Do the advantages of androgynous leadership extend to teams? Development, validation, and testing of a team androgyny instrument

By
Danielle Mercer-Prowse

A Dissertation Submitted to
Saint Mary's University, Halifax, Nova Scotia
In Partial Fulfillment of the Requirements for
The Degree of Doctor of Philosophy in Business Administration (Management)

December, 2017, Halifax, Nova Scotia

Approved:	Catherine Loughlin, PhD Supervisor
Approved:	Kara Arnold, PhD Committee Member
Approved:	Wendy Carroll, PhD Committee Member
Approved:	Stephanie Gilbert, PhD External Examiner
Date:	December 12, 2017

Acknowledgements

There are too many people who have played a role in my PhD adventure over the last five years – to each and every one of them, I am so utterly thankful. Most importantly, Dr. Catherine Loughlin, who has not only been my doctoral supervisor, but also my supporter, encourager, mentor and confidence booster. She has provided so many fantastic opportunities, ensuring I was given diverse opportunities in research and publishing, and ensuring I had funding. I will always have her "you go girl" in the back of my mind.

Dr. Kara Arnold – if it weren't for you, I wouldn't be here in the first place. Thanks for exposing me to the world of academia and encouraging me to complete a PhD. Thanks also to Dr. Wendy Carroll for serving on my dissertation committee, and Dr. Stephanie Gilbert, my external examiner. There have been several other important professors in the Management Department, who have been instrumental in helping me to be the best academic I can – thanks to each and every one of you. In addition, I would like to thank the Social Science and Humanities Research Council of Canada (SSHRC) for your generous financial support in funding this project.

To my cohort, Corinne, Ryan, Rhonda, Kristine, and Ken, I hit the jackpot with each of you. I don't even want to wonder what my experience would have been like with a different cohort. Thanks for your venting sessions, your research guidance, and most importantly, the lifelong friendships we've created. A special shout out to Dr. Cara Scheuer and Corinne McNally, who have provided constant support as I completed my dissertation, be it sharing articles, making suggestions, or helping solve data analysis issues - I'm so happy that we plan to continue to research together.

Of course, family and friends have been a huge support system in ensuring that I complete this journey and I am forever grateful to each of them – especially Mom, Angie, Steve, Hannah, Jack, Will, Linda, & Rick. Mom – you have no idea how much I appreciated our mother-daughter vacations throughout my PhD – these vacations have allowed me to unwind, recharge, and laugh a ton (and of course enjoy some wine!).

To my husband, Mike, and our little dog, Arthur – if it weren't for the two of you, I probably would have gone insane by now. Thank you for your welcomed distractions, your constant motivation, your incredible patience, and your unconditional love. It's been a long time coming – now let's celebrate!

Finally, and most importantly, Dad - my hero from above, I know you were so very proud of me before, but I can imagine you beaming from ear to ear right now. This one is for you.

Table of Contents

List of Appendices

List of Tables

Abstract

Do the advantages of androgynous leadership extend to teams? Development, validation, and testing of a team androgyny instrument

By Danielle Mercer-Prowse

Recent business scandals (e.g., Lehman Brothers) raise questions regarding the role of decision-making in shaping outcomes such as performance. Researchers have examined decision-making of organizational *teams* from a variety of perspectives, however, limited work has studied androgynous decision-making. Researchers have argued that androgynous leaders are more effective than leaders who are masculine or feminine. I argue that this may extend to the team level. The purpose of my dissertation is to use the theory of androgyny (Bem, 1974) as a conceptual basis to support the development of a *new* instrument that examines androgynous decision-making in organizational *teams*. In three studies, my aims are to develop/refine a new instrument to measure team androgyny, provide evidence of convergent/divergent validity with other measures, and demonstrate the utility of my team androgyny instrument by showing its ability to predict outcomes related to risk propensity, ethical responsibility, and objective team performance. Additionally, I examine the interplay of sex composition *and* team gender on performance. Study 1 was exploratory and included a literature review, consultation with subject matter experts, a pilot study and main study to refine the item pool and test the factor structure of team androgyny. In Study 2, using the refined instrument, the items factored onto the two hypothesized subscales (i.e., masculinity and femininity) and demonstrated convergent/divergent validity. At the aggregated level, as hypothesized, team masculinity predicted risk propensity, and team femininity predicted ethical responsibility. Team androgyny was unrelated to team performance; However, masculine teams had the *highest* performance via team grades. Study 3 used an experimental design to replicate/confirm previous results. Team sex composition (i.e., male) *and* team masculinity negatively related to performance, no other main effects emerged. This dissertation provides support for my team androgyny instrument and demonstrates that team sex *and* gender impacts performance, but how depends on the task.

December 12, 2017

1: Introduction and Literature Review

Background of the Dissertation

In 2008, Lehman Brothers Holdings Incorporated filed for the largest bankruptcy in history and intensified the financial crisis. This company was represented *almost* exclusively by male leaders who engaged in unethical and irrational decision-making, with critical moral lapses (Turner, 2012). Contrast this with an Icelandic company called Audur Capital – the only financial services firm to emerge from the 2008 financial crisis unscathed. Audur Capital was a female founded and led firm committed to incorporating feminine values into its decision-making in the financial sector (i.e., risk awareness, profit with principles, emotional capital, etc.). In Iceland specifically, the outcome of the crisis led to a multitude of women being elected into office *and* some of the biggest micro-financiers only lending to women (Newton-Small, 2016). Scenarios like this have prompted some to ask – what might have been different if "Lehman brothers had been Lehman sisters?" (e.g., Moss-Kanter, 2010, para. 5).

Around the globe, legislators are initiating calls for changes to increase the representation of women on corporate boards (e.g., Canada – Status of Women Canada, 2014; Norway, Spain, and France – Sealy & Vinnicombe, 2010) from a justice perspective (i.e., because women comprise 50% of the population and they should be proportionally represented), and as a means for improving performance (Singh, Terjesen, & Vinnicombe, 2008). Furthermore, increasing the representation of women on decision-making teams is expected to improve not only the heterogeneity of attitudes, behaviors, and perspectives brought to the table (Eagly, 2005) but also improving outcomes such as ethical responsibility, risk orientation, collaboration, and open communication in decision-making bodies (e.g., Woolley, Chabris, Pentland, Hashmi, & Malone, 2010).

Although many researchers argue that more women on organizational teams will improve decision-making, and thus, performance, to date there has been little research directly addressing this issue at the *team level of analysis* (Groysberg & Bell, 2013).

Research examining the performance of mixed-sex teams has been inconsistent (Post & Byron, 2014). Some researchers see positive effects when women are added to teams (Catalyst, 2003; 2007; Clark, 2013) whereas others see no effect (Rohner & Dougan, 2012), or a negative effect (Adams & Ferreira, 2009; Ahern & Dittmar, 2012; Webber & Donaue, 2001). At the board level, studies typically examine macro level outcomes such as financial performance and utilize secondary data (e.g., company annual reports). At the organizational team level, empirical and/or experimental studies testing the impact of biological sex composition on team performance are equivocal (e.g., Apesteguia, Azmat, & Iriberri, 2010; Hoogendoorn, Oosterbeek, & Praag, 2013).

In this dissertation, I will explore links between sex composition, team gendered decision-making styles (i.e., team androgyny), and objective performance. My dissertation does *not* use sex and gender as interchangeable terms; *sex* relates to the biological categories of male and female (e.g., Powell, 2012) whereas *gender* refers to the "social-psychological categories of masculinity and femininity" (Berdahl, 1996, p. 23) – while gender tends to be stereotypically associated with biological sex, the relationship is much more complex (Abele, 2003). Likewise, Bem's (1974) pioneering work on gender roles argued that masculinity and femininity were independent constructs and an individual could espouse both traits (i.e., androgyny), regardless of biological sex. For some time now, researchers have argued that androgynous leaders (those high on masculinity and femininity) are more effective than leaders who are predominately masculine or feminine (e.g., Higher quality decision – Kirchmeyer, 1996, Gershenoff &

Roseanne, 2002; Effective decision-making – Radecki & Jaccard, 1996; Leader emergence – Kent & Moss, 1994; Kolb, 1997; Broader range of leadership skills – Stephens, 2005) because they are able to adapt their behaviors depending on the decision at hand. My dissertation argues that this *conceptualization* may extend to the team level of analysis. Specifically, I use the theory of androgyny (e.g., those high on masculinity and femininity – Bem, 1974; Spence, Helmreich, & Stapp, 1974) as a *theoretical* basis for the development of a *new* androgyny measure, which is designed to specifically examine gendered decision-making of organizational *teams*. Consequently, I hypothesize that sex balanced (as opposed to all male or all female) and androgynous teams (utilizing both masculine and feminine behaviors) *may* perform better than primarily masculine or feminine teams, given their ability to situationally adapt. I define *team androgyny* as the extent to which a team perceives it is utilizing both masculine (e.g., risk taking, confidence, assertive etc.) and feminine (e.g., ethical reasoning, collaboration, open communication etc.) decision-making styles.

Purpose and Contribution of this Dissertation

Given the lack of systematic/empirical research in regard to *sex, gender,* and organizational team performance, the purpose of my research is to use the theory of androgyny (i.e., being high in masculinity and femininity – Bem, 1974; Spence et al., 1974) to develop a *new team-level* measure that could be used by future researchers interested in investigating gendered decision-making styles (i.e., masculine and feminine) of organizational teams. The theory of androgyny was initially proposed by Bem (1974) and her conceptualization was carried out through a measure called the Bem Sex Role Inventory (BSRI). Since that time, interest in androgyny increased dramatically, and several scales have been designed, developed, and used to examine gender role behaviors

(e.g., Personal Attributes Questionnaire [PAQ] – Spence et al., 1974; Traditional Masculinity-Femininity scale [TMF] – Kachel, Steffens, Niedlich, 2016). Despite the extensive use of the theory of androgyny in past literature, relatively no work has been done in regard to androgyny at the team level of analysis. Therefore, my aim was to use the theory of androgyny from a conceptual standpoint in helping to define *team* androgyny in the specific context of business organizations. While I use past gender role scales as a foundation to inform my literature review and item search strategy, my first objective was to develop and validate a new measure of androgyny examining the gendered decision-making styles of organizational teams. The development of this instrument allowed me to examine, refine, and confirm the factor structure of team androgyny. Additionally, I collected evidence on convergent/divergent validity for my team androgyny instrument by examining existing psychometric measures related to masculinity and femininity (e.g., risk propensity, ethical responsibility, collaboration/cooperation etc.). Specifically, I included existing scales related to risk propensity and ethical responsibility given one of my interests in this dissertation stemmed from the Lehman Brothers example (a case study will be provided later in this dissertation). There is a foundation in the literature that supports the notion of masculinity/risk and femininity/ethics and I aimed to test if my team androgyny instrument correlated with these constructs.

The second purpose of my dissertation was to investigate specific outcomes of team androgyny at the *aggregated level of analysis*. Specifically, I aimed to assess if my team androgyny instrument (i.e., team masculinity and team femininity) and/or team sex (i.e., all male, all female, or mixed sex) predicted risk propensity, ethical responsibility, and most importantly, objective team performance using decision-making scenarios (i.e.,

team course project – Study 2; Winter Survival Exercise – Study 3) - The most crucial being team performance given the inconsistent current research surrounding sex composition and firm performance of companies (Post & Byron, 2014). This could have serious theoretical and practical contributions to knowledge. Practically speaking, organizations understanding the importance of both sex balance and gender balance of their respective teams may lead to "smarter" decision-making (e.g., balancing risk/ethics), and ultimately performance. As such, they may be better equipped to design their teams accordingly.

