Spring 2019

Gender Bias in Self Perceived Confidence and Competence

Danielle Bibeault
danielle.bibeault@trincoll.edu

GENDER BIAS IN SELF PERCEIVED CONFIDENCE AND COMPETENCE

By

Danielle Bibeault

A Thesis Submitted to the Department of Economics
of Trinity College in Partial Fulfillment of the
Requirements for the Bachelor of Science Degree

Economics 498-99
April 17th, 2019

Abstract

Numerous studies point to a significant gap in confidence between men and women (Beckmann & Menkhoff, 2008; Beyer, 1990; Bleidorn et al., 2016; Carlin, Gelb, Belinne, & Ramchand, 2018; Erkut, 1983; Fleisher, Schoder, & Bayer, 2017; Fox & Firebaugh, 1992; Hirschfeld, Moore, & Brown, 1995; Jakobsson, Levin, & Kotsadam, 2013; Kamas & Preston, 2012; Kay & Shipman, 2014; Lundeberg, Fox, & Punccohar, 1994; Sarsons & Xu, 2015). The goal of this study is first to determine whether a gender gap in confidence exists, and then to examine whether Social Norms Marketing can be used to increase the confidence level of women. In order to do this, two experiments are conducted, both utilizing a trivia contest in which participants indicate their confidence level on a scale from 1 (least confident) to 10 (most confident) for each trivia response. If the answer is correct, the confidence level is added to their overall score. If the answer is incorrect, the confidence level is deducted from their overall score.

The first experiment is a field experiment taking place in West Hartford Center, CT and Buckland Hills Mall in Manchester, CT. In the control treatment, a five-question trivia game is administered. In the experimental treatment, in addition to a five-question trivia game, a normative message was verbally delivered to the participants stating that women performed better in the first trivia contest, which was the control treatment.

Overall, we find no difference in performance or confidence levels between men and women in the control treatment and subsequently in the experimental treatment. However, the confidence level of both genders increased significantly after the SNM intervention. Both genders also performed significantly better on the trivia contest after the SNM intervention.

The second experiment conducted was a laboratory experiment in The Theory of Games and Experimental Game Theory class (COLL 210) at Trinity College. In the control treatment, a

ten-question trivia game is administered. In the first experimental treatment, in addition to a ten-question trivia game, a normative message was delivered to the participants stating that women performed better in the first trivia contest (control treatment). This experiment also included a second experimental treatment, where a normative message stating that men performed better in the second trivia contest (first experimental treatment) was delivered to participants in addition to a ten-question trivia game.

Even though confidence levels for both men and women increased (7.1 and 9.8 points respectively) after the first normative message, stating that women scored higher on the previous trivia contest, the results are not statistically significant. We interpret these results with caution and attribute the findings to the small sample size, since only 5 females and 27 males participated. We find no difference in performance between men and women in the laboratory experiment. The results changed drastically after the second normative message, which stated that men scored higher on the previous trivia contest, was delivered. Even though confidence levels for both men and women went down between the first and second experimental treatments, men were more confident and performed better than women in the second experimental treatment.

Altogether, we find Social Norms Marketing an effective tool in affecting confidence levels of both genders. Women responded positively in terms of confidence (though the same way as men) to the message stating that women scored higher than men on the previous trivia contest and negatively in terms of both confidence and performance to the message that men scored higher on the previous trivia contest. Therefore, we conclude that SNM can be functional in boosting performance and confidence level of women at least in select instances.

Acknowledgments

I would like to express sincere gratitude to my advisor, Professor Arthur Schneider, for all the time and effort he has invested in me over the past four years. You are an incredible professor, advisor, and mentor, and I cannot thank you enough for your support, guidance, and patience throughout this process as well as my time at Trinity. Thank you for continuing to inspire my interest in Economics and for encouraging me to challenge myself academically.

I would also like to thank all of the professors in the Economics department for their advice and instruction throughout my time at Trinity. Each of you has a passion for sharing knowledge and helping others has been immensely valuable, and I will never forget it. I sincerely appreciate the classes you have provided me with that enabled me to write this thesis. I would especially like to thank Professor Ahmed, for her patience, understanding, and willingness to help. Your guidance has been immensely helpful to me as well as the other thesis students. Thank you sincerely for your feedback and advice in the seminar throughout the past two semesters.

Finally, I would like to thank my family and friends for their support, both in this thesis and throughout my undergraduate experience. To my mom, thank you for all of the opportunities you have given me. I'd also particularly like to thank my roommates for all of their support throughout this thesis and help with the data collection. Without you, this would not have been at all possible.

Table of Contents

Introduction

Many studies have determined that a confidence gap exists between males and females at all age groups, with men often having more self-confidence and a better perception of their own abilities than women (Beckmann & Menkhoff, 2008; Beyer, 1990; Bleidorn et al., 2016; Lundeberg et al., 1994). Past confidence experiments have included market simulations (Eckel & Füllbrunn, 2013; Powell & Ansic, 1997), surveys of working and academic professionals (Beckmann & Menkhoff, 2008; Sarsons & Xu, 2015), questionnaires before and after test-taking (Jakobsson et al., 2013), and risk behavior studies (Pawlowski, Atwal, & Dunbar, 2008).

A confidence gap between men and women has overarching implications, both for women as individuals and the world economy as a whole. Gender differences in confidence may mean women are less willing to enter higher-paying, more competitive fields, compete for promotions, or take risks that may pay off (Beckmann & Menkhoff, 2008; Carlin et al., 2018; Kamas & Preston, 2012; Kay & Shipman, 2014). Furthermore, the confidence gap costs firms as they cannot access the complete pool of skilled talent available when women do not participate fully in the labor force (Carlin et al., 2018; "Gender Forward Pioneer Index: World's Most Reputable Companies Have More Women in Senior Management", 2016; Horowitz, Igielnik, & Parker, 2018; Noland, Moran, & Kotschwar, 2016; Woetzel et al., 2015). Additionally, this confidence gap may contribute partially to a wage gap between men and women (Carlin et al., 2018; Kay & Shipman, 2014; Noland et al., 2016; Woetzel et al., 2015).

Past experiments have attempted increase women's confidence levels using blindness to gender differences (Martin & Phillips, 2017), college and online courses (Carlin et al., 2018), and corporate training programs (Nelson, Porth, Valikai, & McGee, 2015). We did not find any experiments using social norms marketing, a strategy that utilizes normative interventions to

correct individuals' misperceptions about their peers' behaviors. Social norms marketing argues that people often form their actions and decisions around these perceptions, and changing an individual's perceived social norms can alter their behaviors (Berkowitz, 2004; Cialdini & Trost, 1998; Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005). In the past, this strategy has been used to increase positive behaviors like recycling, reduce electricity consumption, and decrease binge drinking on college campuses (Clapp, Lange, Russell, Shillington, & Voas, 2003; Mollen, Rimal, Ruiter, & Kok, 2013).

Literature Review

I. Confidence Gap

Compared to men with similar abilities, women tend to evaluate themselves lower and degrade their own potential by not taking credit for their personal successes and performance (Beyer, 1990). This confidence gap between men and women heightens in unfamiliar situations, where women's confidence suffers the most (Sarsons & Xu, 2015). These confidence gaps are especially prominent in stereotypically 'male' fields, such as mathematics, science, and economics (Beyer, 1990; Jakobsson et al., 2013; Lundeberg, Fox, & Puncchar, 1994; Pope, 2017).

For example, a study of SAT scores and an associated questionnaire by Pope in 2017 found that women are less confident on the mathematics portion of the exam than men with the same score, but equally confident on the verbal portion. Despite this, other studies have also found confidence gaps to exist within the social sciences and humanities fields (Kamas & Preston, 2012). Across the United States, women only make up about 30 percent of all undergraduate Economics majors (Fleisher et al., 2017). Similar gaps appear in science and mathematics degree programs and are often attributed to lower confidence levels in women that obstruct them from selecting these degrees, which are typically known as more challenging. (Fox & Firebaugh, 1992).

In a 2013 experiment comparing gender confidence differences between young school children in Sweden and El Salvador, Jakobsson et al. found that confidence gaps between boys and girls in mathematics begin during early childhood. Sarsons & Xu (2015) found that this gap extends through adulthood and up to the PhD level. In an internet survey of nearly 1 million individuals across 48 nations, women were found to almost always have lower self-esteem than

men of the same age (Bleidorn et al., 2016). The confidence gap between men and women likely exists across almost all cultures, age groups, and subjects.

II. Implications and Costs of a Confidence Gap

Lower confidence levels may make women less willing to take risks or enter competitions, such as applying for jobs they are not completely qualified for, asking for promotions, choosing a difficult major, entering a competitive industry, or speaking up in a meeting (Beckmann & Menkhoff, 2008; Carlin et al., 2018; Eckel & Fullbrunn, 2013; Fox & Firebaugh, 1992; Gneezy, Niederle, & Rustichini, 2003; Horowitz et al., 2018; Kamas & Preston, 2012; Kay & Shipman, 2014; Kray & Kennedy, 2017; Nelson et al., 2015; Palomino & Peyrache, 2010; Pawlowski et al., 2008; Powell & Ansic, 1997). The confidence gap potentially leads women to underestimate their expected performance and competence, while men tend to overestimate these attributes (Beyer, 1990; Erkut, 1983; Gneezy et al., 2003).

Powell and Ansic (1997) performed two computerized financial decision-making laboratory experiments on undergraduate students to assess risk-seeking differences between men and women. They found that regardless of familiarity with a theoretical situation and costs associated, females were less willing to take risks than men, but both genders were able to achieve similar performance outcomes (Powell & Ansic, 1997). Comparable outcomes have also been observed in field studies: among mutual fund managers, females make fewer trades compared to males and tend to avoid competition and risk (Beckmann & Menkhoff, 2008; Niessen-Ruenzi & Ruenzi, 2006). Although they achieve similar results, lower willingness to compete may cost women in the long run (Beckmann & Menkhoff, 2008; Niessen-Ruenzi & Ruenzi, 2006). Expected rank compared to others has been found to be the most significant factor in whether or not an individual decides to enter a competition (Kamas & Preston, 2012).

Because of their lower confidence levels, women may be less willing to participate in competitive environments, even if they have the same capabilities as men, costing them opportunities for growth and development that could benefit the economy as a whole (Eckel & Fullbrunn, 2013; Gneezy et al., 2003; Kamas & Preston, 2012; Kay & Shipman, 2014; Pawlowski et al., 2008).

This lower self-confidence and lower expected performance can also worsen actual performance. For example, on the Economics GRE Subject Test, there exists a 40-point gap between the scores of men and women that can be explained by women's lower confidence levels in a stressful environment, but not by gender differences in academic or economics ability (Hirschfeld et al., 1995). Low confidence in women costs individuals significantly throughout their lifetimes, but it also costs firms, who are not able to access the largest possible talent pool and miss out on many females' diverse thoughts and skill sets (Carlin et al., 2018; Fleisher et al., 2017).

Firms with females in leadership positions are more profitable and have higher stock values than firms without females in executive positions ("Gender Forward Pioneer Index: World's Most Reputable Companies Have More Women in Senior Management", 2016; Noland et al., 2016). Despite this, females only account for 1.9 percent of senior executives even though they make up half of the workforce ("Gender Forward Pioneer Index: World's Most Reputable Companies Have More Women in Senior Management", 2016). Some attribute this absence to the confidence gap between men and women that makes women less likely to compete for a promotion or enter more competitive, higher-paying industries (Pawlowski et al., 2008; Kamas & Preston, 2012). Palomino and Peyrache (2010) found that differences in confidence levels between men and women may also contribute to the gender wage gap, as women self-select into less competitive industries, which tend to pay less. Lower confidence is one of several factors

hindering women from fully participating in global labor markets, but labor equality between men and women could add up to $12 trillion to the global economy if every country individually improved gender equality to match the best country in its region (Woetzel et al., 2015).

