged stereotype content and professor evaluation ratings, though previous research suggests that female professors are expected to be warmer than their male counterparts, and that a failure to meet this stereotype may adversely affect students' evaluations (Basow & Silberg, 1987; Kierstead, et al., 1988; Sprauge & Massoni, 2005). This study investigates how students may form impressions of professors on the warmth/competency scales, how implicit SCM is related to these impressions, and how these perceptions may affect their recommendations of a

professor. It will extend previous findings into an experimental paradigm that is familiar to students: lecture and lecturer evaluation.

Hypotheses:

1) Female and male lecturers' competency and warmth scores will mirror results from Study 1 and Study 2. Specifically, female professors will be seen as warmer and less competent than are male professors.
2) Explicit competency and warmth scores will predict corresponding implicit associations. Specifically, (a) higher associations on the IAT between male and competence will predict higher explicit ratings of male professors' competence, and lower ratings of female professors' competence and (b) higher associations on the IAT between female and warmth will predict higher explicit ratings of female professors' warmth, and lower ratings of male professors' warmth.
3) (a) Implicit measurements (from the ScIAT) will be better predictors for (a) over all ratings for lecturers and (b) behavioral intentions than explicit measurements of SCM (c) gender will moderate the prediction of overall ratings from warmth ratings.

Methods

Participants

134 participants initiated the online computer protocol. Twenty-three were removed from analysis for failing to complete the study, and 10 were removed for having previous taken a similar study. Of the remaining 101, 60 were shown a female lecturer,

and 41 were shown a male lecturer[12]. The mean age of participants was 24.11 (SD=8.54), and 64.4% of participants identified as female. 69.3% were currently enrolled in an undergraduate institution, with the remaining 30.7% having at least completed a bachelor's degree. Participants had declared a variety of majors, with 39.6% majoring a social science, 23.8% majoring in a science, and 16.8% majoring in humanities. Participants were ethnically diverse, with 46.5% identifying as Asian, 30.7% identifying as White/EuropeanAmerican, 7.9% identifying at Pacific Islander, and 5.9% identifying as Multi-Racial. The majority of participants had been raised in Hawaii (61.4%). Participants were recruited using targeted online strategies that included message board postings as well as more traditional classroom recruitment.

Materials

Lecture. Three brief lectures (approximately three minutes) were developed about advanced, but basic, scientific processes from three disciplines (the lac operon/biology, steel corrosion/chemistry, and ideal gas law/physics). These processes were chosen in consultation with PhDs with expertise knowledge in the relevant discipline. College-level textbooks were also consulted to ensure a professor in the field could reasonably be expected to cover this material at some point in his or her career. These written lectures were matched on for reading level using the Flesch-Kincaid Grade Level and for ease of reading using the Flesch-Kincaid Reading Ease Score.

Lecturer. Six photos (three male, three female) of lecturers in business dress were pre-tested and matched on age and ratings of attractiveness. These photos were front facing, and were placed on a neutral background. These photos were inserted into an

[12] This raises concerns about the validity of the reported results.

introductory explanation detailing information about the lecturer (Dr. Anderson) and the computer-based learning module that the lecturer developed.

Ratings. Ratings were collected on both the lecturer and the lecture. The four items for the lecture assessed interest, attractiveness, motivation, and comprehensibility. Sample items for the lecture included, “How comprehensive did you find that lecture?”

Target items for the lecturer measured their perceived warmth and competency. These six items were modified warmth and competence questions from Study 1 and Study 2. Sample items for the lecturer included, “How competent did you find the lecturer?”

Behavioral Intent. Two behavioral intent questions were included to investigate how likely students were to take a course with or recommend this lecturer. These questions were "*Based on this lecture, how likely would you be to recommend this lecturer to someone who wanted to learn about this topic?*" (BI1) and *"Based on this lecture, how likely would you be to take a required course from this lecturer?"* (BI2).

ScIATs. For this study, the two sets of single-category ScIATs replicated those used in Study 2. These four ScIATs implicitly measure the relationship between gender (male/female) and field (science/humanities) and SCM (competence and warmth).