Literature Review

Theoretical Framework: Androgyny

Prior to the 1970s, femininity and masculinity were argued to be bipolar ends of a single dimension (Bukowski, Panarello, & Santo, 2016); an individual could be masculine or feminine but not both (Johnson, 1988). More specifically, the absence of one trait meant the presence of the other trait (Foushee, Helmreich, & Spence, 1979). Bem's (1974) groundbreaking work on gender roles argued that masculinity and femininity were independent constructs and that one individual could enact both traits – a term coined psychological androgyny. According to Bem (1974), regardless of biological sex, an individual could be high or low in masculinity (i.e., defined as assertive, competitive, risk oriented, and direct), high or low in femininity (i.e., defined as communal, collaborative, and communicative), or be high in both (i.e., androgyny). She argued that by assuming gender roles were bipolar ends of a single dimension, people were categorized as one or the other, most often in accordance to their biological sex – men were expected to adhere to masculine norms and women to feminine norms. Many other feminist researchers shared the belief that a unidimensional scale created negative consequences and

restrictions for both men and women, and encouraged gender stereotypes (Donnelly & Twenge, 2016).

Bem's (1974) conceptualization of psychological androgyny meant that men or women had the *flexibility* to identify with various gendered behaviors. More importantly, Bem and her colleagues (Bem & Lewis, 1975; Bem, 1985) argued that being androgynous was advantageous because an individual could "cross gender boundaries" (Martin, Cook, & Andrews, 2016, p. 3) and change his/her behavior depending on what the situation required. For example, an androgynous individual could be assertive *and* affectionate, as such, that individual had greater adaptability and would be more effective, competent, and a better problem solver than individuals that only identified with masculinity *or* femininity (Bukowski et al., 2016).

To empirically test psychological androgyny, Bem (1974) developed the Bem Sex Role Inventory (BSRI), which consisted of 60 self-descriptive items, and measured "the extent to which a person divorces himself from those characteristics that might be appropriate for the opposite sex" (p. 156). The inventory consisted of three subscales – a masculine, a feminine, and a *filler* scale (e.g., assumed to be gender neutral) and was rated on a 7-point Likert Scale ranging from 1 (never or almost never true) to 7 (always or almost true). Originally, Bem's operationalization of the BSRI was that if an individual did not distinguish between masculinity and femininity (e.g., scored low *or* high in both) than they were classified as androgynous. However, scholarly critiques (i.e., Spence et al., 1975) of the BSRI noted that Bem did not account for those who scored *high* versus those who scored *low* in the masculine and feminine subscales. Bem (1977) addressed these scoring critiques in an empirical study with 665 undergraduate students. Individuals who scored low in masculinity and femininity would be classified as "undifferentiated"

because of their inability to be behaviorally flexible. Furthermore, "androgyny" would be defined as having *high* absolute levels of both masculinity and femininity. Most importantly, the development of the BSRI showed that individuals did not have to be defined by one dimension or the other (Martin et al., 2016).

Despite the BSRI being used extensively and being considered the instrument of choice in measuring masculinity and femininity over the last several decades (Powell & Butterfield, 2015), it has received some criticism about item development and measurement procedures (Heilbrun, 1976; Spence et al., 1974; 1975). For instance, some researchers argue that the items used to describe masculinity and femininity were not universal or accurate representations (Johnson, 1988) of the constructs. Others claimed that the overall instrument lacked validity (i.e., Bem's lack of clarity about what the BSRI is intended to measure) *and/or* reliability (i.e., due to scoring procedures) (Hoffman & Borders, 2001). Concerns regarding the measure also relate to the changes that *may have* occurred in the roles of males and females since the 1970s.

Although controversy ensued regarding the BSRI, most researchers agree with the theory of androgyny and "no one has discredited her conceptualization of femininity and masculinity as distinct constructs that compliment rather than contradict each other" (Bukowski et al., 2016, p. 2). For example, Holt & Ellis (1998) conducted a partial replication of the method Bem used to validate the BSRI. Using 138 undergraduate psychology students, the authors found that all of the masculine items were still more desirable for a man, and all but two feminine items (i.e., loyal and childlike) were still more desirable for a woman.

While Bem (1979) did have concerns with the instrument as reinforcing gender stereotypes by using the terms masculine and feminine, Spence (1985) addressed this

concern by emphasizing the distinction between biological sex and gender. An individual could tap into socially desirable gendered traits of masculinity and femininity whether they were a man or a woman. Nonetheless, an individual's behavior could be gendered, regardless of their biological sex. To date, Bem and other similar researchers (i.e., Spence et al., 1975) conceptualizations of androgyny are still seen as a viable method in studying gender (Bukowski et al., 2016; Leaper, 2015). Specifically, androgyny continues to be examined in the realm of leadership (e.g., Kark, Waismel-Manor, & Shamir, 2012), management (McGregor & Tweed, 2001), well-being (Woodhill & Samuels, 2003) and mental health (Leftkowitz & Zeldow, 2006), and creativity (Norlander, Erixon, & Archer, 2000), to name a few.

In my dissertation, I use Bem's theory of androgyny as a conceptual foundation in defining masculinity and femininity as two distinct constructs. Bem's (1974) conceptualization centered around the notion that individuals can identify with masculine and/or feminine sex role traits, regardless of biological sex, and those who are high in both gendered traits are considered androgynous. However, operationally I extend the conceptualization to both the *team level* and focus specifically on the context of *organizational decision-making styles* (i.e., risk propensity and ethical responsibility). Specifically, I argue that teams can have situational adaptation in relation to gender roles, and have the capacity to be androgynous (i.e., *using* both masculinity and femininity). Therefore, while my instrument has a masculine and feminine subscale, these scales are specific to teams within organizations. In support of this, Bem (1977) noted that androgyny could be relevant to many different experiences beyond her research. I argue that *teams* that utilize an *androgynous* decision-making style (i.e., one using both masculine and feminine gender role behaviours) may outperform (i.e., on an objective

task) teams who are exclusively masculine or feminine. Additionally, given my focus on the context of business organizations in terms of recent scandals (e.g., Lehman Brothers, Audur Capital, etc.), I argue that ethical responsibility will be associated with a more feminine decision-making style whereas risk propensity a more masculine style (literature supporting this discussed below), regardless of sex composition of the team.

In my operationalization of team androgyny, each individual does not have to be high in masculine or feminine traits, nor do I require an equal number of masculine and feminine individuals. I define team androgyny as a team that perceives it is utilizing *both* masculine and/or feminine decision-making styles. In the following sections, I demonstrate how decision-making behaviors (e.g., risk propensity, ethical responsibility etc.) may be associated with both *sex* and *gender* and how these could have an impact on organizational performance especially in light of the increase in team-based structures.

Biological Sex and Decision-Making

Do men and women differ on decision-making styles?

For decades, research on decision-making *and* performance has focused explicitly on differences between men and women. Whether due to nature or nurture (Booth & Nolen, 2012), *in general*, at the individual level of analysis, men are viewed as more risk seeking and women more risk-averse (Powell & Ansic, 1997). Harris, Jenkins, & Glaser (2006) examined the likelihood of males versus females engaging in risky activities (e.g., gambling, health, recreation, and social). Using 389 females and 268 males, the authors found that women were significantly less likely to engage in risky behaviors in an effort to lessen negative outcomes (e.g., financial loss). Powell & Ansic (1997) found that women had significantly lower preference for risk in financial decisions. The study also suggested that males and females adopt different strategies in financial decision-making

(see also Byrnes, Miller, & Schafer, 1999 meta-analysis). Some scholars believe that men's propensity to embrace risk makes them better decision-makers (Powell & Ansic, 1997). Others argue that women's ability to be more conservative in relation to risk is key to successful decision-making outcomes (Adams, 2016). More generally, in terms of the behavioral aspects of decision-making, men are also perceived as more proactive in crisis situations (see Alonso-Almeida & Bremser, 2015).

Several researchers also argue that men and women differ on the value they place on moral reasoning and ethical decision-making. A recent study by Ho, Li, Tam, & Zhang (2015) examined CEO gender and both *risk aversion* and *ethical sensitivity* in regard to accounting conservatism. Using COMPUSTAT (i.e., database of financial and statistical information on global organizations) data (1996 to 2008) of firms with assets greater than 10-million dollars, the authors found that in organizations with high litigation and risk, the presence of a female CEO positively correlated with accounting conservatism. A similar study by Chan, Jamilah, & Rusinah (2012) found that women tend to be more sensitive to ethics whereas men focus more on the end result regardless of ethical conduct. Likewise, Glover, Bumpus, Sharp & Munchus (2002) examined biological sex and individual decision-making. The authors concluded that women were more likely to make more ethical decisions than men and were also more consistent in their ethical choice behavior regardless of the issue in question. Studies in the context of organizational decision-making show that women are stricter than men in their ethical stance and behavior (e.g., Lund, 2008; Valentine & Rittenburg, 2004), are higher in moral orientation (Galbreath, 2011), and are more likely to practice whistle-blowing (Rothschild & Miethe, 1999). Women are also more likely to take a 'care approach' (i.e., concern for human welfare – Bampton & Maclagan, 2009), and utilize egalitarianism or collaboration

when making a decision (Andreoni & Vesterlund, 2001; Cadsby, Servatka, & Song, 2010; Dufwenberg & Muren, 2006; Guth, Schmidt, & Sytter, 2001; Miller & Ubeda, 2012).

Does the ~~gender~~ biological sex composition of the team affect decision-making?

Similar instances of biological sex differences in decision-making behavior have been found within teams. For example, in regard to risk orientation, Bogan, Just, & Dev (2013) found that team composition influences financial decisions in relation to risk and loss. Having men on a team increased the likelihood of choosing a high-risk investment. As well, all male teams were more risk seeking than all female teams (See also Castillo, Leo, & Petrie, 2012; Ertac & Gurdal, 2012). Likewise, Karakowsky & Elangovan (2001) conducted a study with 163 business students aiming to examine decision making under risk and uncertainty at both the individual level and the team level. Participants were assigned to 28 groups of different sex compositions and individually completed four decision-making scenarios, then were required to complete one response for each scenario within their team. Unsurprisingly, men appeared to be greater risk takers than women. Furthermore, in team settings, only male members influenced group decision-making processes in regard to being more risk tolerant. Therefore, male group members rather than female group members determined the risk preference of the team further supporting the notion that risk propensity may be viewed as masculine. On the other hand, female team representation tends to lead to (more strategic) decisions that mitigate risks. Nonetheless, a recent meta analytic study by Jeong & Harrison (2016) explored how female representation on top management teams and CEO positions affected firm performance. Drawing on over 140 studies in several countries over many decades, the authors deemed that there is no evidence of women decreasing long-term performance. Furthermore, there are positive associations between a firm's fiscal outcomes and female

representation in CEO positions of top management teams. A contributing factor in why female representation (i.e., CEO) in top management teams might result in positive financial performance is that they *may* reduce extensive risk-taking and encourage strategic decision-making. In a study by Perryman, Fernando, & Tripathy (2016), they also found that as top management teams became more sex balanced, risky behavior decreased and performance of the firm increased.