In addition, women's lower affinity for risk-taking and different approach compared to men could establish more stable global markets if they were to participate more equally in labor markets (Eckel & Fullbrunn, 2013; Syed, 2008). Syed (2008) argues that more women in highly competitive, typically masculine trading roles could have prevented the economic crisis of 2008. In experimental asset markets, women take fewer risks than men and predict lower prices, leading to significantly fewer bubbles and lower prices in all-female markets compared to all-male markets (Eckel & Füllbrunn, 2013).

III. Past Strategies to Improve Confidence of Women

Several experiments have been performed to test strategies to reduce or reverse the confidence gap. Carlin et al. (2018) created a college course aimed at increasing confidence among underconfident male and female undergraduate students, which was found to improve overall confidence indicators for both genders, suggesting that it may transfer effectively to workplaces and other environments. Firms could offer programs to teach their employees life skills like negotiation and leadership to improve women's confidence (Carlin et al., 2018; Nelson et al., 2015; Woetzel et al., 2015). In McKinsey's 2015 *Power of Parity* report, Woetzel et al. (2015) recommend leadership, confidence building, and negotiations training like after-school programs that teach skills to create economic opportunities and build capabilities for women of all ages. Kay and Shipman (2014) consider that male managers may not be aware of how many women feel and might need additional training to provide accurate and complete feedback without being overly critical. They also highlight the importance of encouraging girls, especially

at the high-school age, to compete and be willing to move past their failures (Kay & Shipman, 2014). When supermarket chain Asda discovered that women comprised 70 percent of their hourly salesforce but only 30 percent of employees suitable for promotion, they surveyed to uncover what specifically was holding women back then established more positive female role models in leadership positions (Nelson et al., 2015; Woetzel et al., 2015). Similar strategies have been implemented by other firms and schools in order to improve outcomes for women.

Blindness to Gender Differences

Martin and Phillips (2017) proposed that minimizing and downplaying natural gender differences may be effective in increasing women's confidence, competitiveness, and willingness to negotiate. In order to compare gender blindness and gender awareness strategies, they first surveyed how women at Amazon's MTurk felt gender differences affected their abilities to lead effectively and hold influence in the workplace (Martin & Phillips, 2017). Martin and Phillips (2017) then compared individuals they identified in surveys as gender-blind to a control group and found an association between gender-blindness and action-taking.

They ultimately discovered that gender-blindness would lead to more agency, confidence, and higher self-perception among women (Martin & Phillips, 2017). Martin and Phillips (2017) found that women are more confident when they're blind to gender differences between themselves and men than when they are made aware of these gender differences. This effect is especially apparent in male-dominated environments, so gender blindness could be most helpful to women who work in these contexts (Martin & Phillips, 2017).

Workplace Education at Gap Inc.'s Factories

Clothing company Gap created a workplace educational program for female workers in its garment factories throughout the world (Nelson et al., 2015). The program involves class-based

training on practical skills like time management and financial literacy as well as communication tools, the influence of gender norms, and effective decision-making (Nelson et al., 2015). These classes centered around improving how the women felt about themselves and their abilities (Nelson et al., 2015).

Six years after Gap's initiative began, the International Center for Research on Women evaluated the program and reported a 50 percent increase in self-esteem compared to before (Nelson et al., 2016). The classes had clear benefits for the empowered women but also served factory owners: with increased confidence, the women were more productive and efficient, and more likely to stay at their factory (Nelson et al., 2015). In Cambodia, the participating women were 66 percent more likely to stay employed with that specific factory while in India, participants were 58 percent more likely to be promoted (Nelson et al., 2015).

College Class for Confidence

In order to test their recommendations that managers concern themselves with women's low self-confidence and work to improve it, Carlin et al. (2018) created a classroom learning opportunity for underconfident male and female undergraduate business majors. The "Internships for Introverts" course aimed to increase self-confidence and improve communication skills (Carlin et al., 2018).The activities to improve self-esteem included mock interviews, public speaking, and networking followed by accurate feedback (Carlin et al., 2018).

At the end of Carlin et al.'s 2018 experiment, the participants demonstrated significantly increased confidence compared to a control class for all business majors in general. Participants volunteered more, joined more student organizations, and applied for more internships (Carlin et al., 2018). Although this was limited to a classroom setting, Carlin et al. (2018) believe this could be applied to firms using meetings or videotaped presentations targeted at women or all underconfident individuals. The authors also suggested that firms should consider training all of

their employees to review how they give feedback to women compared to men, as different external perceptions can further impact self-confidence levels (Carlin et al., 2018).

IV. Social Norms

Social norms are beliefs, rules, and standards that members of a group or a culture share (Cialdini & Trost, 1998; Lapinsky & Rimal, 2005). Individuals use the social norms they perceive to decide their own behaviors and beliefs, culminating in how they should act in situations, especially unfamiliar ones (Cialdini & Trost, 1989; Cialdini & Goldstein, 2004).

Social norms can be divided into three categories based on what sort of information they provide individuals with. Descriptive social norms use actions and examples to inform individuals about what is typically done by others in a particular social situation (Cialdini & Trost, 1989, Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005; Rimal & Lapinsky, 2015). In comparison, injunctive social norms are simpler, and inform individuals about whether an action is approved or disapproved by the general community they associate with (Cialdini & Trost, 1989, Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005; Rimal & Lapinsky, 2015). Third, subjective social norms tell individuals what people who matter most to them, like friends and family, would specifically think about their actions (Cialdini & Trost, 1989, Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005; Rimal & Lapinsky, 2015).

Occasionally, different social norms will contradict each other, potentially creating an internal decision-making conflict (Cialdini & Trost, 1989, Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005; Rimal & Lapinsky, 2015). For example, a college student might see their peers binge drinking and receive the descriptive norm that they should drink in order to fit in with the community, but at the same time they have been made aware of an injunctive norm that tells them binge drinking is disapproved by the wider community. This student might also consider

the subjective norm of what their parents or other family members would think of their behavior. Cialdini & Trost (1998) argue that in this situation, individualistic value orientations come into consideration, where people will move past the social norm and focus on their own personal beliefs and priorities.

Social norms and beliefs are not official rules or legal norms, instead gaining power from the culture they come from or their importance to efficiency and survival (Cialdini & Trost, 1998; Elster, 1989). People follow social norms in order to avoid social consequences like ostracization and disapproval, but that does not mean individuals are always rational in deciding to adhere to a norm (Elster, 1989). Often, individuals who follow social norms are not entirely outcome-focused and follow norms instead because they "should" or always have (Elster, 1989). In following social norms, people are not typically acting out of self-interest, as they gain very little by following a norm (Elster, 1989). However, individuals will occasionally use social norms to their advantage and act in a rational, self-interested way by capitalizing on social expectations, such as doing a favor for a friend with the expectation of receiving one back (Elster, 1989). Most often though, people follow social norms not out of rational self-interest, but because they believe violating them will result in social sanctions or consequences (Cialdini & Trost, 1998; Elster, 1989).

Typically, individuals follow social norms in efforts to connect and identify as part of a larger group (Lapinsky & Rimal, 2005). People who use social norms to affiliate themselves with a group will interpret and copy the group's social norms and behaviors (Cialdini & Goldstein, 2004). This may allow someone to feel closer to the group and believe the group will see them in a more positive light (Cialdini & Goldstein, 2004). Here, social norms are followed because they are shared collectively with a group from which an individual wants approval (Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005; Lapinsky & Rimal, 2015). Mostly, people will follow

social norms due to self-concept, or a desire to see themselves in a positive light (Cialdini & Trost, 1998; Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005). In order to feel like a "good" person and maintain a positive self-concept, individuals will choose to follow what they perceive as social norms in conjunction with their own value ideals (Cialdini & Trost, 1998; Cialdini & Goldstein, 2004; Lapinsky & Rimal, 2005). Because people want to be liked and respected by the people they admire, they will continue to follow social norms (Elster, 1989). Whether injunctive, descriptive, or subjective, a social norm only has power because people see enough benefits from following it or enough drawbacks to violating the social norm (Elster, 1989).

While norms demonstrated by the media and norms that affect society as a whole are collective, an individual's own interpretation of them is a perceived norm (Lapinsky & Rimal, 2005). When individuals interpret social norms from the collective level to the perceived level, they often misunderstand and the intention behind the norm can change (Lapinsky & Rimal, 2005). Often, perceptions of peers' beliefs and behaviors are actually incredibly inaccurate, especially as social distance increases: individuals may have some ability to estimate family and close friends' beliefs and behaviors but are often hugely mistaken about the beliefs and behaviors of larger groups with more social distance, like an entire town or student body (Berkowitz, 2004). Pluralistic ignorance, where individuals incorrectly perceive the attitudes and behaviors of their peers and greater community, can cause beliefs and perceived social norms, and therefore actions to be misguided (Berkowitz, 2004).

V. Social Norms Approach

Social norms marketing interventions use normative messaging to correct individuals' misperceptions about their peers and communities that establish their behaviors. These interventions have been used in efforts to improve individual behaviors, such as recycling,

electricity consumption, and eating habits (Berkowitz, 2004; Cialdini & Trost, 1998; Cialdini & Goldstein, 2004; Clapp et al., 2003; Elster, 1989; Harries, Rettie, Studley, Burchell, & Chambers, 2013; Mollen et al., 2013; Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007; Turner, Perkins, & Bauerle, 2008; Wechsler, Nelson, Lee, Seibring, Lewis, & Keeling, 2003). Social norms marketing experiments assume that individuals overestimate their peers' negative behaviors and underestimate positive behaviors, and then follow those normative behaviors in their own actions (Schultz et al., 2007). By correcting misperceptions about their peers, the theory is that the individuals will change their own behaviors in response to their new understanding of social norms (Berkowitz, 2004).

In order to perform these experiments, researchers use a normative messaging intervention on experimental treatment groups and compare the results to a control group to judge the effectiveness of a social norms marketing intervention. Sometimes, multiple normative messages, such as a descriptive and an injunctive norm, are applied simultaneously (Harries et al., 2013; Meeker et al., 2016). Researchers often distribute posters, flyers, and other paraphernalia expressing facts, approval, and peer comparisons to influence subjects under the social norms marketing treatment (Clapp et al., 2003; Harries et al., 2013; Kilmartin et al., 2008; Mollen et al., 2013; Morewedge et al., 2015; Nolan, Schultz, Cialdini, Goldstein, & Griskevicius, 2008).

Many individuals who undergo a social norms marketing treatment fail to realize how significantly the intervention affected them (Nolan et al., 2008). In one study of energy conservation, respondents rated social influences as the "least motivating" factor compared to fact sharing and other messaging, when in reality it was the most significant factor in reducing their energy consumption (Nolan et al., 2008). This can grant social norms interventions extra power in that some subjects may not realize their effects (Nolan et al., 2008). However, it also

may mean researchers should be extra careful when crafting how they intervene to change individuals' perceptions (Schultz et al., 2007).

Some studies that use normative messaging to influence behaviors can actually worsen negative behaviors or reduce positive behaviors in certain individuals (Schultz et al., 2007). In this case, individuals who were already performing positive actions, like conserving energy, at an above-average rate may reduce their efforts in order to meet their peers' levels (Schultz et al., 2007). This "boomerang effect" means that researchers must take caution when they distribute messaging in order to have only a positive effect (Schultz et al., 2007). It also means that purely descriptive messaging may not be the most effective, therefore many researchers choose to add an injunctive message to further signify general approval or disapproval (Schultz et al., 2007).