Procedure

Participants were directed to the Millisecond website (www.millisecond.com) to take the online computer protocol by clicking on a link provided in a recruitment email. After receiving informed consent, they were instructed to click “START” if they wished to participate. The computer protocol randomly assigned participants to either the male or female lecturer condition (lecturer photographs and discipline were also randomly

assigned). Participants then read about Dr. Anderson's lecture series, and then read a lecture. After the lecture, participants responded to several questions about the lecture and lecturer. The order of lecture and lecturer rating sections was counterbalanced, and ratings were randomized within each section. Participants then took the four ScIATs. The order of the ScIATs were counterbalanced to diminish any order effects. Participants also filled out a brief demographics form. When they had completed the experiment, a thank you screen debriefed them and thanked them for their participation.

Data Reduction

Explicit ratings. Competence items and warmth items had an acceptable internal reliability (α= .72, .76) and were collapsed into composite competence and warmth scale indexes with a potential minimum score of three, and a maximum score of 21. The four items assessing the lecture were highly reliable (α= .9) and were collapsed into a composite lecture-rating index (possible minimum four, maximum 24).

ScIATs. Reaction times to the ScIATs were transformed into D-scores using an adaptation of the improved D score algorithm (Greenwald, et al., 2003).[13] This signal analysis resulted in only 67 participants having interpretable D scores. Analyses that include D scores will use this reduced N size.

[13] Trials with response times longer than 1000ms were deleted, and participants who had > 10% of trials faster than 300ms or >10% of categorization errors were removed from analysis. Compatible blocks were subtracted from their incompatible blocks (such that larger D scores correspond with a stronger association between the "compatible" categories). These blocks were then divided by their pooled standard deviation, and the related practice and test blocks were averaged.

Results

Mean ratings for composite competence, warmth, lecture index scale, and the two behavioral intent questions can be found in Table 3.1.

Table 3.1.
Means and Standard Deviations of Dependant Variables

	Competence	Warmth	Lecture Index	BI1	BI2
Mean	15.83	12.47	22.49	4.29	3.70
SD	2.86	2.91	7.88	1.53	1.59

An exploratory one-way MANOVA on gender of participant revealed significant effects on ratings of competency and warmth, so participant gender was included in all further analysis. Competence and warmth scores were submitted to a 2 (gender of participant) X 2 (gender of professor) MANOVA. There were no significant multivariate effects of gender of professor or gender of participant ($p>.05$). The interaction between gender of participant and gender of professor approached significance ($p=.88$).

Sc*IAT Analysis*

The means for the ScIATS are presented in Table 3.2. The computed D scores were submitted to one-sample t-tests. Warmth was implicitly associated with humanities ($t(66)=10.59$, $p=.00$, $r=.41$), as was competence ($t(51)= -6.82$, $p=.00$, $r=.33$). Warmth and competence were associated with neither gender ($p>.05$).

Table 3.2
ScIAT Means

	d_gc	d_gw	d_fc	d_fw
Mean	-0.03	0.07	-0.22	0.42
SD	0.3	0.37	0.27	0.33
t	-.77	1.66	-6.82***	10.59***

Note: Higher d_gc scores indicate an association between male and competence, higher d_gw scores indicate an association between female and warmth. Higher d_fc scores indicate an association between science and competence, higher d_fw scores indicate an association between humanities and warmth. ***p<.001

To test if implicit associations predicted explicit evaluations two regressions were run. The first predicted professor competence ratings by gender of professor (g), implicit association of competence and gender (d_gc), and the interaction between the two (g*d_gc). The overall model did not significantly predict explicit competence evaluations ($F(4, 62)=1.6$, $p=.19$). The second submitted gender of professor (g), implicit association of warmth and gender (d_gw), and the interaction between the two (g*d_gw). The overall model did not significantly predict explicit warmth evaluations ($F(3, 63)=.24$, $p=.87$).

To assess hypothesis 3(a) warmth scores, competency scores, and all ScIAT scores were entered into a forward stepwise regression predicting overall lecture scores (regression coefficients are presented in Table 3.3). The final model included warmth scores and competency scores ($F(2,64)=24.4$, $p<.001$) and accounted for 41% of the variance in the model. As ratings of professor warmth increased one point, overall lecture ratings increased by 1.29 points. Similarly, ratings of professor competence increasing by a point predicted a .86 increase in overall lecture ratings.

Table 3.3.