Other researchers argue that in team decision-making tasks, positive performance relates to *competitive* behavior. An experimental study by Ivanova-Stenzel & Kübler, (2009) sought to explore pay for performance versus competition between teams based on sex composition. Participants were asked to solve numerous memory games in a certain time frame. Payoffs were analyzed by adding up the total number of games the teams solved then given a payout based on that number. Team competition was analyzed by comparing results to another randomly selected second team and calculating total points. The findings revealed that the sex of the team accounted for difference in performance related to both incentives and competition. More specifically, male teams' performance was significantly higher than females in terms of payouts because men tended to exert more effort in winning the experiment. Furthermore, competition among all male or all female teams showed a significant gap in performance – men performed better than women relative to payouts. Similarly, in team situations, men tend to enact power displays more frequently (e.g., interrupting team members to get their point across), this may help or hinder the team's performance. Karakowsky, McBey, & Miller (2004) examined the sources of influence on power displays in mixed-sex work teams and found that in male-dominated teams, men showed high levels of power displays (i.e., interrupting others); however, the power displays of men lessened when the team was sex

balanced. Weis, Kolbe, Grote, Dambach, Marty, Spahn, & Grande (2014) studied two-person teams in a simulated environment and also found that agency (i.e., masculinity) was positively associated with speaking up behaviors. Other researchers argue that in team settings men operate more confidently (Barber & Odean, 2001; Hugelshafer & Achtziger, 2014) in terms of making a decision that directly affects performance.

Similarly, at the individual level, male decision-makers (i.e., leaders) are traditionally defined as being assertive. In examples where assertiveness has been studied at the team level, we also see that males *may* employ assertive tendencies when making decisions that have an outcome related to performance. LePine, Hollenbeck, Ilgen, Colquitt, & Ellis (2002) specifically examined a team's sex composition in relation to decision-making accuracy of a masculine typed task. The study involved a computerized decision-making task (i.e., Team Interactive Decision Exercise for Teams Incorporating Distributed Expertise) with eighty teams of four undergraduate students. Teams were randomly assigned to various conditions (e.g., all males, 25% male, 50% male, 75% male, and 100% male). Unlike past studies that show that all male teams outperform sex balanced teams (e.g., Wood, 1987), LePine et al. (2002) found that decision-making accuracy (i.e., extend of overaggressive decision-making) was higher in all female, majority female, or sex-balanced teams. Most noteworthy is the finding that men tend to act unintentionally overaggressive in team settings sometimes leading to poorer performance. In contrast, Apesteguia, Azmat, & Iriberri (2012) investigated whether the sex composition of a team would affect economic performance through the use of an experimental game (i.e., L'Oréal e-Strat Challenge – online business simulation). The authors found that all female teams underperformed in comparison to all male and mixed

sex teams' due to differences in decision-making styles. More specifically, the female teams tended to be less assertive in pricing strategies and were less likely to invest.

More recently, research examining the inclusion of women on organizational teams (e.g., corporate boards) supports the notion that they encourage more ethically responsible decisions. For instance, Isidro & Sobral (2015) investigated the indirect and direct effects (i.e., financial performance and ethical compliance) of women on corporate boards relative to firm value by using a simultaneous equation model with 992 boards throughout Europe. The authors concluded that there is a positive effect of women on the board relative to higher ethical/social compliance of the firm. Furthermore, since women on the board improved ethical responsibility, indirectly this had a positive effect on the overall firm's value. Likewise, other researchers show that when women are added to teams they not only increase ethical compliance, but *may* also increase the relational aspects of decision-making. Galbreath (2011) investigated women on boards and corporate sustainability and a noteworthy finding was that women contribute positively to issues of sustainability due to their ability to establish relationships within the board and with external stakeholders. Additionally, Bear, Rahman, & Post (2010) examined how biological sex diversity of boards affects corporate social responsibility ratings. Using data from Fortune's 2009 Most Admired List, the authors found that the number of women on the board positively correlated with corporate social responsibility ratings. More specifically, women induced greater *sensitivity* and *participative* decision-making styles within the board (See also Rao & Tilt, 2016). Finally, Woolley et al. (2010) conducted two studies that sought to examine the collective intelligence of groups. Using 699 people working in mixed-sex groups on a variety of cognitive related tasks (e.g., visual puzzles, making collective moral judgments etc.), the authors found that collective

intelligence is not correlated with the average individual intelligence of group members, but rather with the average social sensitivity, conversational turn taking, and number of women in the group.

The studies above demonstrate how men and women may possess specific behaviors that lead to more masculine or feminine decision-making styles. More importantly, these behaviors are also evident within teams. However, the literature surrounding decision-making styles of teams and performance is mixed. For example, some researchers argue that risk tolerance leads to higher performance whereas others argue that risk consciousness leads to more careful decision-making, and thus indirectly improves performance. Furthermore, while there is some controversy regarding sex composition of teams (specifically all male or all female teams), the examples above show the potential benefits of *mixed sex* teams in terms of performance outcomes and decision-making (e.g., Hoogendoorn, Oosterbeek, van Praag, 2011). Therefore, I argue that these findings support my claim that sex balanced teams may benefit from the strengths of each sex. As such, I hypothesize that team sex (i.e., mixed sex) will positively predict team performance.

Additionally, consistent with my earlier theoretical arguments around risk propensity, ethical responsibility, and gender/sex, I argue that my team androgyny instrument will demonstrate similar findings to the individual based findings discussed above. Specifically, because some researchers have found controversial results surrounding risk/ethics and sex/gender composition of teams and/or did not distinguish between biological sex/gender in their studies, I contend that risk may be perceived as a masculine decision-making style and ethics a feminine decision-making style. I therefore hypothesize that teams classified as masculine will have higher levels of risk propensity

whereas feminine teams will demonstrate less risk propensity (regardless of biological sex). Likewise, the more a team associates with feminine decision-making styles, the higher it will be on ethical responsibility, and the more masculine a team, the lower on ethical responsibility. In the section that follows, I investigate the leadership literature to further support how androgyny may be advantageous for individuals and teams within business organizations.

Success of Androgynous Leadership

To date, although there have been countless studies examining sex in various disciplines including leadership, studies specific on *gender* and decision-making in team settings are quite scarce and there have been few instances in the literature examining groups. Consequently, the majority of this section will borrow from the *individual level* leadership literature. I argue that showing the success of androgynous leaders who are at the cornerstone of decision-making will guide my argument to extend the conceptualization of androgyny to organizational teams.

In one study explicitly focused on androgynous decision-making within teams, 160 undergraduate business students were asked to participate in a group decision-making exercise and complete the Bem Sex Role Inventory (BSRI). Kirchmeyer (1996) found that groups with more androgynous members made higher quality decisions than masculine or feminine groups. On the other hand, Radecki & Jaccard (1996) examined the relationship between gender role identification and decision-making skills of *individuals*. Using self-report questionnaires, the authors found that participants perceived both masculinity and femininity as important for effective decision-making, thus supporting androgyny.

Other literature has examined links between androgyny and leader emergence (e.g., Brenner, Tomkiewicz, & Schein, 1989; Goktepe & Scheier, 1989; Korabik, 1990). For example, Kent & Moss (1994) found that androgynous individuals were more likely to emerge as a leader in small group settings. Furthermore, gender was deemed a better predictor of leader emergence than biological sex. Similarly, Kolb (1997) explored self and group reported assessments of leader emergence in relation to biological sex and gender roles. The findings reported no significant differences in the assessments of leader emergence in relation to biological sex, but androgynous individuals were more likely to emerge as leaders. More recently, Gershenoff & Roseanne (2002) studied the effect of gender role and intelligence on leadership emergence in all female groups. The findings revealed that an androgynous-intelligent personality led to leadership emergence.

Other researchers have found androgynous leaders to be considered more effective than highly masculine or feminine leaders. Stephens (2005) investigated the relationship between leaders' gender-role orientation and ratings of their effectiveness across multiple leadership skills (e.g., task, people and conceptual skills) rated by themselves and their direct reports. The author found that leaders classified as androgynous were able to manage a broad range of leadership skills in comparison to masculine or feminine leaders. That is, androgynous individuals fully capitalized on task (i.e., agentic) and/or people (i.e., communal) orientations depending on the situation. Second, Powell & Butterfield (2015) conducted a study examining the role of androgyny in leader prototypes from 1979-2015. The authors found that perceptions of a good manager have become more androgynous over time, as more people believe that managers should possess a balance of both masculine and feminine traits.

While past research has examined the sex composition of teams and performance/decision-making, the same cannot be said for gender (i.e., androgyny). However, adopting the leadership literature relevant to gender roles at the individual level of analysis suggests that androgynous teams *may also* have a wider repertoire of behaviours at their disposal, and thus, *may* be more flexible in using either style (i.e., masculine and/or feminine decision-making) that lead to optimal performance across situations. With that said, I hypothesize that this may extend to the team level, as such, in the context of organizational gendered decision-making, team androgyny may be a better predictor of performance than team femininity or masculinity. To reiterate, I define team androgyny as a team that perceives it is utilizing both masculine and feminine decision-making styles.

Good Decision-Making Today – More Androgynous?

In the introduction, I briefly discussed the impact that poor decision-making can have on organizational outcomes such as performance (e.g., Lehman Brothers Holdings Incorporated). Some of the worst business decisions have been the result of excessive risk taking without consideration of ethical responsibility. For example, Volkswagen is now paying a high price for failing to consider the environment versus its own profits in decision-making (Bansal, King, & Seijts, 2015). Firestone's longstanding success halted overnight when a new kind of tire was introduced by its competitor, Michelin. While Firestone reacted promptly, it ignored warnings that their decision-making was quite risky, resulting in a loss of over 200 million and eventually leading them to be acquired by Bridgestone (Sull, 1999). Likewise, Firestone and Ford Motor Company's controversy regarding tire installation and the Ford Explorer, resulted in hundreds of fatalities and injuries due to cheap manufacturing and failing to act ethically in their production.

Similarly, situations in oil and gas, for example, Deepwater Horizon, resulted from a series of questionable decisions and unsafe practices by BP whereby drilling continued, despite the results from pressure tests suggesting they should halt production (causing a blowout that killed several workers and released millions of barrels of oil into the Gulf of Mexico; National Academy of Engineering, 2012).

While these examples do not explicitly discuss sex and/or gender role behaviours in conjunction with decision-making, they do focus on *risk* and *ethics.* I have already demonstrated the links between decision-making such as risk propensity and ethical responsibility, sex and/or gender. Consequently, these examples (and many others), point in the direction of the importance of balancing *and/or* simultaneously considering risk and ethics in decision-making. Likewise, the popular press increasingly describes "good" decision-making as aligning risk and ethical leadership/decision-making practices (Disparte, 2016). According to Disparte (2016), scandals such as Volkswagen and Deepwater Horizon could have been prevented if their organizational teams (e.g., boards), changed the way they thought about and responded to risk, and were better equipped with codes of conduct and ethical value systems. Similarly, Neal & Spetzler (2015) discuss the importance of considering an organization-wide approach to good decision-making. In their recommendations, they note that a thorough analysis of risk and organizational ethical values are two key factors in increasing the likelihood of quality decisions.

In support of my argument relative to androgynous decision-making and team performance, risk propensity and ethical responsibility *may* play an important role given they are traditionally seen in the literature as being associated with males and females at both the individual level and team level of analysis. As discussed previously, Woolley & Malone's (2011) research on group collective intelligence demonstrated that while there

was little correlation between a group's collective intelligence and individual IQs, increasing the number of women in the group lead to a rise in collective intelligence. The researchers attributed these findings to the greater number of women increasing the amount of turn taking, listening to each other, and being democratic. I argue that this *may* be the case relative to risk and ethics as well – sex balance may allow teams to engage in *both* masculine and feminine decision-making styles. Additionally, I argue that empirically investigating their links to gender and/or of organizational teams will further support the need for sex balanced teams (e.g., sex targets), and positively influence objective performance. To revisit the Lehman Brothers example from the introduction, balancing masculine (e.g., risk propensity) and feminine (e.g., ethical responsibility) decision-making styles in teams (e.g., corporate boards) may be increasingly required. While I argue that sex and gender are not necessarily interchangeable terms, perhaps a quick way to achieve androgynous decision-making of organizational teams is to ensure that they also have sex balance.