Successful Experiments

Reducing Male Sexism

In a 2008 experiment, Kilmartin et al. tested the attitudes of 65 undergraduate males using several sexism scales. Then, they divided the participants evenly into intervention and control groups and had each participant compare their own attitudes about sexism to what they perceived to be the attitudes of their peers in the room (Kilmartin et al., 2008). The initial surveys of these men showed that they overestimated the sexism of others in the room compared to their own indicated sexism levels (Kilmartin et al., 2008).

After a brief descriptive intervention was done, a two-way multivariate ANOVA between the control and intervention groups showed a significant difference in each group's perception of others in the second survey (Kilmartin et al., 2008). As researchers predicted, the brief

intervention changed how men perceived their peers: they realized their peers were less comfortable with sexism and less sexist than they had initially thought (Kilmartin et al., 2008).

This demonstrates that short interventions may be helpful in changing how individuals perceive the opinions of their peers (Kilmartin et al., 2008). Expanding the experiment, Kilmartin et al. (2008) tested how the men were able to perceive the opinions of peers who they were close friends with. They found that the undergraduate men were no more accurate when predicting their friends' attitudes compared to strangers, providing a potential counterpoint to theories that say social norms perceptions worsen with social distance, such as described by Berkowitz in 2004 (Kilmartin et al., 2008).

Alcohol Misuse Consequences

In a 2008 study, Turner, Perkins, and Bauerle described the three years they spent between 1999 and 2002 reaching out initially to first-year undergraduate students and, in the final year, to the entire student body on a college campus to intervene in the misuse of alcohol. In doing so, they chose to focus on the consequences that come from misusing alcohol and binge drinking, such as driving under the influence of alcohol, getting injured, and making decisions one otherwise would not (Turner et al., 2008). They created posters, web postings, presentations, and student newspaper advertisements that focused on healthy normative behaviors students had indicated in an earlier survey, like using the buddy system and being careful of how much alcohol they consumed (Turner et al., 2008). The posters also corrected students' perceived estimations about how much and how often their peers were drinking heavily (Turner et al., 2008).

Throughout their three-year experiment, the odds of an undergraduate student experiencing none of the negative consequences of alcohol misuse that they surveyed for increased by 113% (Turner et al., 2008). Specifically, 1,972 fewer students were injured in alcohol-related events and 1,511 fewer students drove under the influence of alcohol (Turner et al., 2008). This shows that a social norms marketing intervention may be useful for reducing alcohol consumption on a college campus (Turner et al., 2008). However, the 2008 experiment by Turner et al. was unique in that it focused on the consequences that result from binge drinking and misusing alcohol, instead of simply on correcting outside perceptions of alcohol. This could have a different effect when compared to a more traditional social norms intervention, focused directly on binge drinking.

Decision-Making

This study was done to see how a short, one-time social norms intervention can alter decisions that have been influenced by untrue biases in a theoretical example (Morewedge et al., 2015). Experimental groups were each subjected to a short, one-time training intervention: one group watched a 30-minute video and another group played a computer game (Morewedge et al., 2015). Morewedge et al. (2015) found that people's biased decisions are not based on facts, but mistaken beliefs. Therefore, individuals are using what they perceive and believe to be "normal" to make decisions about actions that affect their neighbors, peers, and society as a whole (Morewedge et al., 2015). When these participants were subject to the brief interventions to correct biases, their beliefs and perceptions changed and so did their decision-making (Morewedge et al., 2015).

Antibiotic Prescriptions

Doctors' over-prescription of antibiotics for unnecessary causes has been thoroughly documented and the negative effects are well-known, but little has been done to reduce unnecessary antibiotic prescriptions (Meeker et al., 2016). In 2016, Meeker et al. decided to assess if a social norms intervention could work to reduce inappropriate antibiotic prescriptions among participating doctors. Each doctor was subjected to no intervention, a single intervention, or a combination of interventions (Meeker et al., 2016).

The interventions included suggesting non-antibiotic alternatives in electronic order sets, requiring entry of free-text justifications for prescribing antibiotics, and, most notably for social norms marketing, peer comparison emails where a doctor could see how their unnecessary antibiotic prescribing rates compared to 'top performing' doctors with the lowest inappropriate antibiotic prescribing rates (Meeker et al., 2016). Of the three interventions, the free-text justification and peer comparison emails resulted in groups with the lowest rates of inappropriate antibiotic prescriptions compared to their peers (Meeker et al., 2016).

Unsuccessful Experiments

A Comparison of Colleges With and Without SNM Programs for Reducing Alcohol Use

In a 2003 social norms marketing survey, Wechsler et al. used survey data from students at 118 different colleges to determine how much alcohol they consumed as well as the existence and extent of any social norms marketing program the schools had in place. They then compared the 57 schools that reported using social norms marketing to the 61 that said they did not use social norms marketing to determine any changes in drinking behavior over time (Wechsler et

al., 2003). The researchers noticed that many schools reported that they did not have an official program but were using similar marketing tactics in efforts to reduce excessive alcohol consumption (Wechsler et al., 2003). Among the 118 schools they studied, schools that had an official social norms marketing program were likely to be larger and less likely to be religious than schools that lacked a program (Wechsler et al., 2003).

Ultimately, their trend analyses showed no reduction at the social norms marketing schools in drinking behaviors, leading the researchers to conclude that the social norms intervention programs were ineffective (Wechsler et al., 2003). However, many of the schools that used social norms marketing had higher alcohol consumption rates at the initial survey, indicating that social norms marketing may be a common strategy that schools with drinking problems turn to in efforts to combat it (Wechsler et al., 2003). Notably, the study did not enact its own experimental program and instead created a high-level comparison between schools with a social norms marketing program and schools without one (Wechsler et al., 2003). Therefore, some inadequacies and failures in individual programs may have been overlooked – while the social norms marketing programs were ineffective overall, this may not mean that social norms marketing is ineffective when used carefully in appropriate situations (Wechsler et al., 2003). Additionally, some individual schools utilizing the programs did see a decrease in excessive drinking, although a change was not seen overall (Wechsler et al., 2003).

Domestic Electricity Consumption in the U.K.

In a 2013 study in the U.K., Harries et al. studied if social norms marketing can reduce household electricity consumption. They separated 367 participating households into a control group (N=121), an individual group (N=124), and a social norms group (N = 122) (Harries et al.,

2013). The control group received no messaging on their electricity consumption, while the individual group received messaging about only their own electricity consumption, and the social norms group received messaging that compared them to other households nearby (Harries et al., 2013). The households received the messaging through 18 emails that were sent throughout the study (Harries et al., 2013).

While energy consumption in both the individual and social norms intervention groups declined by 3 percent relative to the control group, there was not a statistically significant difference between the three groups (Harries et al., 2013). Although the interventions were statistically insignificant, it was worth noting that the participants in the social norms group downloaded their emails 19.8 times compared to only 13.4 downloads in the individual intervention group, meaning the social norms condition may have had some impact on how individual households thought about their energy consumption (Harries et al., 2013). The experiment's lack of success may also be attributed to its basis in the U.K., where electricity consumption is significantly lower at 13 kWh on average per household per day compared to 31 kWh daily per household in the United States, where more social norms marketing experiments to reduce electricity consumption have had success (Harries et al., 2013).

Food Selection

In a 2013 study of 697 students at an on-campus university food court, Mollen, Rimar, Ruiter, and Kok compared the effectiveness of healthy and unhealthy descriptive norms and healthy injunctive norms for decision-making. To share healthy and unhealthy descriptive norms, they described the popularity of salads and hamburgers respectively using noticeable signs in the

dining hall (Mollen et al., 2013). The same signs were used to express approval for making a salad choice to create a healthy injunctive norm (Mollen et al., 2013).

Their experiment ultimately had mixed results. They found that the healthy injunctive norm did not create more salad choices and the unhealthy descriptive norm did not significantly alter hamburger consumption (Mollen et al., 2013). However, the healthy descriptive norm did create more salad consumption when compared to both the control and the unhealthy descriptive norm condition (Mollen et al., 2013). Therefore, the extent of the impact and the effect social norms marketing had on the participants in this study cannot be determined (Mollen et al., 2013).

Alcohol Consumption at a Large University

In a 2003 experiment, Clapp et al. utilized two residence halls at a large university to establish a control and experimental treatment. In the experimental social norms hall, they distributed posters, signs, stickers, bookmarks, and notepads stating truthfully that 75 percent of students at the school drank 4 or fewer drinks when they partied (Clapp et al., 2003). Although comparing surveys of students in both halls showed that the experimental group perceptions of how much alcohol other students consume had changed, actual alcohol consumption stayed the same in both residence halls, meaning that the social norms marketing experiment had failed (Clapp et al., 2003).

Research Methodology

I. Field Experiment - Between the subject design

Trivia Game and Scoring

To test participants' confidence levels, they were asked to complete a brief survey that involved answering five trivia questions. An example of this survey can be found in Appendix A. The contest asks participants their gender and whether they expect to do better or worse than male participants, female participants, and average participants, and their confidence level for each expectation before they see the trivia questions. The decision to only allow participants to choose "better" or "worse" and exclude a "same" option was purposeful in order to establish participants' real bias.

After responding to the preliminary questions, participants write in responses for five trivia questions and select their confidence level on a scale from 1 (least confident) to 10 (most confident). At the end of the game, the participants say whether they now believe they performed better or worse than the average male participant, the average female participant, and the average person, and their confidence level in each comparison.

The use of five questions compared to ten or another number was decided using a pilot game consisting of seven students in an undergraduate Economics senior seminar. The pilot survey utilized ten questions and asked more demographic information than gender. Ultimately, it was determined to be too long to use in a field experiment game of this sort. Next, 25 trivia questions for a total of five different surveys (to prevent cheating) were chosen from a group of 100 potential trivia questions.

These 25 final questions were chosen by providing surveys with 10 questions each to 40 undergraduate students and asking them to answer each question, then rate the difficulty from 1 (easiest) to 10 (most difficult). If they answered incorrectly, their rating was changed to a 10. Using the average difficulty of each question, one question each was selected from the categories of sports, geography, and pop culture, and two questions were selected in the science category, to get to an average total difficulty for each survey between 4.5 and 5 out of 10.

The survey was distributed on paper with an instructions sheet, which provided an advising professor's information for any questions, explained that cheating was not allowed, and detailed how the trivia game was to be scored. For each of the five questions that participants answered correctly, their confidence level was added to their score. For each incorrect question, their confidence level was subtracted from their score. Therefore, scores could range between -50 and +50 points.

Participants and Procedure

For the control treatment survey, the trivia game described above was distributed three times in West Hartford Center, CT in October and November 2018 to random passerby and one time in November 2018 at the Buckland Hills Mall. All of the experiments took place on Saturdays or Sundays. 30 of the 139 control group participants were recruited at the Buckland Hills Mall. No difference was found in the confidence level or performance of participants between West Hartford Center and the Buckland Hills Mall, so the data was pooled to form the control group. The 139 control group participants were comprised of 71 females and 68 males.

Participants for both the control treatment and the social norms marketing treatment were recruited through tabling in West Hartford Center, CT and Buckland Hills Mall in Manchester, CT with a poster, where undergraduate students asked volunteers to participate in a trivia survey for a senior thesis. Participants were also offered Lindt chocolates and told the participant with

the highest score would win an Amazon gift card if they chose to share their email address or phone number. Otherwise, the survey was completely anonymous.

The participants were then categorized based on their scores as follows:

Confident: An above average score when they said they thought they had performed "Better than the Average Person" or a below average score when they said they thought they had performed "Worse than the Average Person".