Dependant Overall Lecture Score

	b	SE	ß
Step 1			
Constant	1.69	3.55	
Warmth	1.67	0.27	0.61***
Step 2			
Constant	-7.05	4.74	
Warmth	1.29	0.3	0.47***
Competence	0.86	0.33	0.28**

Note. $R^2= .36$ for Step 1: $\Delta R^2= .06$ (ps<.05). * p<.05, ** p<.01, ***p<.001

To assess hypothesis 3(b), warmth scores, competency scores, and all ScIAT

scores were entered into a forward stepwise regression predicting B1 (regression coefficients are presented in Table 3.4). The final model included only warmth scores ($F(1,65)=27.94$, $p<.001$) and it accounted for 29% of the variance in the model. As ratings of professor warmth increased one point, B1 increased by .29 points.

Table 3.4

Dependent B1

	b	SE	ß
Constant	.59	.71	
Warmth	.29	0.05	0.55***

* p<.05, ** p<.01, ***p<.0001

To assess hypothesis 3(c), warmth scores, competency scores, and all ScIAT scores were entered into a forward stepwise regression predicting B2 (regression coefficients are presented in Table 3.5). The final model included only warmth scores ($F(1,65)=48.02$, $p<.001$) and it accounted for 41% of the variance in the model. As ratings of professor warmth increased one point, B2 increased by .37 points.

Table 3.5

Dependant B2

	b	SE	ß
Constant	-.94	.71	
Warmth	.37	.05	0.55***

* p<.05, ** p<.01, ***p<.0001

To analyze the effect of gender of professor and explicit warmth scores on overall lecture ratings a regression was run to predict overall lecture ratings with gender of professor, explicit warmth scores, and the interaction between the two[14]. The overall model was significant ($F(3,97)=18.24$, $p<.001$) and explained 34% of the variance in

[14] Warmth scores were centered to increase interpretability.

overall lecture scores, however only explicit warmth ratings significantly predicted overall lecture evaluations (t=6.37, p=.00). For every point above average warmth scores that a professor was rated, their overall lecture score went up by almost two points. Gender and the interaction effect were both nonsignificant (regression coefficients are presented in Table 3.6, ps>.05).

Table 3.6

Dependant Overall Lecture Scores

	b	SE	ß
Constant	23.49	.816	
Professor Gender	-1.43	1.28	-.09
Warmth	1.79	0.28	0.66***
Gender*Warmth	-0.57	0.46	-1.24

Note. ***p<.0001

Discussion

The first hypothesis, that female professors would be perceived as warmer and less competent than male professors, was not supported by the data. In fact, there were no significant differences between male and female professors. Why should this be when two previous studies have demonstrated this effect? There are two plausible, if not mutually exclusive, explanations for this: category versus exemplar effects and lack of ambiguity.

The first two studies in this investigation asked participants to think about general groups of professors while this study presented participants with a specific person to evaluate. Researchers have demonstrated that people's evaluations of an exemplar of a category are often evaluated differently than the general social group, especially if the exemplar is non-normative (as a female physics professor might be) (Groom, Sherman,

Lu, Conrey, & Keijzer, 2005). It is possible that this cognitive bias is driving the difference between studies. Future research should investigate this possibility.

Similarly, previous studies have asked people to evaluate general groups of *professors* without offering any anchor points regarding their competence. This study, however, presented an explicitly titled *Dr.* which may have conveyed a high level of competence. Bias is most likely to be present in ambiguous situations, and it is possible that the title of *Dr.* does not leave enough room for ambiguity, especially amongst a population largely made up of undergraduates. Future research should investigate the experimental effects of a more ambiguous situation. This could include removing the *Dr.* designation in the current paradigm, focusing on a different subject population (fellow experts), or similarly changing the scenario to involve interaction with a potential peer group member.

As in study two, explicit competency and warmth scores were not related to implicit scores. This is not surprising, since the ScIAT is measuring cognitively simpler categories than the explicit measures. The ScIAT has been used mainly to assess attitudes that are clearly positive or negative (Karpinski & Steinman, 2006; Nevid & McClelland, 2010), and it is possible that it is more suitable for good/bad attitude measurement. The continuing association between humanities and competence raises questions about the validity of the ScIAT, however this too warrants further research, as it would be interesting indeed if people explicitly report sciences as more competent but implicitly associate competence with the humanities. Future research should systematically address this question.

In keeping with previous educational research, overall ratings of the lectures, as well as intention to take or recommend a class, were solely predicted by explicit ratings of professor warmth. No effects for gender of professor were found, suggesting that warmth trumps competence when evaluating a lecturer.