A case study to illustrate my point: In 2008, the Lehman Brothers investment bank filed for the largest bankruptcy in history and intensified the financial crisis. If we examine the background of Lehman Brothers, we see that the company was a prime example of a stereotypically *masculine* environment. For example, in 2007, the board of directors was solely *male* (11 men in total; Richardson, 2008), and the executive committee was primarily men, with one female, Erin Callan, as chief financial officer. Dick Fuld, the chief executive officer and chairman at the time, was "the poster child for the culture of excessive risk taking at Lehman (Farrell, 2013, para. 1)" and used phrases such as "the bros always wins!" in correspondence with other members of the executive team (Mahapatra, 2013, para. 3).

Acquiring five mortgage lenders, and experiencing a massive surge in revenues, Lehman Brothers continued to operate recklessly despite signs of the bubble bursting. In several instances, executives of the company acted unethically (e.g., accounting fraud), were aggressive in the market (e.g., subprime mortgages), and took dramatic risks. Considering this example from a masculine standpoint, we see that the company focused on competition, power, as well as the task/goal at hand (i.e., being the leading global financial services firm) and were perhaps too distracted by the goal to recognize their decision-making was quite risky and unethical (Turner, 2012). Consequently, some argue that the collapse *may* have had a different outcome if "Lehman brothers had been Lehman sisters" (Moss-Kanter, 2010, para. 5) due to the different traits and behaviors that women may bring to the table. In the case of Lehman Brothers Holdings Inc., some might argue that if the board had consisted of women, avoidance of *any* risky outcomes might have hindered the initial financial success of the organization, and that while executives would have collectively met to develop ideas and build consensus, the focus on participation could have stalled any decisions being made anyhow (Turner, 2012). I argue that the best alternative is not an either/or scenario, but a balance of men, women, and masculine and feminine decision-making styles (i.e., an androgynous team). Masculinity would have allowed them to engage in risk and healthy competition, but femininity would ensure that any risk taken was ethically responsible. Also allowing them to not only focus on the task at hand but also the importance of participation, collaboration, and turn taking in completing that task.

Summary

As evident in the literature review, the interaction between sex, gender, and organizational decision-making/performance is quite complex. Theoretically and empirically, there is a multitude of other individual level, team level, and organizational level literature that could have been examined regarding factors that affect team decision-making processes and gender. See Appendix A for an overall conceptual team decision-making process model. However, because my interest in this dissertation stemmed from the recent attention and/or initiatives (scholarly, politically, and socially) regarding calls for equal representation on decision-making bodies, my focus specifically hones in on gendered decision-making in conjunction with risk propensity and ethical responsibility at the team level of analysis. See Appendix B for a conceptual model of this dissertation.

Therefore, the main purpose of this program of research was to explore, develop, and refine the construct of team androgyny. Specifically, with the aims of examining the validity of the team androgyny instrument's masculinity and femininity scales, and assess the links of team gender *and* sex composition to risk propensity, ethical responsibility, and performance.

My dissertation addresses these aims in a pilot study and three main studies. Several hypotheses have been developed, which are described in detail the following chapters. Study 1 involves an extensive literature review to create an initial item pool, uses subject matter experts to confirm each item, and two exploratory principal components analyses with different samples to examine the factor structure and refine the item pool. Study 2 uses a cross-sectional design to further refine the factor structure, investigates the convergent and divergent validity of the team androgyny scales, and assesses links to performance (i.e., team project grades), risk propensity, and ethical responsibility at the team level. Finally, in Study 3, an experimental design was used to

control the sex composition of the team, and test objective performance using a decision-making scenario. These data were used to confirm and validate the factor structure, and evaluate the effects of team sex and team gendered decision-making styles (i.e., masculinity, femininity, and androgyny) on performance, risk propensity, and ethical responsibility.

2: Study 1 – Item Development, Reduction, and Factor Structure

Study 1a: Item Development

Developing a valid measure of any underlying construct is a necessary process. Many researchers agree in deductively conceptualizing a theoretical definition as well as a thorough review of relevant literature prior to item construction (e.g., Clark & Watson, 1995; Hinkin, Tracey, & Enz, 1997). Clark & Watson (1995) note that by conducting a comprehensive literature review of the target construct, researchers are better able to understand the content and most importantly, determine if the proposed scale already exists.

In discovering the scope of my content domain, I developed my instrument from the *theory* of androgyny (Bem, 1974). One of my main objectives of this dissertation was to develop a valid measure of masculine and feminine decision-making styles at the team-level that can be used by organizational researchers. Given the recent nature of team-based organizational structures as well as global calls for change for more women on decision-making bodies, I argue that creating this measure is not only timely, but allows for a conceptually consistent *and* empirically driven construct. Therefore, my main objective in Study 1a was to develop an initial item pool to capture team masculinity and femininity and have the items reviewed by subject matter experts.

Method

Participants

Participants consisted of a group of four subject-matter experts from two universities in Eastern Canada, primarily selected due to convenience. The subject matter experts were Masters or PhD students in management and/or psychology, and have had experience in studying gender and/or leadership, as well as quantitative research design.

Procedure

Using the conceptualization of team androgyny discussed previously, I developed a pool of items as the preliminary version of the scale. Once the initial pool was created, a paper-and-pencil questionnaire was provided to the subject-matter-experts, who were asked to define each item as masculine or feminine as well as note if any items were ambiguous, redundant, and/or double-barreled.

Data Analysis

Item Generation

I conducted a literature search by examining potential pre-existing published measures (individual level and/or team level) directly and indirectly related to masculinity and femininity. Spector (1992) recommends that when developing a new construct, researchers should examine and when possible, adapt psychometrically sound items from existing scales. My extensive item search strategy began in the field of gender roles (e.g., BSRI – Bem, 1974). After reviewing gender role literature, I broadened my search to include behaviors that are indirectly associated with masculinity and femininity as well as decision-making. These included risk, ethical/moral responsibility, collaboration, emotion, conflict, open communication, influence/confidence, and trust. See table 2.1 below for a description of scales/measures considered for item generation of my team androgyny instrument.

Table 2.1 Literature search of relevant measures for item generation

Construct	Measure	Citation
Gender roles	Agentic and Communal Values	Trapnell & Paulhus, 2011
	Bem Sex Role Inventory [BSRI]	Bem, 1974
	Bidimensional Impression Management Index [BIMI]	Blasberg, Rogers, & Paulhus, 2013
	Personal Attributes Questionnaire [PAQ]	Spence, Helmreich, & Stapp, 1974
Risk	General Risk Aversion Scale	Carter & Yeqing, 2005
	Risk Orientation Questionnaire	Rohrman, 1997, 2005
	Risk Propensity Scale	Meertens & Lion, 2008
Ethical/moral responsibility	Ethical Climate Questionnaire	Cullen & Victor, 2008
	Ethical Values Assessment	Padilla-Walker & Jensen, 2016
Collaboration	Assessment of Interprofessional Team Collaboration Scale [AITCS]	Orchard, King, & Bezzina, 2012
	Work Group Characteristics Measure	Campion, Medsker, & Higgs, 1993
Emotion	Team-Referent Emotional Intelligence Scale	Wei, Liu, & Allen, 2016
Conflict	Team Conflict	Jehn & Mannix, 2001
	Rahim Organizational Conflict Inventory-II	Rahim, 1983
Open communication	Team Open Communication	Barry & Stewart, 1997; Plunkett-Tost, Gino, & Larrick, 2013
	Performance Assessment Tool for Interpersonal Communication and Teamwork [PACT]	Chiu, 2014
Influence/confidence	Empowering Leadership Questionnaire	Arnold, Arad, Rhoades, & Drasgow, 2000
	Team Multifactor Leadership Questionnaire	Bass & Avolio, 1996
	Team Leadership Questionnaire [TLQ]	Moregeson, DeRue, & Karam, 2010
Trust	Formative and Reflective Indicators of Team Trust	Costa & Anderson, 2011

Item Confirmation

The items were submitted to the subject-matter experts, who were asked to review each item (which were in random order), and give their first impression as to whether each would be prominent in a masculine or feminine team, without priming them in defining masculinity or femininity.

Results

Item Generation

Once I completed my literature review of relevant scales, I developed a large preliminary item pool comprised of 64 items grouped into two scales representing masculine and feminine team-based decision-making styles. These were written in a way that individuals could respond by rating their level of agreement with a statement regarding their team. For instance, an example of a masculine item was "our team enjoys risky decisions" and a feminine item was "our team encourages open communication." Generally, these items were developed from the scales above and reworded to relate specifically to gendered decision-making styles. See Appendix C for list of items used for item confirmation.

Item Confirmation

All subject-matter experts identified each of the *64 items* as aligned with masculine or feminine decision-making styles. *One item* was perceived by two of the experts as ambiguous so that item was eliminated from the initial pool resulting in *63 items*.

Study 1b: Item Reduction and Initial Factor Structure

This study had two objectives. The first was a *pilot* test with the aim of refining my initial item pool. Worthington & Whittaker (2006) note that initial pools may include four times as many items than will be included on the final instrument, and do not

recommend testing of convergent/divergent validity at this point in scale construction. Additionally, Whetten (1989) noted that items will be refined overtime. Therefore, my aim was to eliminate psychometrically weak and repetitive items, which would lessen the likelihood of response bias due to length of my measure (Clark & Watson, 1995; Hinkin, 1995). The second and *main* objective was to conduct an exploratory principal components factor analysis to test the structure of my refined instrument.

Pilot Study

The pilot test included 90 undergraduate business and psychology students who completed the initial item pool questionnaire online or in class. The average age of participants was 23, and there was a total of 25 males (28%) and 63 females (72%) participants (2 missing).

Participants were given a questionnaire and asked to rate the likelihood of each of the 63 items occurring in a masculine or feminine team (the order was randomized but all participants rated both the masculine and feminine scales). See Appendix D for pilot study materials.

Data analysis was exploratory in nature and included a principal components analysis with varimax rotation using SPSS Version 24. Using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, which was an excellent .90, and Bartlett's Test of Sphericity, which was ($\chi^2 = 7375.\ 88$, $df = 2016$, $p < 0.001$), indicated that the correlations between items were statistically significant from zero. Initial eigenvalues indicated that the first two factors accounted for 47 percent of the variance. On the other hand, the scree plot showed the leveling off of eigenvalues after three factors, but the third factor only accounted for an additional 3 percent of the variance.

When examining the factor loading matrix, beyond two factors, there was an insufficient number of primary loadings as well as a number of cross-loadings making it difficult to interpret subsequent factors. Specifically, several communality values were below .5 indicating that these are not well captured by the component structure, thus, are grounds for removal (Meyers et al., 2013). Furthermore, several researchers note that there should be a difference of at least .2 between the primary loading and cross-loading, and any cross-loading above .3 can be eliminated from the solution (Meyer et al., 2013). As such, I argue that my results and previous theoretical support (i.e., theory of androgyny) could be used to justify a two-factor solution. See Appendix E for pilot study scree plot demonstrating the leveling off of eigenvalues.