Underconfident: A score above average when they said they thought they had performed "Worse than the Average Person".

Overconfident: A score below average when they said they thought they had performed "Better than the Average Person".

Participants were also divided based on whether they were "Competent" or "Incompetent" as defined below:

Competent: A score above -5 (-10 for laboratory experiment, because of 10 questions)

Incompetent: A score below -5 (-10 for laboratory experiment, because of 10 questions)

Social Norms Intervention

For the social norms intervention, participants were recruited at the Buckland Hills Mall in Manchester, CT on three different weekend days in February 2019. The social norms intervention experiment yielded 150 participants between the three days. No significant difference in scores or questions answered correctly was found on any of the three days. The participants were recruited the same way as in the control treatment through tabling, offered Lindt chocolates and the chance to win an Amazon gift card in exchange for completing a trivia

game for a senior thesis. The participants consisted of 79 females and 71 males. The survey distributed to them was the same as the survey in the control treatment.

In the social norms intervention treatment, participants were told that females had scored higher than males on the first experiment. No other context was given about past or expected performance besides the verbal messaging when the survey was distributed.

II. Laboratory Experiment - Within the subject design

Participants

In addition to the field experiment, a laboratory experiment with a within the subject design was performed in an undergraduate classroom of 32 students, utilizing the same students in all three treatments: a control treatment and two experimental SNM treatments. This course was titled "Theory of Games and Experiment" and focused on experimental economics and experimental games. In the course, students play economic games and may pay a fee in order to compete for monetary prizes. The students in the course are primarily Economics majors. Of the 32 students, five were female and 27 were male. The low number of women included in the experiment, as well as the potential competitive and confident nature of the students, must be considered when comparing this experiment to the field experiment.

Procedure

The trivia questions used in the classroom experiment were sourced the same way as in the field experiment, with each survey having a difficulty level between 4.5 and 5 out of 10. However, participants in the classroom experiment complete the trivia contest using an online survey. The classroom also utilized ten questions instead of five. The students' responses were scored using the same method as in the field experiment, where their confidence level was

subtracted from their score in the event of an incorrect answer and added to their score in the event of a correct answer. The preliminary questions, where students responded whether they expected to do “better” or “worse” than the average male, female, and individual, and their confidence level from 1 (least confident) to 10 (most confident) were also the same as in the field experiment. The trivia survey was taken during class time and students were given ten minutes to take the survey.

The participants first responded to a control survey that consisted of ten trivia questions. In a subsequent class period, the students were told that women had performed better than men on the first trivia game, and again given ten minutes to complete an online trivia game survey with ten new questions. Finally, a second intervention was performed in a third class period where students were told that men had performed better than women on the last survey.

Results

In analyzing the results, a participant's score is considered to be their 'confidence', and the number of questions a participant answered correctly is considered their 'performance' in the trivia game.

I. Field Experiment - Between the subject design

Descriptive Statistics

71 females and 68 males participated in the control treatment. Female scores averaged 8.563, which was not significantly different than the males' average of 8.544 (one-tail p=0.993). Altogether, all participants in the control treatment had an average score of 8.554, meaning that 39 females and 39 males were above the overall average.

98 of the 139 participants were "Confident", of which 48 were women and 50 were men. Additionally, 17 females and 5 males were "Underconfident", and 6 females and 13 males were "Overconfident".

Of all the females, 52 were "Competent" and only 19 were "Incompetent". For males, 49 were "Competent" and 19 were "Incompetent", leading to a total of 101 "Competent" participants and 38 "Incompetent" participants.

On average, participants answered 2.576 questions correct out of the 5 questions in the game. Females answered 2.51 questions correct on average, while males answered 2.65 questions correct on average. A table comprising of these statistics can be found in Appendix B.

In the social norms treatment, 79 participants were women and 71 were men, for a total of 150 participants. The average confidence in the social norms intervention experiment was 11.94, with the female average score of 12.24 insignificantly higher than the male average of 11.79.

84 individuals were “Confident”, a group that consisted of 44 women and 40 men. Additionally, 19 women and 5 men were “Underconfident” and 16 women and 26 men were “Overconfident”.

Of these participants, 63 men and 68 women were “Competent” while 8 men and 11 women were “Incompetent”.

On average, participants answered 2.84 questions correct, with women performing slightly, but not significantly, better than men and answering 2.899 questions correct compared to men’s 2.788 average. A summary of these statistics can be found in Appendix C.

Confidence

There was not a significant difference between the confidence levels of men and women on average in either the control or social norms treatments. However, in the control treatment Confident women were significantly less confident than Confident men ($p=0.041$) (Appendix I).

All individuals on average were significantly more confident in the SNM treatment than in the control treatment ($p=0.021$) (Appendix E). Specifically, females had a weakly significant increase in confidence level after the confidence intervention ($p=0.055$) (Appendix F). There was no significant difference in the confidence levels of men after the intervention (Appendix F). Therefore, the intervention increased confidence levels for the group as a whole and for women.

Confident, underconfident, and overconfident individuals were all significantly more confident in the SNM treatment compared to the control treatment (Appendix G and Appendix H). For confident individuals, their scores increased from an average of 11.99 to 15.45 ($p=0.023$). Overconfident individuals saw the largest confidence increase, with their scores going from -11.68 to 1.10 on average ($p<0.001$). Underconfident individuals’ average scores improved from 8.55 to 18.63 ($p=0.020$) (Appendix G and Appendix H).

In the control treatment, there was no significant difference in confidence between incompetent males and females, but incompetent females became significantly more confident than incompetent males after the SNM intervention (p=0.048) (Appendix M). This suggests that the intervention in the field experiment was effective for incompetent females.

Performance

Females performed significantly better after the social norms intervention, answering 0.585 more questions correctly on average (p=0.002) (Appendix N). Interestingly, males also performed significantly better after the intervention than in the control treatment. They answered 0.624 more questions correctly on average (p=0.005) (Appendix N). Males' similar performance response suggests they may have become more competitive after the intervention and improved their performance.

Confident, overconfident, and underconfident individuals all performed better after the SNM intervention. Overconfident participants answered 2.17 questions correctly on average after the intervention, compared to 1.42 questions before (p=0.0055). Confident individuals had a weakly significant improvement in performance as well (p=0.051). Underconfident individuals answered 3.38 questions correctly on average after the intervention and 2.59 questions correctly before (p=0.0055). (Appendix O and Appendix P).

Of the Confident individuals, females in particular performed significantly better after the social norms intervention (p=0.0225). (Appendix Q and Appendix R). Underconfident females, however, saw the biggest improvement in performance, answering 0.734 more questions correctly (p=0.011). This suggests that the brief SNM intervention is effective in improving performance for underconfident and confident females.

The social norms intervention was insignificant to the performance of Confident males, but improved performance significantly for both Overconfident and Underconfident males

(Appendix Q, Appendix R). Overconfident males answered 0.808 more questions correctly on average after the intervention (p=0.007), which may further support the possibility that the social norms intervention increased competitiveness in males. Underconfident males also had a weakly significant increase in performance after the intervention (p=0.049) (Appendix Q and Appendix R).

Incompetent individuals performed significantly worse after the social norms intervention (p=0.005). This suggests the intervention may not have the intended effect for Incompetent individuals, or those who scored below a -5 in the field experiment. However, competent individuals had no difference in performance after the social norms intervention (Appendix S and Appendix T).

Individuals who expected to perform "Better than the Average Male" performed significantly better following the social norms intervention (p=0.014) (Appendix U). In contrast, specifically males who expected to perform "Worse than the Average Male" performed significantly worse after the social norms intervention (p=0.038) (Appendix V and Appendix W).

While these males had a decrease in performance, females who expected to perform "Worse than the Average Male" saw an insignificant improvement in their performance between the control and social norms treatments(Appendix Y). This reverse trend suggests that the intervention may have worked as anticipated on this group of people who had low expectations compared to the average male.

II. Laboratory Experiment - Within the subject design

Descriptive Statistics

Among 32 participants in the laboratory experiment, 5 were female and 27 were male. The small number of females may have had an effect on the significance of trends between the

three treatments. In the control treatment, the average score was 25.28, with females having an average score of 13.20 compared to males' average score of 27.52. In this treatment, females answered an average of 4.60 questions out of 10 correct, while males averaged 5.59 questions out of 10 correct, for a group average of 5.44 questions correct out of 10.

In the second treatment, the social norms intervention in which participants were told that females had performed better than males, all of the participants had an average score of 32.84. Females' scores averaged 23.00, while males averaged 34.67. Together, the group of participants had an average of 5.56 questions correct out of 10, with females having an average of 4.40 questions correct and males performing at 5.78 questions correct.

In the third treatment, in which the group was told that males had outperformed females, the average score for all participants was 16.44. Females scored -4.20 on average, while males scored 20.26 points on average. The group got an average of 4.72 questions correct, consisting of the males' 5.07 average and females' 2.80 average. These statistics are summarized in a table in Appendix Z.

Confidence

Between the three treatments, females were the least confident in treatment 3. They were significantly less confident in treatment 3 than in treatment 2 ($p=0.001$) and treatment 1 ($p=0.027$), as seen in Appendix AA. Comparatively, men were also the least confident on average in treatment 3. As seen in Appendix AA, they were significantly more confident in both treatments 1 ($p=0.031$) and 2 ($p<0.001$) than in treatment 3.

Males were significantly more confident than females in treatment 3 ($p=0.001$), and weakly significantly more confident than females in treatment 1 ($p=0.08$) (Appendix AB). This may imply that the first social norms intervention, where the group was told that women performed better than men, improved female confidence levels, even if both groups improved. Overall, the

group's confidence was the highest in treatment 2 and the lowest in treatment 3, although there was no significant difference. This means that, on average, participants were the most confident when they were told women had performed better than men, suggesting the brief intervention may be effective for increasing confidence levels of women as well as men.

In particular, females were the most confident in treatment 2 compared to treatments 1 or 3. They were significantly more confident in treatment 2 than in treatment 3 ($p=0.001$) and insignificantly more confident in treatment 2 than in treatment 1 ($p=0.186$), as seen in Appendix AC. Overall, the intervention was effective in increasing the confidence levels of women in the classroom experiment.

Performance

Women performed significantly worse than men in treatment 3, when the group was told that men had performed better than women ($p<0.001$) (Appendix AD). In treatment 2, men performed weakly significantly better than women ($p=0.051$) (Appendix AD). The average performance was lowest in treatment 3, where it was significantly lower than in treatment 1 ($p=0.007$) or treatment 2 ($p=0.002$) (Appendix AE).

As a group, men performed worst in treatment 3, performing weakly significantly better in treatment 1 ($p=0.07$) and treatment 2 ($p=0.01$) in treatment 2 than in treatment 3, and insignificantly better in treatment 1 than in treatment 3 (Appendix AF). It is possible that, when told men performed better than women in the previous trivia contest, men no longer felt the need to compete and their performance suffered. Similarly, female performance was the worst in treatment 3. Female participants also performed significantly better in treatment 1 ($p=0.018$) and treatment 2 ($p=0.011$) than in treatment 3 (Appendix AF). Here, it is possible that women felt they could not perform better than men after hearing the intervention.

Conclusions and Limitations

I. Conclusions

Overall, females in the field experiment were more confident after the SNM treatment than they were in the control treatment to the 10% significance level (p=0.055) (Appendix F). Males also had a weakly significant increase in confidence after the SNM intervention (p=0.0995) (Appendix F). Confident, Overconfident, and Underconfident individuals were all significantly more confident on average in the SNM treatment than in the control treatment. There was no significant difference in confidence between males and females in either of the treatments. While the SNM intervention resulted in a weakly significant improvement in confidence levels of women, it also caused a weakly significant improvement in male confidence levels.