Some of the potential limitations of this study have already been addressed, however a few more are worth nothing. Although effort was made to recruit science majors and people who were not raised in Hawaii, these groups still comprised a minority of the sample, limiting the generalizability of this study. This study also relies on self-reported behavioral intent questions to assess how people might recommend or take a course from a professor, which may or may not have a relationship with a range of actual behaviors towards a professor.

General Discussion

These three studies investigated stereotype content model in the context of the "leaky pipeline" of women in STEM. Two major discourses are deployed in the academic discussion of this pipeline problem: women's lack of competence (compared to men's) and women's overabundance of warmth (their desire to interact with others and have family priorities). These arguments are put forth as reasons, sometimes evolutionary (Baron-Cohen, 2003), sometimes choice-driven (Ceci & Williams, 2010; 2011), as to why women are a minority in STEM. Though scholars have recently argued that "historical" forms of discrimination are no longer barriers to women in STEM (Ceci & Williams, 2011) it is possible that the talk about women's competence and women's

warmth represent stereotypes held about female scientists, and themselves form barriers to women's advancement.

The first two studies support this contention. Two different samples reported that women in the sciences are perceived as less competent and warmer than their male counterparts. Regardless of whether they were espousing society's opinion or their own, the message stayed the same. Preconceived ideas about warmth and competence are applied to female scientists, at least as a group. How these ideas affect women and their presence in the sciences is a rich area for future research.

The final study, which looked at specific professors as opposed to general groups, did not replicate the gender differences in SCM between male and female professors. While surprising, this finding does open up several areas of potential investigation. When is SCM employed? Does it apply more to social groups than individuals? If SCM is the underlying construct of all stereotypes, does it then behave in the same manner as other forms of bias? What role does ambiguity play in SCM?

I also attempted to establish implicit measures for warmth and competence, the components of stereotype content. Though some of the implicit measures demonstrated significant associations, these implicit scores failed to correlate with explicit scores. A recent meta-analysis found that poor correlation is not uncommon for implicit and explicit measures for socially sensitive issues, however it is less common for attitudes and beliefs that participants have ready access to and are willing to share (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). It is troubling that the implicit measures of warmth and competence were not correlated with explicit reports of these constructs. It is

possible that implicit associations between gender and SCM, and field and SCM do not map upon the more complex idea of gender and scientists.

Despite the initial lack of correlation between implicit and explicit measures, it is possible that implicit measures will predict behavioral intention or actual behaviors. In the same meta-analysis, Greenwald reported that for socially sensitive issues, even when implicit and explicit measures were not strongly correlated, implicit measures were a better predictor of actual behavior than explicit measures (Greenwald, et al., 2009). However, implicit measures in this study failed to predict behavioral intention at all, while explicit measures accounted for a large percentage of variance in the model.

These two discrepancies raise serious questions about the validity of the implicit measure used in this study. Systematic inquiry into how and why implicit and explicit measures of SCM failed to correlate is needed. If SCM does represent two universal aspects of stereotypes, measuring it implicitly (or understanding why it cannot be measured implicitly) will greatly add to our current understanding of bias, both implicit and otherwise.

Abel, M. H., & Meltzer, A. L. (2007). Student ratings of a male and female professors' lecture on sex discrimination in the workforce. *Sex Roles, 57*, 173-180. doi: DOI 10.1007/s11199-007-9245-x

Baron-Cohen, S. (2003). *The Essential Difference: Men, Women, and the Extreme Male Brain*. London: Allen Lane.

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: implications for explaining autism. *Science, 310*(4), 819-823.

Basow, S. A. (1995). Student evaluations of college professors: When gender matters. *Journal of Educational Psychology, 87*(4), 656-665.

Basow, S. A., & Silberg, N. T. (1987). Student evaluations of college professors: Are female and male professors rated differently? *Journal of Educational Psychology, 79*(3), 308-314.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: evidence relating to the question of gender bias in teaching evaluation. *Journal of Educational Psychology, 74*(2), 170-179.

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. (2011). Understanding current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences, 108*(8), 3157-3162. doi: 10.1073/pnas.1014871108

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: Sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218-261.

Centra, J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *The Journal of Higher Education, 71*(1), 17-33.

Correll, J. (2008). 1/f noise and effort on implicit measures of bias. *Journal of Personality and Social Psychology, 94*(1), 48-59.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When professionals become mothers, warmth doesn't cut the ice. *Journal of Social Issues, 60*(4), 701-718.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS Map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92*(4), 631-648.

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in Experimental Social Psychology, 40*, 61-149.