When examining the first two factors, they represented 28 items (i.e., feminine) and 18 items (i.e., masculine) respectively. However, a number of these items were cross-loaded in addition to having low factor loadings. For several of the items, it was difficult to justify whether it primarily loaded onto one factor or another. This may have occurred in part due to the length of the measure causing a response bias, and more importantly, because these items did not accurately represent the target construct. Therefore, I eliminated items with high cross-loadings, low factor loadings and those that were redundant. The initial PCA results of my pilot test provided preliminary support for my proposed instrument. Consequently, I proceeded with data collection for Study 1b.

Method

Participants

The sample consisted of 240 undergraduate psychology and business students who completed the questionnaire online or in class. The average age of participants was 21, and there was a total of 97 males (i.e., 40.3%) and 143 females (i.e., 59.7%) – For details about the sample used in this study as well as Study 2 and 3, see Table 2.2 below.

Procedure

The online and in class survey included a letter of information, the team androgyny instrument, and a brief demographic questionnaire. In the letter of information, the purpose of the survey was described to participants as studying the gendered behaviors of organizational teams. The letter also described the informed consent, confidentiality, freedom to withdraw, risks, and my contact information. As the scales were randomly assigned, the next page was labeled either "Team Masculinity Questionnaire" *or* "Team Femininity Questionnaire." Participants were required to complete the scale twice. Once to capture the masculine items and once to capture the feminine items. After completing the survey, participants received a feedback letter explaining the nature of the study. See Appendix F for Study 1b materials.

Measures

Demographics: Demographic information including the participant's sex, age, and ethnicity, was collected.

Team Androgyny: After shortening the instrument (from the pilot study), the measure included 20 items, 10 feminine, and 10 masculine. I asked participants: "using the following scale below, please rate the likelihood of each statement occurring in a masculine/feminine organizational team." Items were rated on a 5-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Internal reliability for the 10-item feminine scale was .93, and for the 10-item masculine scale was .89.

Table 2.2 Summary of studies, their procedures, and data/sample

Study	Actions and variables	Data/sample
Study 1a	- Initial item pool = 64 items - Item pool after subject matter expert review = 63 items	*N* = 4 subject matter experts (Graduate students in management/psychology)
Study 1b – pilot	- Item reduction to 20 items using exploratory PCA	*N* = 90 undergraduate students *M* age = 23 years 72% female
Study 1b – main	- Item reduction to 17 items using exploratory PCA	*N* = 240 undergraduate business/psychology students *M* age = 21 years 59.9% women
Study 2	- Item reduction to 12 items using Exploratory structural equation model (ESEM); Convergent/divergent validity: Team androgyny (masculinity and femininity), risk propensity, ethical responsibility (benevolence, egoism, and values), collaboration, counterproductive work behaviors - Aggregation analysis; multiple hierarchical regression: Team masculinity, team femininity, team masculinity x team femininity, risk propensity, ethical responsibility (benevolence and egoism), objective performance (i.e., team grades), and controls	*N* = 238 undergraduate business/psychology students (organized into 62 teams) *M* age = 22 years 47.9% women
Study 3	- ESEM and Confirmatory factor analysis (CFA); convergent/divergent validity: Team androgyny (masculinity and femininity), communality/agenticism, risk propensity, ethical responsibility (benevolence and egoism) - Aggregation analysis; multiple hierarchical regression: Team masculinity, team femininity, team masculinity x team femininity, risk propensity, ethical responsibility (benevolence and egoism), objective performance (i.e., Winter Survival Exercise team rank), and controls	*N* = 245 undergraduate business/psychology students (organized into 53 teams) *M* age = 22.88 years 50.6% women 52% Caucasian

Data Analysis

The data was screened for appropriate descriptive statistics (i.e., Skewness and Kurtosis), inter-item correlations, and univariate and multivariate outliers (Meyers, Gran, & Guarino, 2013).

Factor Structure

To assess the initial 20-item factor structure, I conducted an exploratory principal components factor analysis with varimax rotation using SPSS Version 24. Principal components analysis (PCA) is often used to evaluate the psychometric properties of the scale(s) as well as for data reduction (Tabachnick & Fidell, 1996). As this was my first analysis with a sufficient sample size, PCA was advantageous because I was able to test the underlying factor structure of my team androgyny instrument without constraining the number of items/factors to be extracted.

In order to conduct a factor analysis, a minimum number of cases relative to number of variables must be satisfied (e.g., Rule of 10 –Everitt, 1975; Kunce, Cook, & Miller, 1975). The number of participants in my study was 240. However, there were 480 cases (because participants completed the scale twice), providing a ratio of 24 cases per variable.

I used the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's Test of Sphericity as indicators to ensure that PCA was appropriate for my dataset. Meyers et al. (2013) recommend a KMO value of at least .70 or above and a Bartlett's Test of Sphericity that is statistically significant from zero demonstrating that none of the items are significantly correlated. Additionally, a recommended criterion is that components of eigenvalues above one or more should be retained and communalities of each variable below .5 should be eliminated from the analysis (Meyers et al., 2013). Finally, varimax rotation, which is the most commonly used exploratory rotation method,

was also chosen for this analysis. Additionally, at this stage of my research, I felt that varimax rotation was appropriate given that the factors were assumed to be uncorrelated (Meyer et al., 2013).

Descriptive Statistics

I computed the relevant means, standard deviations, skewness and kurtosis, internal reliability, and bivariate correlations of the composite scores for each factor (i.e., masculine scale and feminine scale).

Results

Factor Structure

The results of the PCA with varimax rotation had a KMO measure of sampling adequacy of .96, and Bartlett's test of sphericity was statistically significant (χ^2 = 5636.28, df = 190, p < .00). Finally, the communalities were all above .3 confirming that each variable shared variance with the other variables. Therefore, conducting a PCA on all 20 items was deemed appropriate.

Initial eigenvalues produced two components greater than 1.00, cumulatively accounting for 58.8% of the total variance. Furthermore, the scree plot indicated that the eigenvalues leveled off (Neill, 2008) after two factors. See Appendix G for Study 1b scree plot results. Loadings ranged from .49 - .80 and there were a few cross-loadings – these items included "to our team, success means winning," "our team tends to use power to reach a solution," and "our team prefers to stay on the safe side when completing the task." Since these items had relatively high cross-loadings above .35, I eliminated these items from my team androgyny instrument, which resulted in a remaining 17 items. See Table 2.3 for factor loadings from PCA with varimax rotation for all 20 items:

Table 2.3
Factor loadings and communalities based on a principle components analysis with varimax rotation for 20 items from the team androgyny scale (*n* = 240)

	Factor 1	Factor 2	Communalities
7. Our team encourages open communication	.801		.700
15. We make decisions based on the interests of all parties involved	.778		.629
1. Our team openly listens to each other's issues	.770		.729
8. We consider the rights of others to arrive at a fair decision	.769		.666
13. We ensure that everyone on our team participates in decision-making	.763		.637
4. We are perceptive to each other's suggestions	.755		.643
20. Kindness is important to the success of the team	.739		.660
5. As a team, we care about each other's well-being	.727		.615
17. Before reaching a solution, our team considers all possible alternatives	.642		.473
10. Our team prefers to stay on the safe side when completing a task	.495	-.487	.481
9. Our team is willing to take risks when making decisions		.758	.609
14. Our team sees any task as a game to be won		.742	.659
3. Our team tends to use power to reach a solution	-.438	.740	.739
11. We use our competitiveness to successfully complete a task		.730	.618
6. Our team will do whatever it takes to perform well		.702	.531
19. We use our assertiveness to reach a solution		.697	.531
2. To our team, success means winning	-.398	.691	.636
12. Our team follows the motto 'nothing ventured, nothing gained' when making a decision		.642	.462
18. We avoid using our emotion when finalizing a decision		.564	.418
16. Our team tends to generate a solution as quickly as possible		.563	.329

Note. Factor loadings below .35 are suppressed

Descriptive Statistics

Once the data was analyzed using an exploratory PCA, and a few items were eliminated, composite scores were created for each of the two factors (i.e., masculine scale and feminine scale). I computed descriptive statistical analysis without the cross-loaded items, which resulted in a feminine scale of nine items, and a masculine scale of eight items. The skewness and kurtosis were within the acceptable range, the Cronbach's alphas were very good, .93 for the feminine scale, and .87 for the masculine scale.

Additionally, the bivariate correlation between the masculine and feminine scale was significantly negative (r = -.6, p = .000), but this was expected as participants were directly asked to rate items as masculine or feminine. See Table 2.4 for the descriptive statistics for the masculine and feminine scales of the team androgyny instrument.

Table 2.4
Descriptive statistics for the masculine and feminine scales (n = 240)

	No. of items	*M (SD)*	Skewness	Kurtosis	Cronbach's Alpha
Feminine	9	3.45 (.84)	-198	-.836	.927
Masculine	8	3.51 (.78)	-.218	-.717	.867

Discussion

Although several researchers have suggested that androgynous leadership may be advantageous in terms of effective decision-making (e.g., Kirchmeyer, 1996; Stephens, 2005), to date, no study has considered this phenomenon at the team level of analysis. The primary purpose of Study 1a and 1b was to begin developing and operationalizing the construct of team androgyny by creating a large item pool, refining the item pool, and testing the factor structure through exploratory means. First, a thorough review of the literature identified several potential items to be defined as masculine or feminine. Subject matter experts classified the items into the appropriate categories. Second, the factor analytical technique of principal components analysis reduced and refined the data from 63 items to 20 items loading onto two unique scales. Finally, another exploratory principal component analysis was conducted using the 20 items (and a new sample) to assess the factor structure without actually constraining the structure of the measure. Based on the conceptualization of androgyny, masculinity and femininity are two independent constructs, and my factor structure supported this notion as one factor

represented team masculine decision-making, and the other factor team feminine decision-making styles.

Results of the exploratory principal components analysis with varimax rotation in my pilot study suggested that the items in the team androgyny instrument generally assessed three factors; however, the first two factors represented over 47 percent of the explained variance, with an additional 3 percent of variance from the third factor. Several items loaded onto a single factor and had communalities above the recommended cutoff score of .5 (Worthington & Whittaker, 2006). Furthermore, although the pilot study sample size was relatively small, Velicer & Fava (1998) indicated that in sample sizes of 150 or more communalities of .5 are sufficient. As expected, the two main factors consisted of items representing either masculinity or femininity. More specifically, 28 items loaded onto the feminine factor and 18 onto the masculine factor. The remaining factors included items that were cross-loadings, redundant or had low factor loadings and eigenvalues below one (Kaiser, 1958). Consequently, these were eliminated from the pool. Because the initial instrument was so time consuming, potentially imposing a response bias in terms of scale length (Hinkin, 1995), I wanted to make the measure as short as possible without affecting internal consistency or reliability. Cook, Hepworth, Wall, & Warr (1981) noted that reliability can be reached with as few as three items, and scales with many items tend to have issues in terms of cross-loadings and tapping into more than one dimension. Acceptable justification for item reduction relates to poor conceptualization or redundancy (Hinkin, 1995). Therefore, I eliminated several other items from the feminine and masculine components based on low factor loadings and redundancies resulting in a shorter measure with ten masculine items and ten feminine items. I argue that these items are adequate in terms of content validity, and internal

consistency and reliability (Kenny, 1979) of my team androgyny instrument. Nonetheless, I felt it was important to further explore my instrument prior to examining convergent and divergent validity, and more complex experimental analysis.