Participants in the field experiment on average performed significantly better in the SNM treatment than in the control treatment, a trend that applied across all classifications -- confident, overconfident, and underconfident. Amongst females, performance was significantly higher in the SNM treatment than in the control treatment as they answered 0.585 more questions correctly on average. Males also performed significantly better in the SNM treatment than in the control treatment, answering 0.624 more questions correctly on average after the social norms intervention (p=0.005) (Appendix N).

This data from the field experiment may indicate that the brief SNM treatment used in this experiment was effective in improving confidence levels of women across all classification groups. The social norms intervention also improved confidence levels of men, a trend that was shared with the laboratory experiment. After the social norms intervention where participants

were told that women performed better, overall confidence and performance increased for both men and women in the laboratory and field experiments.

The laboratory experiment validates the field experiment by demonstrating that the same individuals perform better on average and are more confident after the intervention than they did in the control treatment. This means that informing individuals that women performed better than men may increase confidence levels of women, but it could also increase confidence of men who now feel the need to be more competitive.

The field experiment did not include a third intervention as in the laboratory experiment, where participants were told that men had outperformed women. In the laboratory experiment, both males and females performed worse and were less confident after this intervention. It is possible that this message discouraged the female participants from behaving competitively, while males no longer felt the need to compete, lowering confidence and performance for all participants on average.

In both the control and SNM treatments, women tended to be more underconfident and less overconfident, while men tended to be more overconfident and less underconfident. Therefore, even after the intervention, women had a significant tendency to be underconfident ($p<0.05$) (Appendix AG). This suggests that while the intervention began to improve female confidence levels, further steps are needed to see more significant improvements. Additionally, a more specific intervention may be necessary in order to improve female confidence levels without simultaneously altering male confidence levels.

II. Limitations

When utilizing a survey experiment, it is impossible to capture the wide scope of characteristics that make up any one participant. An individual's education level, income, marital

status, or other demographic information could all impact how they perform during the trivia game and their confidence level. Without including further demographic information, this study is limited in its ability to properly judge what influences an individual's confidence level.

In the laboratory experiment, the study was limited by the size, particularly the low number of females in the course. This may have hindered the study's ability to accurately compare males and females within the participant group.

III. Questions for Future Research

Future experiments should consider comparing this extremely brief intervention to a longer social norms intervention with the same impact. The extent and effect of the two interventions, whether performed on a random sample from the field or in the laboratory, could be compared.

Given the opportunity to expand this research, it would be interesting to see how other factors, such as income and demographic information, affect confidence levels and if a brief social norms intervention like this one has a significant impact across different demographics by considering participants for more than just their gender. Future experiments may also consider replicating this study with different information about individuals instead of or in addition to gender. For example, a future study could compare the confidence levels of college-educated individuals to those with only high school degrees, or of high-income individuals compared to lower-income individuals.

The interventions utilized here affected both men and women, so future experiments could attempt to find an intervention that would affect only the intended target group, women and Underconfident women, instead of all participants on average.

References

Beckmann, D., & Menkhoff, L. (2008). Will Women Be Women? Analyzing the Gender Difference among Financial Experts. *Kyklos,61*(3), 364-384. doi:10.1111/j.1467-6435.2008.00406.x

Berkowitz, A. D. (2004, August). The Social Norms Approach. Retrieved from http://www.alanberkowitz.com/articles/social_norms.pdf

Beyer, S. (1990). Gender differences in the accuracy of self-evaluations of performance. *Journal of Personality and Social Psychology,59*(5), 960-970. doi:10.1037/0022-3514.59.5.960

Bleidorn, W., Arslan, R. C., Denissen, J. J., Rentfrow, J., Gebauer, J. E., & Potter, J. (2016). Age and Gender Differences in Self-Esteem—A Cross-Cultural Window. doi:10.31219/osf.io/b65sa

Carlin, B. A., Gelb, B. D., Belinne, J. K., & Ramchand, L. (2018). Bridging the gender gap in confidence. *Business Horizons,61*(5), 765-774. doi:10.1016/j.bushor.2018.05.006

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance.

Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology,55*(1), 591-621. doi:10.1146/annurev.psych.55.090902.142015

Clapp, J. D., Lange, J. E., Russell, C., Shillington, A., & Voas, R. B. (2003). A failed norms social marketing campaign. *Journal of Studies on Alcohol,64*(3), 409-414. doi:10.15288/jsa.2003.64.409

Eckel, C. C., & Füllbrunn, S. (2013). Thar She Blows? Gender, Competition, and Bubbles in Experimental Asset Markets. *SSRN Electronic Journal*. doi:10.2139/ssrn.2271873

Elster, J. (1989). Social Norms and Economic Theory. *Journal of Economic Perspective,3*(4), 99-117. doi:10.1007/978-1-349-62397-6_20

Erhardt, N. L., Werbel, J. D., & Shrader, C. B. (2003). Board of Director Diversity and Firm Financial Performance. *Corporate Governance,11*(2), 102-111. doi:10.1111/1467-8683.00011

Erkut, S. (1983). Exploring sex differences in expectancy, attribution, and academic achievement. *Sex Roles,9*(2), 217-231. doi:10.1007/bf00289625

Fleisher, C., Schoder, D., & Bayer, A. (2017, September 8). The gender gap in economics. Retrieved from https://www.aeaweb.org/research/gender-gap-economics-profession-interview-with-amanda-bayer

Fox, M. F., & Firebaugh, G. (1992). Confidence in Science: The Gender Gap. *Social Science Quarterly*.

Gender Forward Pioneer Index: World's Most Reputable Companies Have More Women in Senior Management. (2016). Retrieved from https://www.webershandwick.com/news/gender-forward-pioneer-index-most-reputable-companies-have-more-senior-wome/

Gneezy, U., Niederle, M., & Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics,118*(3), 1049-1074. doi:10.1162/00335530360698496

Harries, T., Rettie, R., Studley, M., Burchell, K., & Chambers, S. (2013). Is social norms marketing effective? *European Journal of Marketing,47*(9), 1458-1475. doi:10.1108/ejm-10-2011-0568

Hirschfeld, M., Moore, R. L., & Brown, E. (1995). Exploring the Gender Gap on the GRE Subject Test in Economics. *The Journal of Economic Education,26*(1), 3. doi:10.2307/1183461

Horowitz, J. M., Igielnik, R., Parker, K., Horowitz, J. M., Igielnik, R., & Parker, K. (2018, November 13). How Americans View Women Leaders in Politics and Business. Retrieved from http://www.pewsocialtrends.org/2018/09/20/women-and-leadership-2018/

Jakobsson, N., Levin, M., & Kotsadam, A. (2013). Gender and Overconfidence: Effects of Context, Gendered Stereotypes, and Peer Group. *Advances in Applied Sociology,03*(02), 137-141. doi:10.4236/aasoci.2013.32018

Kamas, L., & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization,83*(1), 82-97. doi:10.1016/j.jebo.2011.06.013

Kay, K., & Shipman, C. (2014, May). The Confidence Gap. Retrieved from https://www.theatlantic.com/magazine/archive/2014/05/the-confidence-gap/359815/

Kilmartin, C., Smith, T., Green, A., Heinzen, H., Kuchler, M., & Kolar, D. (2008). A Real Time Social Norms Intervention to Reduce Male Sexism. *Sex Roles,59*(3-4), 264-273. doi:10.1007/s11199-008-9446-y

Kray, L. J., & Kennedy, J. A. (2017). Changing the Narrative: Women as Negotiators—and Leaders. *California Management Review,60*(1), 70-87. doi:10.1177/0008125617727744

Lapinski, M. K., & Rimal, R. N. (2005). An explication of social norms. *Communication Theory,15*(2), 127-147.

Lundeberg, M. A., Fox, P. W., & Punccohar, 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

Martin, A. E., & Phillips, K. W. (2017). What "blindness" to gender differences helps women see and do: Implications for confidence, agency, and action in male-dominated environments. *Organizational Behavior and Human Decision Processes,142*, 28-44. doi:10.1016/j.obhdp.2017.07.004

Meeker, D., Linder, J. A., Fox, C. R., Friedberg, M. W., Persell, S. D., Goldstein, N. J., . . . Doctor, J. N. (2016). Effect of Behavioral Interventions on Inappropriate Antibiotic Prescribing Among Primary Care Practices. *Jama,315*(6), 562. doi:10.1001/jama.2016.0275

Mollen, S., Rimal, R. N., Ruiter, R. A., & Kok, G. (2013). Healthy and unhealthy social norms and food selection. Findings from a field-experiment. *Appetite,65*, 83-89. doi:10.1016/j.appet.2013.01.020

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing Decisions. *Policy Insights from the Behavioral and Brain Sciences,2*(1), 129-140. doi:10.1177/2372732215600886

Nelson, J., Porth, M., Valikai, K., & McGee, H. (2015). A Path to Empowerment: The role of corporations in ... Retrieved from https://www.uschamberfoundation.org/sites/default/files/Path to Empowerment Report Final.pdf

Niessen-Ruenzi, A., & Ruenzi, S. (2006). Sex Matters: Gender Differences in the Mutual Fund Industry. *SSRN Electronic Journal*. doi:10.2139/ssrn.1099768

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative Social Influence is Underdetected. *Personality and Social Psychology Bulletin,34*(7), 913-923. doi:10.1177/0146167208316691

Noland, M., Moran, T., & Kotschwar, B. R. (2016). Is Gender Diversity Profitable? Evidence from a Global Survey. *SSRN Electronic Journal*. doi:10.2139/ssrn.2729348

Palomino, F., & Peyrache, E. (2010). Psychological bias and gender wage gap. *Journal of Economic Behavior & Organization,76*(3), 563-573. doi:10.1016/j.jebo.2010.07.004

Pawlowski, B., Atwal, R., & Dunbar, R. (2008). Sex Differences in Everyday Risk-Taking Behavior in Humans. *Evolutionary Psychology,6*(1), 147470490800600. doi:10.1177/147470490800600104

Pope, D. G. (2017, August 08). Women who are elite mathematicians are less likely than men to believe they're elite mathematicians. Retrieved from https://www.washingtonpost.com/news/wonk/wp/2017/08/08/women-who-are-elite-

mathematicians-are-less-likely-than-men-to-believe-theyre-elite-mathematicians/?noredirect=on&utm_term=.a3c4f9e0e28c

Powell, M., & Ansic, D. (1997). Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology,18*(6), 605-628. doi:10.1016/s0167-4870(97)00026-3

Rimal, R. N., & Lapinski, M. K. (2015). A Re-Explication of Social Norms, Ten Years Later. *Communication Theory,25*(4), 393-409. doi:10.1111/comt.12080

Roth, L. M. (2008). *Selling Women Short: Gender Inequality on Wall Street*. Princeton University Press.

Sarsons, H., & Xu, G. (2015). Confidence men? Gender and confidence: Evidence among top economists. *Harvard University, Department of Economics, Littauer Center,*1-26.

Schultz, W. P., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science,18*(5), 429-434.

Steele, C. M. (2011). *Whistling Vivaldi: How stereotypes affect us and what we can do*. New York: W.W. Norton & Company.