Diekman, A. B., Brown, E. R., Johnston, A. M., & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*. doi: 10.1177/0956797610377342

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review, 109*(3), 573-598.

Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles, 47*(3/4), 99-114.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82*(6), 878-902.

Foundation, N. S. (2009). *Women, minorities and persons with disabilities in science and egneineering: 2009*. (NSF 09-305). Arlington, VA; January 2009 Retrieved from http://www.nsf.gove/statistics/wmpd/.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology, 45*(2), 410-414.

Good, C., Aronson, J., & Harder, J. A. (2008). Problems in the pipeline: Stereotype threat and women's achievement in high-level math courses. *Journal of Applied Developmental Psychology, 29*, 17-28.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit assocations task. *Journal of Personality and Social Psychology, 74*(6), 1464-1480.

Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Attitudes and social cognition, 85*(2), 197-216.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implcit Association Test: III. Meta-analysis of predictive validity. *Journal of Personalitty and Social Psychology, 97*(1), 17-41.

Groom, C. J., Sherman, J. W., Lu, L., Conrey, F. R., & Keijzer, B. S. C. (2005). Judging compound social categories: Compound familiarty and compatibility as determinants of processing mode. *Social Cognition, 23*(4), 291-323.

Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, math. *Science, 320*(1164), 1165.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *8, 1*(1-51).

Hedges, L. V., & Nowell, A. (1995). Sex differences in mental test scores, variability, and numbers of high-scoring individuals. *Science, 269*(5220), 41-45.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6581-592).

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139-155.

Karpinski, A., & Steinman, R. B. (2006). The Single Category Implicit Association Test as a measure of implicit social cognition. *Journal of Personalitty and Social Psychology, 91*(1), 16-32.

Kierstead, D., D'Agnostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology, 80*(3), 342-344.

Kimura, D. (2007). "Underrepresentation" or misrepresentation. In S. J. Ceci & W. M. Williams (Eds.), *Why aren't more women in science?* Washington, D.D: American Psychological Association.

Kirk, Roger E. (1995) *Experimental Design: Procedures for the Behavioral Sciences, Third Edition*. Monterey, California: Brooks/Cole Publishing.

Levy, J., & Kimura, D. (2009). Women, me, and the sciences. In C. H. Sommers (Ed.), *The Science on Women and Science*. Washington, D. C.: American Enterprise Institute.

Mason, M. A., & Goulden, M. (2004). Marriage and baby blues: Redefining gender equity in the academy. *Annals of the American Academy of Political and Social Science, 596*, 86-103.

Miller, J., & Chamberlin, M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology, 28*(4), 283-298.

Nevid, J. S., & McClelland, N. (2010). Measurement of implicit and explicit attitudes toward Barack Obama. *Psychology & Marketing, 27*(10), 989-1000. doi: 10.1002/mar.20369

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice, 6*(1), 101-115. doi: 1089-2699.6.1.101

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Thierry Devos, Ayala, A., et al. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement *Proceedings of the National*

Academny of Science of the United States of America, 106(26), 10593-10597. doi: 10.1073/pnas.0809921106

Oldmeadow, J., & Fiske, S. T. (2010). System-justifying ideologies moderate status=comptetence stereotypes: Roles for belief in a just world and social dominance orientation. *European Journal of Social Psychology, 37*, 1135-1148.

Parks-Stamm, E. J., Heilman, M. E., & Hearns, K. A. (2008). Motivated to penalize: Women's strategic rejection of successful women. *Personality and Social Psychology Bulletin, 34*(2), 237-247.

Sommers, C. H. (Ed.). (2009). *The Science on Women and Science*. Washington, D. C.: American Enterprise Institute.

Sprauge, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles, 53*(11/12), 779-793.

Summers, L. H. (2005). Remarks at NBER Conference on Diversifying the Science & Engineering Workforce Retrieved July 9, 2010, from http://www.president.harvard.edu/speeches/summers_2005/nber.php

Tierney, J. (2010). Daring to discuss women in science. *New York Times*, D1. Retrieved from http://www.nytimes.com/2010/06/08/science/08tier.html?pagewanted=2&_r=1&ref=science

Wade, M. L., & Brewer, M. B. (2006). The structure of female subgroups: An exploration of ambivalent stereotypes. *Sex Roles, 54*(753-765).

West, M. S., & Curtis, J. W. (2006). AAUP faculty gender equity indicators 2006. Washington, D.C.: American Association of University Professors.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592-598.