As expected, a second exploratory principal components analysis in Study 1b showed that the reduced item pool generated two factors representing masculine and feminine decision-making styles of teams. There were a few cross-loadings that were a cause for concern because their factor loadings were above .35, and did not seem to load primarily on one factor or the other. The substantial cross-loadings of the items likely related to lack of clarity in the wording. For example, the first item surrounding 'staying on the safe side when reaching a decision', resulted in loadings below .5 on each of the two factors; these loadings are both quite low, as such I felt it was grounds for elimination. The second item regarding using power in reaching a solution had a relatively high primary loading on the masculine factor, but the cross-loading was above .45 on the feminine factor indicating that it may pertain to both factors (in the positive direction for masculinity and the negative direction for femininity); as such, following the recommendation of Worthington & Whittaker (2006), I deleted the item since both factors had high absolute values above .32. The third item, 'success meant winning', also had values above .32 on the masculine *and* feminine factor; and as such it was removed. The ambiguity generated from these three items may have caused confusion for some participants (Johnson, Bristol, & Schneider, 2011). Other than these three items, the remaining 17 items of the team androgyny instrument loaded onto two factors with loadings ranging from .563 - .770, demonstrating consistency and dimensionality of the component structure, as hypothesized.

Broadly speaking, the initial factor structure has several important implications. First of all, the dimensionality of the component structure showed two predominate factors, which could be categorized as masculine and feminine. Second, the reliabilities of both the femininity and masculinity subscales of the team androgyny instrument were high, showing support for internal consistency. Researchers such as Bem (1974) and Spence & Helmreich (1974) argued that gender roles are not opposing constructs but rather independent dimensions that can occur within the same individual; As such, these should be separate and orthogonal. The negative correlation between the two subscales can be attributed to explicitly asking participants to rank the likelihood of each item occurring in a masculine *versus* feminine organizational team. Therefore, participants may have interpreted the need to trade off; as femininity increased, masculinity decreased and vice versa. Overall, the team androgyny instrument was shown to be reliable after refinement of the initial item pool.

Limitations and Future Research

Despite some methodological strengths of Study 1a and 1b (e.g., item generation using subject matter experts, two exploratory studies with different samples), there are also a few limitations. First of all, the pilot study consisted of only 90 participants (who completed the scale twice – 180 responses), with a much higher proportion of female respondents, which could have initially biased the feminine versus masculine items. However, randomizing the order of the scales may have reduced this limitation. Additionally, the relatively small sample size *may* have had an effect on statistical power. However, these concerns were mitigated in Study 1b with a larger sample size (240 respondents with 480 cases) that was more representative in relation to sex of participants. Additionally, statistical power was unlikely of concern because in both the

pilot study and in Study 1b, the factor structure was as hypothesized. Future research could achieve higher statistical power by collecting data from a greater number of respondents.

Second, the generalizability of the findings may be limited given that the data was collected from one university in both the pilot study and the main study. Future research could attain higher generalizability by collecting data from more diverse samples, not only covering other universities, but also data from a diverse population of organizations.

Finally, one main concern that required further investigation was the question of whether the scales being highly negatively correlated was due to asking participants to rate both feminine versus masculine items in an effort to ensure an adequate response rate. Similarly, *individual* students were asked to rate items, as opposed to students in actual teams, although, given that the majority of students were in psychology and business, many of them participated in teams in their classes. It is also noteworthy that Study 1's purpose was to evaluate and optimize the measure prior to more complex analysis; By using subject matter experts, and conducting two exploratory studies, this was accomplished. Overall, Study 1 demonstrated initial support for my team androgyny instrument, but further investigation of the measure was necessary to continue to examine the structure, correlations between the scales using actual teams (without disclosing the gendered nature of the measure), and most importantly, validating the feminine and masculine subscales of team androgyny.

3: Study 2 – Instrument Validation and Testing

The first objective of Study 2 was to further test and refine the factor structure of my team androgyny instrument, as well as its convergent and divergent validity in correlating with risk-taking propensity, ethical benevolence, egoism, and values, collaboration, and counterproductive work behaviors. Based on the Bem's (1974) theoretical conceptualization of androgyny as well as my extension of masculinity and femininity to the team level (and findings from Study 1), I expected a 2-factor solution (i.e., a masculine scale and a feminine scale) that would be moderately and positively correlated. Bem (1974) herself noted that masculinity and femininity were meant to compliment rather than contradict one another. Additionally, based on my earlier theoretical arguments, I also aimed to compare my team androgyny instrument with existing psychometric measures that were indirectly associated with masculinity and femininity.

The following hypotheses were developed regarding the factor structure and convergent/divergent validity of the team androgyny scale:

Hypothesis 1a: Team androgyny will represent a two-factor solution, one factor representing masculinity, and one factor representing femininity.

Hypothesis 1b: Masculinity and femininity will be moderately and positively correlated.

Hypothesis 2a: A moderate negative correlation between femininity scores and ethical egoism.

Hypothesis 2b: A weak positive relationship between masculinity scores and ethical egoism.

Hypothesis 3a: A moderate positive correlation between femininity scores and ethical benevolence.

Hypothesis 3b: A moderate negative relationship between masculinity scores and ethical benevolence.

Hypothesis 4a: A strong positive correlation between femininity scores and ethical values.

Hypothesis 4b: A weak negative correlation between masculinity scores and ethical values.

Hypothesis 5a: A moderate positive correlation between masculinity scores and risk propensity.

Hypothesis 5b: A weak negative relationship between femininity and risk propensity.

Hypothesis 6a: A strong positive correlation between femininity scores and collaboration (i.e., work group participation).

Hypothesis 6b: A non-significant relationship between masculinity scores and collaboration (i.e., work group participation).

Hypothesis 7: A non-significant relationship between masculinity and femininity scores, and counterproductive work behavior.

Once evidence of convergent/divergent validity was supported, my second purpose of this study was to partially examine my conceptual model by empirically testing my team androgyny instrument for its usefulness as a predictor of risk propensity, ethical responsibility, and objective performance at the *aggregated level of analysis.* I must stress that this study is considered a *pilot* because I was unable to include sex composition of the team as a control (data was collected from student teams at the end of the semester). The five hypotheses investigated in this study were:

Hypothesis 8a: Team masculinity will be positively related to risk propensity in comparison to team femininity, which will be negatively related.

Hypothesis 8b: Team femininity will be positively related to ethical benevolence in comparison to team masculinity, which will be negatively related.

Hypothesis 8c: Team masculinity will be positively related to ethical egoism in comparison to team femininity, which will be negatively related.

Hypothesis 8d: Team androgyny will be positively related to both risk propensity and ethical benevolence.

Hypothesis 8e: Team androgyny will be a better predictor of objective performance outcomes (i.e., have higher team grades on course project) than team masculinity or team femininity.

Method

Participants

The sample consisted of 238 undergraduate business students organized into 62 teams (two – seven members). These teams were required to work together for the entire semester. Specifically, classes included organizational behavior, marketing, strategic management, and human resources. Participants were comprised of 47.9 percent female and 52.1 percent male with an average age of 22 years. On average, teams consisted of two to nine members and had worked together on a team presentation and/or written project until the end of the term.

Procedure

The in-class paper-and-pencil survey was administered at the end of the fall semester of 2016 and included a letter of information, the team androgyny instrument, other measures (see below) to test convergent and divergent validity, and a brief demographic questionnaire. In the letter of information, the purpose of the survey was described to participants as assessing the decision-making styles of their team; the letter also stated that their team grade would be collected after submission to the registrar. Participants were then asked to individually complete all measures on the questionnaire while thinking about their team. Team grades were collected from professors two months after survey administration to ensure that they had been finalized at the registrar. See Appendix H for Study 2 survey materials.

Measures

Demographics: Demographic information including the participant's sex, age, ethnicity, and GPA was collected. Additionally, the participant's team ID number and final team project grade was collected for data analysis purposes.

Team Androgyny: After the removal of inconsistent items (i.e., removed due to concerns of face validity as some participants questioned how they related to team decision-making), my team androgyny scale consisted of 12 items – 6 masculine and 6 feminine. Sample items included "we were perceptive to each other's suggestions" and "our team encouraged open communication (feminine);" "our team was willing to take risks," and "we used our assertiveness to reach a solution (masculine)." Items were rated on a five-point Likert scale from 1 – strongly disagree to 5 – strongly agree. The instrument showed acceptable reliability for the femininity scale $\alpha = .79$, and the masculinity scale $\alpha = .73$.

Ethical Responsibility: The ethical responsibility portion of the questionnaire included two different measures. First, the Ethical Climate Questionnaire developed by Victor & Cullen (1987) and later modified by Cullen, Victor, & Bronson (1993) was used to assess different aspects of ethical climates. For the purpose of this dissertation, I used the subscales of egoism (3 items) and benevolence (3 items) because they represent Kohlberg's (1989) ethical standards of how organizations can make decisions morally. Egoism relates to efficiency, profits, and performance whereas benevolence relates to maximizing joint efforts and being morally responsible within a company. However, the wording of the items was modified from "company" to "team." Items included "in our team, people are mostly concerned for themselves," and "in our team, our major concern is what is fair for everyone" and were rated on a five-point Likert scale from 1 – mostly false to 5 – completely true. This scale displayed acceptable reliability for the egoism subscale $\alpha = .70$ and good reliability for the benevolence subscale $\alpha = .81$.

I also used an adapted version of the Ethical Values Assessment, which taps into moral psychology related to the ethics of autonomy, community, and divinity (Padilla-Walker & Jensen, 2016). While this is quite a new measure, the ethical autonomy subscale was appropriate because it focused on taking responsibility, being respectful, and achieving goals and had good internal reliability ($\alpha = .82$) in the study by Padilla-Walker & Jensen (2016). Four items were included and were adapted by replacing "I" with "our team" or "we." An example of one of the items is as follows, "our team takes responsibility for each other." The items were rated on a five-point Likert scale from 1 – not at all importance to 5 – completely important. This adapted scale showed acceptable reliability, $\alpha = .76$.

Risk Propensity: Risk propensity included a measure by Meertens & Lion (2008) called the Risk Propensity Scale adapted to the team level. The Risk Propensity Scale included nine items that tap into different aspects of risk taking. For the purpose of this dissertation, I included the three items that examined the general tendency to take risk (i.e., "I take risks regularly" was adapted to "Our team took risks"). The other items were eliminated because they did not relate to organizational risk taking (e.g., "I do not take risks with my health). Items were rated on a nine-point Likert Scale from 1 – totally disagree to 9 – totally agree. This adapted scale showed good reliability, $\alpha = .80$.

Collaboration: The degree of participation on the teams was measured using an instrument by Campion, Medskey, & Higgs (1993) called the Work Group Characteristics Measure. The measure consists of 19 subscales focusing on the functioning of work groups within organizational settings. I included one subscale since other subscales focused on factors unrelated to this dissertation (e.g., training, managerial support, task variety, task significance, and task identity). I also adapted the item wording from "work group" to "team." The participation scale included three items (e.g., "my team was designed to let everyone participate in decision-making). All items were rated on a five-point Likert scale from 1 – Strongly Disagree to 5 – Strongly Agree. The scale displayed a good internal reliability of $\alpha = .80$.

Counter Productive Work Behaviors: For the purpose of divergent validity, I included an adapted version of the abuse subscale of the Counter Productive Work Behaviors checklist by Spector, Fox, Penney, Bruursema, Goh, & Kessler (2006). The abuse subscale consists of nine items such as "started harmful rumors," "been nasty or rude to each other," and "blamed each other for mistakes." Items were rated on a five-point

Likert scale from 1 – strongly disagree to 5 – strongly agree. The scale displayed acceptable reliability of $\alpha = .74$.

Team Performance: Each student team completed a presentation/project at the end of the term relevant to their course. Grades were only assigned one per team; as such this was a team-level construct. Grades ranged from 62 to 95 with a mean of 82.34 and a standard deviation of 7.13 ($n = 230$).