Syed, M. (2008, September 30). What caused the crunch? Men and testosterone. Retrieved from https://www.thetimes.co.uk/article/what-caused-the-crunch-men-and-testosterone-xdfmc637vw0

Turner, J., Perkins, H. W., & Bauerle, J. (2008). Declining Negative Consequences Related to Alcohol Misuse Among Students Exposed to a Social Norms Marketing Intervention on a College Campus. *Journal of American College Health,57*(1), 85-94. doi:10.3200/jach.57.1.85-94

Wechsler, H., Nelson, T. E., Lee, J. E., Seibring, M., Lewis, C., & Keeling, R. P. (2003). Perception and reality: A national evaluation of social norms marketing interventions to reduce college

students heavy alcohol use. *Journal of Studies on Alcohol,64*(4), 484-494. doi:10.15288/jsa.2003.64.484

Woetzel, J., Madgavkar, A., Ellingrud, K., Labaye, E., Devillard, S., Kutcher, E., Manyika, J., Dobbs, R., & Kirshnan, M. (2015). How advancing women's equality can add $12 trillion to global growth. Retrieved from https://www.mckinsey.com/featured-insights/employment-and-growth/how-advancing-womens-equality-can-add-12-trillion-to-global-growth

Wu, A. H. (2017). Gender Stereotyping in Academia: Evidence from Economics Job Market Rumors Forum. *SSRN Electronic Journal*. doi:10.2139/ssrn.3051462

Appendix

Appendix A: Trivia Survey

Trivia Game

1. Check off your answers and select your confidence level for the preliminary questions

2. Write your answers and select your confidence level for each trivia question

3. Check off your answers and select your confidence level for the follow-up questions

No cheating or help from others

Scoring

For correct answers, your confidence level will be added to your overall score

For incorrect answers, your confidence level will be subtracted from your overall score

Therefore, scores can range from -50 to +50

The highest score will receive a $50 Amazon gift card
(please include your email address to be entered to win)

For questions about this research project, please contact Arthur Schneider at the Department of Economics at Trinity College after May 1, 2019 at Arthur.Schneider@trincoll.edu or (860) 297-5158

5

Email (for prize):
Gender:

(Please check): How do you think you will perform on this quiz in comparison to a **male** participant?:

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

(Please check): How do you think you will perform on this quiz in comparison to a **female** participant?:

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

(Please check): How do you think you will perform on this quiz in comparison to the **average** participant?:

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

1. **Which NFL quarterback has been in the most Super Bowls?**_____

   My confidence level is:
   1 2 3 4 5 6 7 8 9 10

2. **What does the N stand for in "NATO"?** _____

   My confidence level is:
   1 2 3 4 5 6 7 8 9 10

3. **What is the longest river in the world?** _____

   My confidence level is:
   1 2 3 4 5 6 7 8 9 10

4. **New York City was originally known by which Dutch name?** _____

   My confidence level is:

   1 2 3 4 5 6 7 8 9 10

5. **Ken Jennings is a popular contestant who won $2.5 million on which show?** _____

   My confidence level is:

   1 2 3 4 5 6 7 8 9 10

(Please check): Do you believe that you performed better or worse than **male** participants?

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

(Please check): Do you believe that you performed better or worse than **female** participants?

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

(Please check): Do you believe that you performed better or worse than the **average** participant?

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

(Please check): Do you think that you did better or worse than you expected?

__ Better

__ Worse

My confidence level is:
1 2 3 4 5 6 7 8 9 10

Appendix B: Summary of Descriptive Statistics for Control Treatment

	Male	**Female**	**Total**
Count	68	71	139
Average Score	8.544	8.563	8.554
Average Questions Correct	2.65	2.51	2.58
Confident	50	48	98
Underconfident	5	17	22
Overconfident	13	6	19
Incompetent	19	19	38
Competent	49	52	101

Appendix C: Summary of Descriptive Statistics for Social Norms Intervention Treatment

	Male	**Female**	**Total**
Count	71	79	150
Average Score	11.79	12.24	11.93
Average Questions Correct	2.79	2.90	2.84
Confident	40	44	84
Underconfident	5	19	24
Overconfident	26	16	42
Competent	63	68	131
Incompetent	8	11	19

Appendix D: Average Score in Control Treatment and SNM Treatment

Descriptive Statistics

Dependent Variable: OverallScore

Treatment	Gender	Mean	Std. Deviation	N
Control	Female	8.56	13.655	71
	Male	8.54	13.879	68
	Total	8.55	13.715	139
SNM	Female	12.24	13.094	79
	Male	11.61	15.419	71
	Total	11.94	14.197	150
Total	Female	10.50	13.444	150
	Male	10.11	14.714	139
	Total	10.31	14.046	289

Appendix E: Difference between Average Scores in Control Treatment and SNM Treatment

Dependent Variable: OverallScore

(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference	
					Lower Bound	Upper Bound
Control	SNM	-3.369*	1.651	.042	-6.620	-.119
SNM	Control	3.369*	1.651	.042	.119	6.620

Appendix F: Male and Female Confidence in Control Treatment Compared to Intervention Treatment

Pairwise Comparisons

Dependent Variable: OverallScore

Gender	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Female	Control	SNM	-3.677	2.292	.110	-8.188	.834
	SNM	Control	3.677	2.292	.110	-.834	8.188
Male	Control	SNM	-3.062	2.378	.199	-7.742	1.619
	SNM	Control	3.062	2.378	.199	-1.619	7.742

Appendix G: Summary of Statistics for Confident, Overconfident, and Underconfident Participants in Control and SNM Treatments

Descriptive Statistics

Dependent Variable: OverallScore

Treatment	Gender	Classification	Mean	Std. Deviation	N
Control	Female	Confident	10.15	13.247	48
		Overconfident	-11.17	14.275	6
		Underconfident	11.06	8.743	17
		Total	8.56	13.655	71
	Male	Confident	13.76	10.052	50
		Overconfident	-11.92	8.616	13
		Underconfident	9.60	7.956	5
		Total	8.54	13.879	68
	Total	Confident	11.99	11.806	98
		Overconfident	-11.68	10.307	19
		Underconfident	10.73	8.407	22
		Total	8.55	13.715	139

SNM	Female	Confident	14.25	13.513	44
		Overconfident	-.44	10.770	16
		Underconfident	18.26	4.357	19
		Total	12.24	13.094	79
	Male	Confident	16.78	15.656	40
		Overconfident	2.04	11.123	26
		Underconfident	20.00	6.892	5
		Total	11.61	15.419	71
	Total	Confident	15.45	14.539	84
		Overconfident	1.10	10.925	42
		Underconfident	18.63	4.862	24
		Total	11.94	14.197	150
Total	Female	Confident	12.11	13.460	92
		Overconfident	-3.36	12.462	22
		Underconfident	14.86	7.616	36
		Total	10.50	13.444	150
	Male	Confident	15.10	12.857	90
		Overconfident	-2.62	12.219	39

	Underconfident	14.80	8.904	10
	Total	10.11	14.714	139
Total	Confident	13.59	13.214	182
	Overconfident	-2.89	12.208	61
	Underconfident	14.85	7.809	46
	Total	10.31	14.046	289

Appendix H: Comparison of Confidence between Control and SNM Intervention Treatments for Confident, Overconfident, and Underconfident Participants

Pairwise Comparisons

Dependent Variable: OverallScore

Classification	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Confident	Control	SNM	-3.560*	1.774	.046	-7.051	-.068
	SNM	Control	3.560*	1.774	.046	.068	7.051
Overconfident	Control	SNM	-12.345*	3.499	.000	-19.232	-5.458
	SNM	Control	12.345*	3.499	.000	5.458	19.232
Underconfident	Control	SNM	-8.802*	4.263	.040	-17.193	-.411
	SNM	Control	8.802*	4.263	.040	.411	17.193

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix I: Comparison of Confident, Overconfident, and Underconfident Male Participants to Confident, Overconfident, and Underconfident Female Participants

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

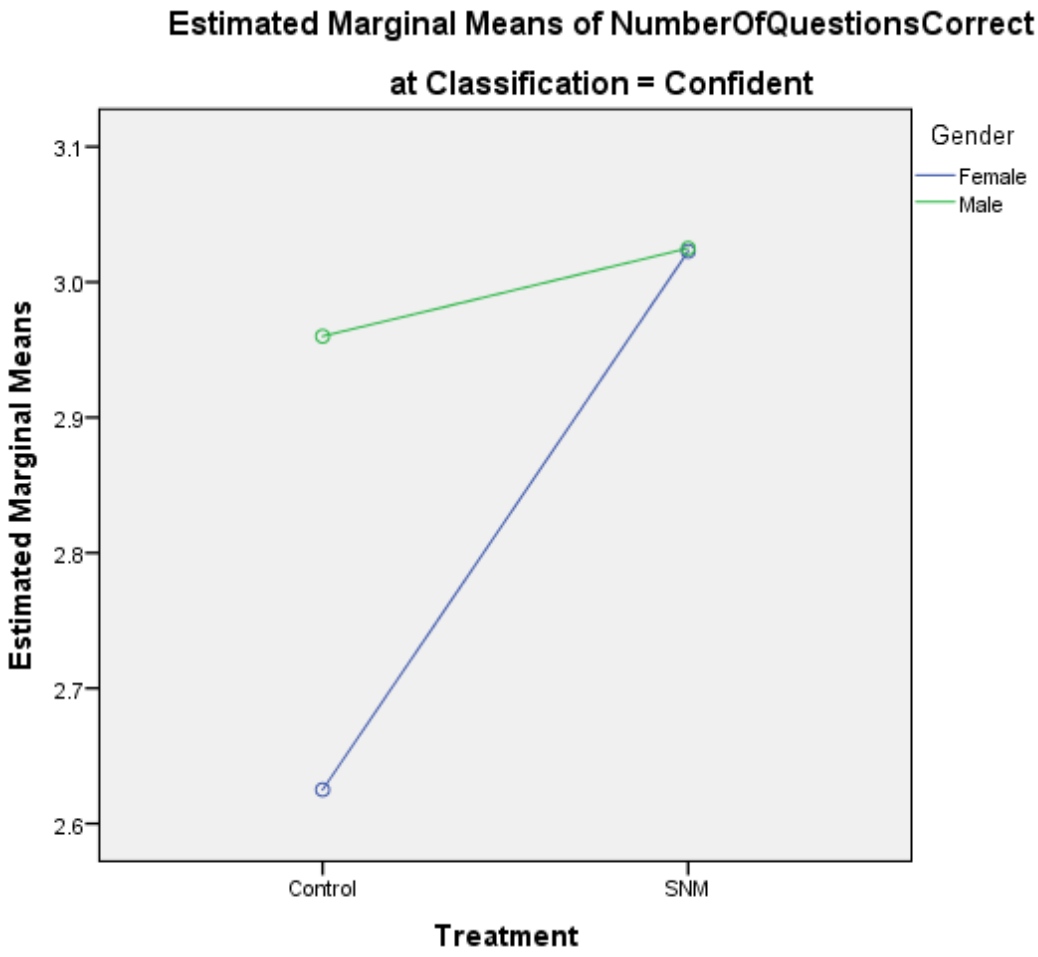
Treatment	Classification	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
Control	Confident	Female	Male	-.335	.192	.081	-.712	.042
		Male	Female	.335	.192	.081	-.042	.712
	Overconfident	Female	Male	.115	.468	.805	-.806	1.037
		Male	Female	-.115	.468	.805	-1.037	.806
	Underconfident	Female	Male	-.271	.482	.575	-1.220	.679
		Male	Female	.271	.482	.575	-.679	1.220
SNM	Confident	Female	Male	-.002	.207	.991	-.410	.405
		Male	Female	.002	.207	.991	-.405	.410
	Overconfident	Female	Male	-.067	.301	.823	-.660	.526
		Male	Female	.067	.301	.823	-.526	.660

Underconfident	Female	Male	-.537	.477	.261	-1.475	.401
	Male	Female	.537	.477	.261	-.401	1.475

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix J: Plot of Confident Males and Females between Control and SNM Treatments



Appendix K: Comparison of Female Confidence After Social Norms Intervention

Pairwise Comparisons

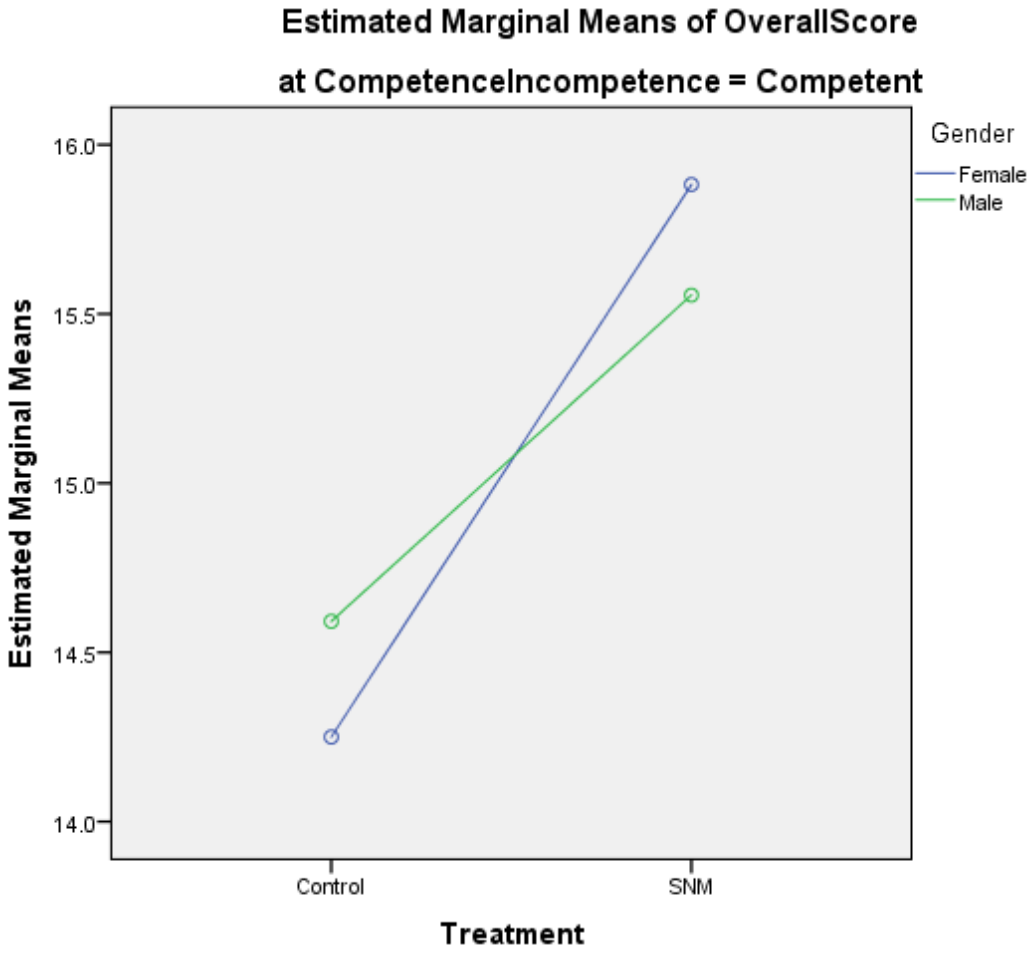
Dependent Variable: OverallScore

Gender	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
Female	Control	SNM	-3.677	2.292	.110	-8.188	.834
	SNM	Control	3.677	2.292	.110	-.834	8.188
Male	Control	SNM	-3.062	2.378	.199	-7.742	1.619
	SNM	Control	3.062	2.378	.199	-1.619	7.742

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix L: Plot Comparison of Competent Males and Females in Control and SNM Treatments



Appendix M: Confidence Comparison between Incompetent Males and Females before and After Intervention

Pairwise Comparisons

Dependent Variable: OverallScore

Treatment	CompetenceIncompetence	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
							Lower Bound	Upper Bound
Control	Competent	Female	Male	-.342	1.990	.864	-4.259	3.576
		Male	Female	.342	1.990	.864	-3.576	4.259
	Incompetent	Female	Male	.053	3.243	.987	-6.331	6.436
		Male	Female	-.053	3.243	.987	-6.436	6.331
SNM	Competent	Female	Male	.327	1.748	.852	-3.114	3.767
		Male	Female	-.327	1.748	.852	-3.767	3.114
	Incompetent	Female	Male	9.227*	4.645	.048	.085	18.370
		Male	Female	-9.227*	4.645	.048	-18.370	-.085

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

Appendix N: Male and Female Participant Performance Before and After Social Norms Intervention

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

Gender	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Female	Control	SNM	-.585*	.196	.003	-.971	-.200
	SNM	Control	.585*	.196	.003	.200	.971
Male	Control	SNM	-.624*	.237	.009	-1.090	-.158
	SNM	Control	.624*	.237	.009	.158	1.090

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix O: Comparison of Confident, OVerconfident, and Underconfident Participants in Control and SNM Treatments

Pairwise Comparisons

Dependent Variable: Number Of Questions Correct

Classification	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Confident	Control	SNM	-.231	.141	.102	-.509	.046
	SNM	Control	.231	.141	.102	-.046	.509
Overconfident	Control	SNM	-.716*	.278	.011	-1.264	-.169
	SNM	Control	.716*	.278	.011	.169	1.264
Underconfident	Control	SNM	-.867*	.339	.011	-1.534	-.199
	SNM	Control	.867*	.339	.011	.199	1.534

Appendix P: Descriptive Statistics of Confident, Overconfident, and Underconfident Participants in Control and SNM Treatments

Descriptive Statistics

Dependent Variable: NumberOfQuestionsCorrect

Treatment	Gender	Classification	Mean	Std. Deviation	N
Control	Female	Confident	2.62	1.160	48
		Overconfident	1.50	.837	6
		Underconfident	2.53	.624	17
		Total	2.51	1.067	71
	Male	Confident	2.96	.807	50
		Overconfident	1.38	.768	13
		Underconfident	2.80	.837	5
		Total	2.65	1.004	68
	Total	Confident	2.80	1.005	98
		Overconfident	1.42	.769	19
		Underconfident	2.59	.666	22
		Total	2.58	1.035	139
SNM	Female	Confident	3.02	1.110	44

		Overconfident	2.12	.885	16
		Underconfident	3.26	.452	19
		Total	2.90	1.020	79
	Male	Confident	3.03	1.097	40
		Overconfident	2.19	.801	26
		Underconfident	3.80	.837	5
		Total	2.77	1.085	71
	Total	Confident	3.02	1.097	84
		Overconfident	2.17	.824	42
		Underconfident	3.38	.576	24
		Total	2.84	1.050	150
Total	Female	Confident	2.82	1.148	92
		Overconfident	1.95	.899	22
		Underconfident	2.92	.649	36
		Total	2.71	1.058	150
	Male	Confident	2.99	.942	90
		Overconfident	1.92	.870	39
		Underconfident	3.30	.949	10

	Total	2.71	1.044	139
Total	Confident	2.90	1.052	182
	Overconfident	1.93	.873	61
	Underconfident	3.00	.730	46
	Total	2.71	1.049	289

Appendix Q: Descriptive Statistics for Confident, Overconfident, and Underconfident Male and Female Participants

14. Treatment * Gender * Classification

Dependent Variable: NumberOfQuestionsCorrect

Treatment	Gender	Classification	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Control	Female	Confident	2.625	.137	2.356	2.894
		Overconfident	1.500	.387	.738	2.262
		Underconfident	2.529	.230	2.077	2.982
	Male	Confident	2.960	.134	2.696	3.224
		Overconfident	1.385	.263	.867	1.902
		Underconfident	2.800	.424	1.965	3.635
SNM	Female	Confident	3.023	.143	2.741	3.304
		Overconfident	2.125	.237	1.658	2.592
		Underconfident	3.263	.218	2.835	3.691
	Male	Confident	3.025	.150	2.730	3.320
		Overconfident	2.192	.186	1.826	2.558

Underconfident	3.800	.424	2.965	4.635

Appendix R: Comparison of Confident, Overconfident, and Underconfident Males and Females in Control and SNM Treatments

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

Gender	Classification	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
							Lower Bound	Upper Bound
Female	Confident	Control	SNM	-.398*	.198	.045	-.787	-.008
		SNM	Control	.398*	.198	.045	.008	.787
	Overconfident	Control	SNM	-.625	.454	.170	-1.518	.268
		SNM	Control	.625	.454	.170	-.268	1.518
	Underconfident	Control	SNM	-.734*	.317	.021	-1.357	-.111
		SNM	Control	.734*	.317	.021	.111	1.357
Male	Confident	Control	SNM	-.065	.201	.747	-.461	.331
		SNM	Control	.065	.201	.747	-.331	.461
	Overconfident	Control	SNM	-.808*	.322	.013	-1.442	-.174

		SNM	Control	.808*	.322	.013	.174	1.442
	Underconfident	Control	SNM	-1.000	.600	.097	-2.180	.180
		SNM	Control	1.000	.600	.097	-.180	2.180

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix S: Means and Standard Deviations of Incompetent and Competent Individuals in Control and SNM Treatments

Estimates

Dependent Variable: NumberOfQuestionsCorrect

Treatment	CompetenceIncompetence	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Control	Competent	2.962	.083	2.800	3.125
	Incompetent	1.553	.135	1.288	1.817
SNM	Competent	3.081	.072	2.939	3.224
	Incompetent	1.119	.193	.740	1.498

Appendix T: Comparison between Control and SNM Treatments for Competent and Incompetent Individuals

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

CompetenceIncompetence	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
Competent	Control	SNM	-.119	.110	.278	-.336	.097
	SNM	Control	.119	.110	.278	-.097	.336
Incompetent	Control	SNM	.433	.235	.066	-.029	.896
	SNM	Control	-.433	.235	.066	-.896	.029

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix U: Comparison between Control and SNM Treatment for Participants who Expected to Perform Better and Worse than Average Males

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

PerformanceMaleBefore	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
Better Than Male Before	Control	SNM	-.308*	.139	.028	-.582	-.034
	SNM	Control	.308*	.139	.028	.034	.582
WorseThanMaleBefore	Control	SNM	.221	.305	.470	-.380	.822
	SNM	Control	-.221	.305	.470	-.822	.380

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix V: Mean and Standard Deviations in Control and SNM Treatment of Females and Males based on Expectations Compared to the Average Male

Estimates

Dependent Variable: NumberOfQuestionsCorrect

Treatment	Gender	PerformanceMaleBefore	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Control	Female	BetterThanMaleBefore	2.587	.153	2.285	2.889
		WorseThanMaleBefore	2.360	.208	1.951	2.769
	Male	BetterThanMaleBefore	2.526	.138	2.255	2.797
		WorseThanMaleBefore	3.273	.313	2.656	3.890
SNM	Female	BetterThanMaleBefore	2.914	.136	2.645	3.182
		WorseThanMaleBefore	2.857	.227	2.411	3.304
	Male	BetterThanMaleBefore	2.815	.129	2.562	3.069
		WorseThanMaleBefore	2.333	.424	1.498	3.169

Appendix W: Comparison between Control and SNM Treatment of Males and Females Sorted by Expected Performance Compared to the Average Male

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

Gender	PerformanceMaleBefore	(I) Treatment	(J) Treatment	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
Female	BetterThanMale Before	Control	SNM	-.327	.205	.112	-.731	.077
		SNM	Control	.327	.205	.112	-.077	.731
	WorseThanMale Before	Control	SNM	-.497	.308	.107	-1.103	.108
		SNM	Control	.497	.308	.107	-.108	1.103
Male	Better Than Male Before	Control	SNM	-.289	.189	.126	-.660	.082
		SNM	Control	.289	.189	.126	-.082	.660
	Worse Than Male Before	Control	SNM	.939	.528	.076	-.099	1.978
		SNM	Control	-.939	.528	.076	-1.978	.099