Control Variables: I included specific control variables in data analysis: Firstly, I included biological sex because I was unable to randomize the sex of each team; as such, the sex of each team could have influenced their gendered decision-making styles. Biological sex was coded (1 = female; 0 = male). Secondly, ethnicity was considered a dummy control variable due to the ethnic diversity of my teams as well as past research indicating that cultural variables may have an effect on the way participants perceived the questionnaire items (Schaffer & Riordan, 2003). Third, individual grade point average (GPA) was captured because research has suggested that prior individual performance (i.e., GPA) may be related to team performance (LePine, 2003). GPA was captured on a self-report from 0 – 4.3. Fourth, because data was collected from two separate universities, I considered this a potential control variable and dummy coded 1 = University A, 0 = University B. Most data were collected from one university; therefore, this was used as the reference category in my analysis. Finally, I included team size as a control variable because teams were of various sizes ranging from two members to nine members. Each of these control variables were included to lessen the likelihood of biasing my results. More importantly, I only included control variables that were conceptually meaningful to my overall research purpose, as each control variable is based on past

literature related to student team performance (Becker, Atinc, Breaugh, Carlson, Edwards, & Spector, 2016).

Data Analysis

Data analysis for Study 2 consisted of several stages including assessing the factor structure, convergent and divergent validity, data aggregation, and hypotheses testing. All data was screened to assess missing data, skewness and kurtosis, and univariate and multivariate outliers (Meyer et al., 2013) prior to conducting further analysis.

Factor Structure

To assess the factor structure, I was originally going to conduct a Confirmatory Factor Analysis (CFA); however, because participants had questions regarding some items, and because some items in my previous study had cross-loadings, I concluded that the team androgyny instrument was still in the exploratory phase. Therefore, the best approach to testing my factor structure at this phase of research was through exploratory structural equation modeling (ESEM). ESEM combines elements of both exploratory *and* confirmatory methods (Kelloway, 2015). Browne (2001) noted:

> "Confirmatory factor analysis procedures are often used for exploratory purposes. Frequently a confirmatory factor analysis, with pre-specified loadings is rejected and a sequence of modifications of the model is carried out in an attempt to improve fit. The procedure than becomes exploratory rather than confirmatory… (p. 113)"

Additionally, when uncertain about a model, conducting a CFA that results in poor fit indices might lead to unnecessary model modifications. Since many researchers advise not to conduct both an exploratory and confirmatory factor analysis on the same data set, using ESEM can be conducted instead of a CFA (Asparouhov & Muthén, 2008).

Another benefit in choosing an ESEM is that because CFA requires zero cross-loadings, it often leads to a poor-fitting model. In psychological research, measures are likely to have many small cross-loadings; therefore, such measures have appropriate EFA structures, but are unsupported in CFA models (Marsh et al., 2009). ESEM will allow less restriction in my model, due to accounting for small secondary loadings.

Much like CFA, goodness of fit is assessed using the root mean square error of approximation (RMSEA), the chi-square test statistic, the Tucker Lewis index (TLI), and the comparative fit index (CFI). For the RMSEA, values less than .08 are deemed acceptable and values less than .05 indicate good fit (Meyers et al., 2013). Likewise, the TLI and CFI values of at least .90 and .95, indicate an acceptable and excellent fit to the data (Marsh, Hau, & Wen, 2004). Finally, researchers recommend a nonsignificant chi-square; however, as sample size increases, power increases, and so too does the likelihood of statistical significance (Kenny & McCoach, 2003).

Convergent and Divergent Validity

I took two approaches in measuring the construct validity of my masculinity and femininity scales of the team androgyny instrument. First, convergent validity is tested using measures that should be related and correlated to the construct (Churchill, 1979). Second, divergent validity, is conducted by demonstrating non-relationships with measures that should not be related. To test both the convergent and divergent validity, I included a variety of measures that should be related and/or unrelated with my masculine and feminine scales, these included: ethical benevolence, ethical egoism, ethical values, risk propensity, collaboration, and counterproductive work behaviors. Additionally, conducting an ESEM allowed me to demonstrate that all scales loaded onto the hypothesized factors. Finally, I should note that the strength of the relationship between

correlation coefficients is also important. Meyers et al (2013) state that correlations of .5 or more indicate a strong relationship, .3 a moderate relationship, and .1 a weak relationship.

Data Aggregation

In order to justify aggregating my predictor variables to the team-level of analysis, I conducted specific statistical analysis as recommended in the literature (e.g., van Mierlo, Vermunt, & Rutte, 2009) on the feminine, masculine, risk propensity, ethical benevolence and egoism scales. I did not have to aggregate team performance (i.e., project/presentation grade) as only one grade was provided per team. Broadly speaking, when undertaking multilevel modeling *or* data aggregation, two indices should be considered: the interrater reliability (IRR) and the interrater agreement (IRA). The IRR can be defined as "the relative consistency in ratings provided by multiple judges of multiple targets" (LeBreton & Senter, 2008, p. 816) whereas the IRA refers to "the absolute consensus in scores furnished by multiple judges for one or more targets" (p. 816). While both indices represent agreement on ratings, choosing which to employ is dependent on the theoretical nature of the construct in question. Chan (1998) discusses four different composition models (i.e., level of construct) inherent in different levels of analysis, deciding which composition model is relevant is the first step in data aggregation.

My feminine and masculine scales are classified as *direct consensus constructs* whereby the measures were collected at the individual level, but the constructs represent the team level of analysis. More specifically, for example, feminine and masculine gendered decision-making styles of teams refer to the shared agreement among

individuals within each team. Therefore, when using a direct consensus model, IRA or within-group agreement, is a viable choice to decide if aggregation can be computed.

To test within-group agreement, I used the $r_{wg(j)}$ index (see James, Demaree, & Wolf, 1984; 1993), which tests the extent that members within a team agree on ratings of a particular construct. High within-group agreement represents consensus, and thus, justifies aggregation. Regarding cutoff values of the $r_{wg(j)}$ index, in range values are from 0 – 1, with median values above .70 indicating higher within-group agreement (LeBreton & Senter, 2008). To estimate $r_{wg(j)}$, I used an Excel tool for computing IRA developed by Biermann, Cole, & Voelpel (2012).

While within-group agreement is a valid choice for direct consensus models, it can be used in conjunction with IRR calculations (i.e., IRR + IRA), specifically, intra-class correlation coefficients. Broadly speaking, the intraclass correlation coefficient (ICC) tests the consistency of responses of members within the same team (Bliese, 2000). Calculating the ICC based on a one-way random effects ANOVA is the most common method employed. While there are numerous versions of ICC, when conducting team research, ICC (1) and ICC (*K*) are of particular importance (Bliese, 2000). According to LeBreton & Senter (2008), ICC (1) considers the level of consensus and consistency of a randomly drawn judge's score in comparison to the mean score within a randomly drawn group. On the contrary, ICC (*K*) tests the measurement reliability of group means scores in a sample (Bliese, 2000; Castro, 2002). Essentially, ICC(*K*) allows a researcher to draw inferences relative to the reliability of mean ratings of the group. Simply put, "ICC (1) informs a researcher as to whether judges ratings are affected by group membership whereas ICC(*K*) tells him or her how reliably the mean ratings distinguish between groups" (LeBreton & Senter, 2008, p. 834). Values for ICC (1) are similar to

recommendations for interpreting effect sizes, whereby a value of .01 is small effect, .05 small to medium effect, .10 a medium effect, and .25 a large effect (LeBreton & Senter, 2008). Likewise, the minimum acceptable level of reliability for ICC(*K*) should be .70 for psychometric measures in the early stages of development (LeBreton & Senter, 2008). To compute ICC (1) and ICC (*K*), I used the statistical software package SPSS by inputting syntax developed by LeBreton & Senter (2008, p. 844).

Descriptive Statistics and Correlations

Once the predictor variables (i.e., masculinity and femininity) were aggregated to the team-level of analysis, I computed the means, standard deviations, as well as the bivariate correlations prior to hypothesis testing.

Hypothesis Testing

Study hypotheses for the second objective of Study 2 were conducted in the final phase of data analysis. Specifically, data was analyzed in SPSS Version 24 using four independent multiple hierarchical regressions. Additionally, to be a thorough as possible in my data analysis, I also conducted the median split method. I must stress that I only included this technique in my analysis as it was recommended by Bem (1974) and Spence et al., (1974) as a previous way to score androgyny (prior to the 1990s). Since that time, there have a been a number of methodological concerns with the technique (e.g., generalizability, Type II error, having to dichotomize continuous variables into categorical – Streiner, 2002). These results can be found in Appendix I, as they are not a part of my main analysis.

When testing each hypothesis relevant to my outcome variables of risk taking propensity, ethical egoism, ethical benevolence, and team performance (i.e., hypothesis 8a – 8e), I regressed the control variables in the first block, including gender, university,

ethnicity, GPA, and team size. In the second block, I inserted the aggregated scores for masculinity and femininity and in step three I input the interaction between masculinity and femininity to be representative of androgyny. As noted by several past researchers (i.e., Burn, O'Neil, & Nederend, 1996; Hall & Taylor, 1985; Holmbeck, 1989; Lubinski, Tellegen, & Butcher, 1981), including Sandra Bem herself, scoring procedures in androgyny research likely results in loss of data due to median-split classification procedures. Therefore, one solution that has been identified is through the use of multiple regression whereby androgyny is the "product of the subject's femininity and masculinity scores" (Burns et al., 1996, p. 5). Using an interaction effect allowed me to test the emergent properties of utilizing both styles (Hall & Taylor, 1985) in accordance with my definition of team androgyny (i.e., the extent to which teams perceive they are using both masculine and female decision-making styles). Please note that my predictor variables were grand-mean centered before creating my interaction term to ensure proper computation of my main effects (Bauer & Curran, 2005).

Results

Factor Structure

An ESEM was conducted to test whether the data fit my hypothesized model. Prior to analysis, I reviewed study one results and removed items that were redundant, unclear, and/or unrelated to decision-making (e.g., "Our team will do whatever it takes to perform well," "We avoided emotion when finalizing our decision," "To our team success meant winning," "As a team, we care about each other's well-being"). The final instrument included twelve items, six masculine and six feminine. Based on the exploratory results in study 1, I proposed a 2-factor solution consisting of a masculine and feminine scale.

As hypothesized (hypothesis 1a), results of the ESEM using Mplus7 indicated that the 2-factor model had an excellent fit. The RMSEA was .03, the CFI and TLI were above .95, and the chi square test statistic was non-significant ($\chi^2 = 53.64$, $p = .13$). Overall, the results demonstrate that the team androgyny can be conceived as consisting of two factors, masculine and feminine.

Convergent and Divergent Validity

For comparison purposes, Table 3.1 contains the coefficient correlation between the masculinity and femininity subscales of the team androgyny instrument and the important convergent/divergent correlates at the individual level of analysis.

As hypothesized (hypothesis 3a and 2a), the results indicate that the feminine scale was shown to have a significant positive correlation with the Ethical Climate Benevolent subscale ($r = .453$, $p < .000$) and a significant negative correlation with the Ethical Climate Egoism subscale ($r = -.421$, $p < .000$). Additionally, the masculine scale was unrelated to the Ethical Climate Benevolent subscale ($r = .072$, $p = .266$), as hypothesized (3b). However, I hypothesized the masculinity would be moderately correlated with Ethical Climate Egoism, but the two were unrelated ($r = -.039$, $p = .552$), as such, hypothesis 2b failed to be supported. With regards to Ethical Values, the feminine scale was significantly and strongly positively correlated to the Ethical Values Assessment scale ($r = .603$, $p < .000$), as expected (hypothesis 4a). Contrary to my hypothesis, the masculine subscale had a significant weak, but positive correlation with the Ethical Values Assessment ($r = .232$, $p = .000$), thus hypothesis 4b was unsupported.