Based on estimated marginal means

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix X: Comparison based on Expectations Compared to Average Male for Females and for Males in the Control and SNM Treatments

Pairwise Comparisons

Dependent Variable: NumberOfQuestionsCorrect

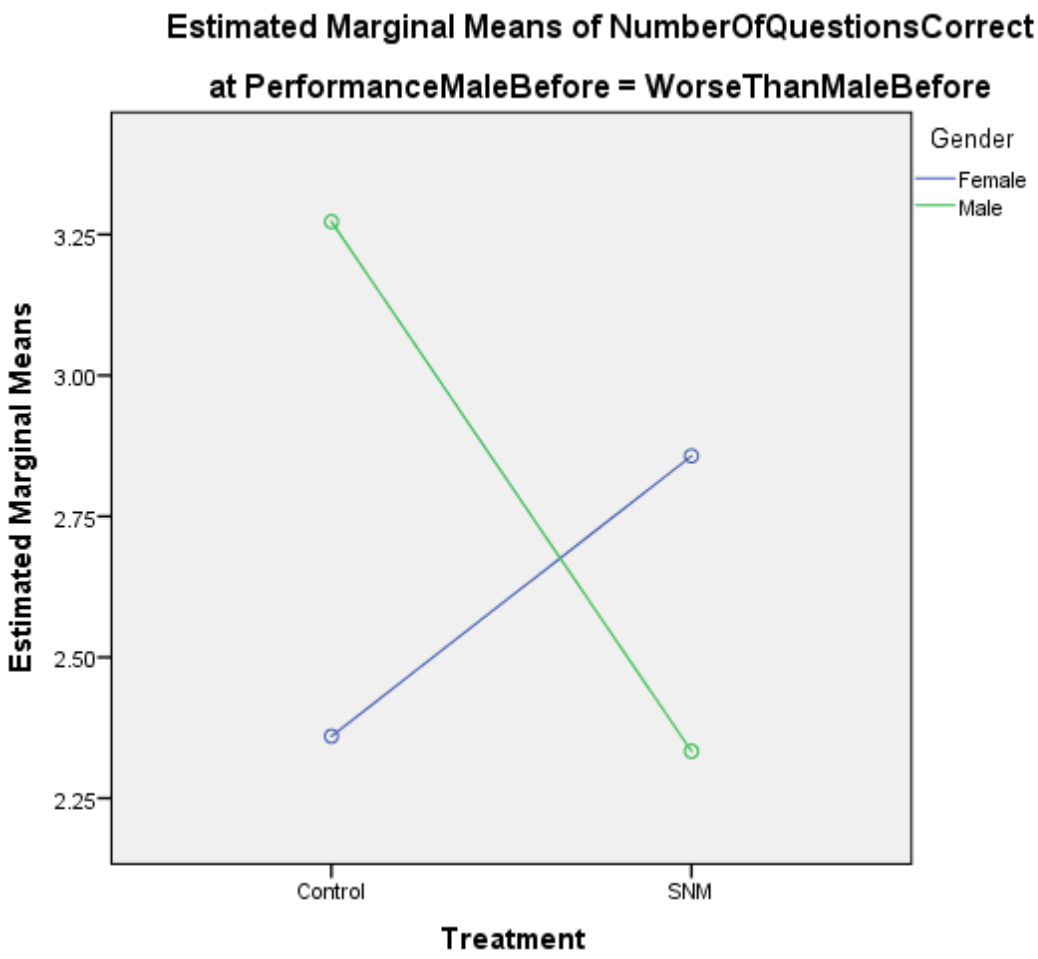
Treatment	Gender	(I) PerformanceMaleBefore	(J) PerformanceMaleBefore	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
							Lower Bound	Upper Bound
Control	Female	BetterThanMaleBefore	WorseThanMaleBefore	.227	.258	.380	-.281	.735
		WorseThanMaleBefore	BetterThanMaleBefore	-.227	.258	.380	-.735	.281
	Male	BetterThanMaleBefore	WorseThanMaleBefore	-.746*	.342	.030	-1.420	-.073
		WorseThanMaleBefore	BetterThanMaleBefore	.746*	.342	.030	.073	1.420
SNM	Female	BetterThanMaleBefore	WorseThanMaleBefore	.057	.265	.831	-.464	.578
		WorseThanMaleBefore	BetterThanMaleBefore	-.057	.265	.831	-.578	.464
	Male	BetterThanMaleBefore	WorseThanMaleBefore	.482	.443	.278	-.391	1.355
		WorseThanMaleBefore	BetterThanMaleBefore	-.482	.443	.278	-1.355	.391

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix Y: Plot of Males Compared to Females who Expected to do Worse than the Average Male in the Control and SNM Treatments



Appendix Z: Table of Descriptive Statistics for Laboratory Experiment

Control Treatment:

	Males	**Females**	**Total**
Average Score	27.52	13.20	25.28
Questions Correct	5.59	4.60	5.44

Social Norms Treatment 1: Female Performance Better

	Males	**Females**	**Total**
Average Score	34.67	23.00	32.84
Questions Correct	5.78	4.40	5.56

Social Norms Treatment 2: Male Performance Better

	Males	**Females**	**Total**
Average Score	20.26	-4.20	16.44
Questions Correct	5.07	2.80	4.72

Appendix AA: Male and Female Performance Comparison between Treatments 1, 2, and 3

Pairwise Comparisons

Measure: MEASURE_1

Gender	(I) TRIVIASCORE	(J) TRIVIASCORE	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
female	1	2	-9.800	10.807	.372	-31.872	12.272
		3	17.400	8.670	.054	-.306	35.106
	2	1	9.800	10.807	.372	-12.272	31.872
		3	27.200*	7.944	.002	10.977	43.423
	3	1	-17.400	8.670	.054	-35.106	.306
		2	-27.200*	7.944	.002	-43.423	-10.977
male	1	2	-7.148	4.651	.135	-16.646	2.350
		3	7.259	3.731	.061	-.360	14.879
	2	1	7.148	4.651	.135	-2.350	16.646
		3	14.407*	3.418	.000	7.426	21.389

3	1	-7.259	3.731	.061	-14.879	.360
	2	-14.407*	3.418	.000	-21.389	-7.426

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AB: Comparison between Male and Female Confidence in Treatments 1, 2, and 3

Pairwise Comparisons

Measure: MEASURE_1

TRIVIASCORE	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
1	female	male	-14.319	9.928	.160	-34.593	5.956
	male	female	14.319	9.928	.160	-5.956	34.593
2	female	male	-11.667	9.785	.242	-31.650	8.317
	male	female	11.667	9.785	.242	-8.317	31.650
3	female	male	-24.459*	7.311	.002	-39.389	-9.529
	male	female	24.459*	7.311	.002	9.529	39.389

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AC: Comparison of Male Scores between Treatments 1, 2, and 3 and of Female Scores between Treatments 1, 2, and 3

Pairwise Comparisons

Measure: MEASURE_1

Gender	(I) TRIVIASCORE	(J) TRIVIASCORE	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
female	1	2	-9.800	10.807	.372	-31.872	12.272
		3	17.400	8.670	.054	-.306	35.106
	2	1	9.800	10.807	.372	-12.272	31.872
		3	27.200*	7.944	.002	10.977	43.423
	3	1	-17.400	8.670	.054	-35.106	.306
		2	-27.200*	7.944	.002	-43.423	-10.977
male	1	2	-7.148	4.651	.135	-16.646	2.350
		3	7.259	3.731	.061	-.360	14.879
	2	1	7.148	4.651	.135	-2.350	16.646
		3	14.407*	3.418	.000	7.426	21.389

3	1	-7.259	3.731	.061	-14.879	.360
	2	-14.407*	3.418	.000	-21.389	-7.426

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AD: Comparison between Male and Female Performance in Treatments 1, 2, and 3

Pairwise Comparisons

Measure: MEASURE_1

QuestionsCorrect	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
1	female	male	-.993	.914	.286	-2.859	.874
	male	female	.993	.914	.286	-.874	2.859
2	female	male	-1.378	.833	.109	-3.079	.324
	male	female	1.378	.833	.109	-.324	3.079
3	female	male	-2.274*	.633	.001	-3.566	-.982
	male	female	2.274*	.633	.001	.982	3.566

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AE: Comparison of Average Performance between Treatments 1, 2, and 3

Pairwise Comparisons

Measure: MEASURE_1

(I) QuestionsCorrect	(J) QuestionsCorrect	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	.007	.472	.988	-.957	.972
	3	1.159*	.443	.014	.254	2.065
2	1	-.007	.472	.988	-.972	.957
	3	1.152*	.358	.003	.421	1.883
3	1	-1.159*	.443	.014	-2.065	-.254
	2	-1.152*	.358	.003	-1.883	-.421

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AF: Male and Female Performance Compared between Treatments

Pairwise Comparisons

Measure: MEASURE_1

Gender	(I) QuestionsCorrect	(J) QuestionsCorrect	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
						Lower Bound	Upper Bound
female	1	2	.200	.867	.819	-1.572	1.972
		3	1.800*	.815	.035	.136	3.464
	2	1	-.200	.867	.819	-1.972	1.572
		3	1.600*	.657	.021	.257	2.943
	3	1	-1.800*	.815	.035	-3.464	-.136
		2	-1.600*	.657	.021	-2.943	-.257
male	1	2	-.185	.373	.623	-.948	.577
		3	.519	.351	.150	-.197	1.234
	2	1	.185	.373	.623	-.577	.948
		3	.704*	.283	.019	.126	1.281
	3	1	-.519	.351	.150	-1.234	.197

2	-.704*	.283	.019	-1.281	-.126

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Appendix AG: Chi-Square Tests for Independence of Variables

Treatment 1 (Control)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Classification * Gender	139	100.0%	0	0.0%	139	100.0%

Classification * Gender Crosstabulation

			Gender		Total
			Female	Male	
Classification	Confident	Count	41	40	81
		Expected Count	41.4	39.6	81.0
		Adjusted Residual	-.1	.1	
	Incompetent	Count	12	11	23
		Expected Count	11.7	11.3	23.0
		Adjusted Residual	.1	-.1	
	Overconfident	Count	7	15	22
		Expected Count	11.2	10.8	22.0
		Adjusted Residual	-2.0	2.0	
	Underconfident	Count	11	2	13
		Expected Count	6.6	6.4	13.0
		Adjusted Residual	2.5	-2.5	

Total	Count	71	68	139
	Expected Count	71.0	68.0	139.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9.135[a]	3	.028
Likelihood Ratio	9.827	3	.020
N of Valid Cases	139		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.36.

Treatment 2 (SNM)

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Classification * Gender	164	100.0%	0	0.0%	164	100.0%

Classification * Gender Crosstabulation

			Gender		
			Female	Male	Total
Classification	Confident	Count	44	37	81
		Expected Count	44.5	36.5	81.0

		Adjusted Residual	-.1	.1	
	Incompetent	Count	11	8	19
		Expected Count	10.4	8.6	19.0
		Adjusted Residual	.3	-.3	
	Overconfident	Count	16	25	41
		Expected Count	22.5	18.5	41.0
		Adjusted Residual	**-2.4**	**2.4**	
	Underconfident	Count	19	4	23
		Expected Count	12.6	10.4	23.0
		Adjusted Residual	**2.9**	**-2.9**	
	Total	Count	90	74	164
		Expected Count	90.0	74.0	164.0