As hypothesized (hypothesis 5a), correlations between the masculine scale and Risk Propensity scale revealed a moderately significant and positive correlation ($r = .339$,

$p < .000$); however, risk propensity was unrelated to the feminine scale ($r = .106$, $p = .103$), thus hypothesis 5b was unsupported.

Additionally, I hypothesized that femininity would be positively associated with Work Group Participation, and masculinity would be uncorrelated with participation. As hypothesized, femininity was significantly correlated with participation ($r = .459$, $p < .000$) whereas masculinity was uncorrelated with participation ($r = .069$, $p = .289$), thus confirming hypotheses 6a and 6b.

Finally, the Counterproductive Work Behavior Checklist was included as an additional measure to test divergent validity. While unrelated to masculinity, as hypothesized ($r = -.017$, $p = .795$), the scale was significantly and negatively correlated with femininity ($r = -.351$, $p < .000$), as such hypothesis 7 was partially supported.

Table 3.1
Overall bivariate correlations between masculinity, femininity, and all other variables at the *individual level of analysis*

Scale	*M*	*SD*	1	2	3	4	5	6	7	8
1. Team Androgyny- Masculine	3.37	.640	(.73)	.329**	-.039	.072	.232**	.339**	.069	-.017
2. Team Androgyny – Feminine	4.23	.546		(.79)	.421**	.453**	.603**	.106	.459**	-.351**
3. ECQ – Egoism	1.88	.791			(.70)	.264**	.294**	-.020	.337**	.278**
4. ECQ – Benevolent	3.76	.875				(.81)	.512**	.127*	.370**	.168*
5. EVA – Autonomy	4.18	.594					(.76)	.038	.463**	.263*
6. Risk Propensity	5.24	1.74						(.80)	.008	.111
7. WGC – Participation	4.23	.638							(.80)	.171**
8. CPWB-C – Abuse	1.08	.157								(.74)

Note. Reliabilities for each scale are presented on the diagonal in parentheses.
** Correlation is significant at the .01 level (2-tailed). *Correlation is significant at the .05 level (2 – tailed).

Data Aggregation

Data aggregation statistics for the feminine, masculine, risk propensity, ethical benevolence and egoism scales are shown in Table 3.2. The $r_{wg(j)}$, ICC (1), and ICC (*K*) all had acceptable values indicating justification for data aggregation.

The median $r_{wg(j)}$ for the feminine scale was .95, demonstrating a very strong within group agreement. ICC (1) was .37, indicating a large effect, and ICC (*K*) was .78, with the *F* statistic being significant as recommended, F [229, 1150] = 4.48, $p < .05$. Regarding the masculine scale, the median $r_{wg(j)}$ was .84, ICC (1) was .30, and ICC (*K*) was .72 with a significant *F* statistic, F [227, 1140] = 3.5, $p < .05$.

Risk propensity had a median $r_{wg(j)} = .77$, with an ICC (1) of .56, and ICC (*K*) of .81) and a significant *F* statistic, F (226, 454) = 5.271, $p < .05$. Likewise, ethical benevolence has a median $r_{wg(j)}$ of .86, indicating a high within group agreement. ICC (1) was .57, and ICC (*K*) was .80, F (224, 450) = 4.909, $p < .05$. Finally, for the ethical egoism scale, the median $r_{wg(j)}$ was .82, with an ICC (1) of .40, ICC (*K*) of .67, and a significant F statistics of F (226, 454) = 3.046, $p < .05$.

Table 3.2
Data Aggregation Statistics

	Feminine Scale	Masculine Scale	Risk Propensity	Ethical Benevolence	Ethical Egoism
Median $r_{wg(j)}$	.95	.84	.77	.86	.82
ICC (1)	.37	.30	.56	.57	.40
ICC (*K*)	.78	.72	.81	.80	.67
F	4.48***	3.465***	5.27***	4.909***	3.046***

Note. n = 230 participants [1](nested in 62 teams). ***p <.000.

Descriptive Statistics and Correlations

[1] A total of eight participants were eliminated from data aggregation and hypothesis testing because they were the only members of their team to complete the questionnaire.

Table 3.3 presents the means, standard deviations, and bivariate correlations for all variables used in hypothesis testing (i.e., control variables and *aggregated* predictor variables).

Table 3.3
Overall bivariate correlations between control variables and *aggregated* variables

			Correlations										
	M	*SD*	1	2	3	4	5	6	7	8	9	10	11
1. Team Androgyny – Feminine	4.22	.36	1	.268**	-.533**	-.472**	-.008	-.037	.021	.008	-.08	-.102	.069
2. Team Androgyny – Masculine	3.38	.38		1	.018	-.047	.403**	-.053	.101	-.042	-.021	.061	-.029
3. ECQ – Egoism	1.9	.54			1	-.424**	.041	-.007	.038	.053	.064	-.200**	-.005
4. ECQ – Benevolence	3.77	.52				1	-.016	-.040	.039	-.073	-.086	.112	.048
5. Risk Propensity	5.23	1.08					1	-.05	.039	-.201**	-.057	.375**	-.224**
6. Team Performance	82.34	7.2						1	.022	-.176**	.057	-.074	.128
7. Sex	.49	.501							1	-.004	.076	-.165*	.146*
8. Ethnicity	2.33	2.87								1	-.106	-.204**	-.029
9. GPA	3.05	.60									1	-.012	-.085
10. University	.169	.376										1	-.380**
11. Team Size	4.24	1.599											1

**Correlation is significant at the .01 level (2 – tailed). * Correlation is significant at the .05 level (2 – tailed).

Hypotheses Testing

Hierarchical Multiple Regression Analyses[2][3][4]

Risk Propensity

Results of hypothesis 8a regarding risk propensity and team masculinity can be seen in the first column of Table 3.4. Gender, ethnicity, university, GPA, and team size, which were entered on the first step, was statistically significant, F (12, 201) = 4.830, p < .001 and accounted for approximately 22 percent of the variance of risk propensity (R^2 = .224). Specifically, university resulted in moderately higher risk propensity (β = .292, p < .001), teams with five members (β = -.158, p < .05) were less likely to use risk in comparison to teams with four members, and ethnicity (i.e., being black) positively predicted risk (β = .128, p < .05).

The second block, which contained the masculine and feminine team scores also showed statistical significance, F (14, 199) = 8.022, p < .001, with a ΔR^2 of .137. Specifically, the second block showed a significant positive relationship between masculinity and risk propensity (β = .408, p < .001), over and above the effects of the control variables. In contrast, femininity negatively related to risk propensity albeit not significantly (β = -.113, p = .085). Thus, hypothesis 8a, which stated that masculine teams would be significantly more likely to take risks than feminine teams, was supported.

---

[2] Please note that I also performed hierarchical multiple regression analyses without the control variable ethnicity (i.e., dummy coded). No new main effects emerged.

[3] As recommended by Becker et al. (2016), I also repeated all hypothesis tests using standard multiple regressions (i.e., inputting all variables in the same block), the results were largely identical to those presented in the main analysis. I argue that using hierarchical multiple regression was necessary in my analysis because of the important potential effects of control variables, the main effects of masculinity and femininity, as well as inclusion of my interaction term (Meyers et al., 2013).

[4] To ensure there were no issues of multicollinearity, I computed the variance inflation factors (VIFs). The recommended rule of thumb for VIF is that if it exceeds 10 than multicollinearity is too high. In my regressions, all VIF values were less than two indicating that multicollinearity was not an issue in my analysis.

Finally, the third block was statistically significant, F (15, 198) = 6.244, $p < .001$, but only accounted for .4 percent of the variance, over and above the effects of the control variables and masculinity/femininity, separately. Specifically, the interaction term between masculinity and femininity was not significant (β = .073, $p = .263$).

Ethical Responsibility

Results of the hypothesis pertaining to ethical benevolence and femininity are depicted in the second column of Table 3.4. After inputting the same control variables in the first block, which were moderately significant, F (12, 201) = 1.992, $p < .05$, $R^2 = .106$ (i.e., the covariate university and ethnicity), the second block, which contained the masculine and feminine scores, also showed statistical significance F (14, 199) = 7.124, $p < .001$, and increased substantially in its predictive power, $\Delta R^2 = .334$. Specifically, femininity was significantly and positively related to ethical benevolence (β = .548, $p < .001$) whereas masculinity was significantly and negatively related ethical benevolence (β = -.209, $p < .001$). As such, hypothesis 8b, regarding femininity and ethical responsibility was supported.

The final block was also significant, F (15, 198) = 6.677, $p < .001$, but only increased slightly in predictive power with a $\Delta R^2 = .002$. Importantly, the interaction term between masculinity and femininity did not significantly contribute to ethical benevolence (β = -.052, $p = .433$).

Results of the hypothesis related to ethical egoism and its effects on masculinity and femininity are located in the third column of Table 3.4. The first block, which included the control variables were significant F (12, 201) = 2.850, $p < .01$, and accounted for 14.5 percent of the initial variance. Namely, the control variable university,

was significantly and negatively related to egoism (β = -.248, $p < .01$); on the other hand, ethnicity (i.e., being Asian) (β = .175, $p < .05$*)* and team size (i.e., two members) (β = .224, $p < .01$) led to a significant positive effect on ethical egoism. All other covariates were not significant. When masculinity and femininity were entered on the second block, the prediction model was also statistically significant, F (14, 1989) = 10.583, $p < .001$. In the second block, feminine teams significantly and negatively related to ethical egoism (β = -.608, $p < .001$) whereas masculine teams were significantly and positively related to egoism (β = .239, $p < .001$), with a ΔR^2 of .281. The third block, while statistically significant F (15, 198) = 9.836, $p < .001$, did not increase in predictive power, $\Delta R^2 = .000$. The masculinity and femininity interaction term was not significant (β = -.016, $p = .793$), over and above the effects of the covariates and the main effects of masculinity and femininity. In sum, this analysis provides support for the study hypothesis (8c) pertaining to gender and ethical egoism.

Team Performance

Results of the hypothesis pertaining to team performance are depicted in the final column of Table 3.4. The control variables, entered on the first block were statistically significant, F (12, 201) = 3.014, $p < .001$, and accounted for 15% of the predictive variance. Namely, team size (i.e., five members) had a significant and positive effect on team performance (β = .236, $p < .01$), and ethnicity had a moderately significant and negative effect (i.e., Asian, β = -.179, $p < .05$; or Arabic β = -.174, $p < .05$). When masculinity and femininity were entered on the second block, the prediction model was statistically significant, F (14, 199) = 2.568, $p < .01$; However, it was not significant over and above the effects of the covariates – masculinity (β = .025, $p = .728$) and femininity

(β = -.003, *p* = .972). Finally, the third block, which contained the masculinity and femininity interaction term, also did not contribute to higher team performance (β = .042, p = .576) after considering the control variables, and main effects of masculinity and femininity. However, the regression coefficient for the interaction effect (i.e., androgyny) in step three of the regression was in the hypothesized positive direction.

Table 3.4

Multiple Hierarchical Regression Analysis of Aggregated Variables

	Risk Propensity			Ethical Benevolence			Ethical Egoism			Team Performance		
	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β
Constant	5.274	.397		4.196	.200		1.654	.207		81.93	2.689	
Sex	.272	.141	.124*	.063	.071	.061	-.002	.073	-.002	-.382	.954	-.027
University	.828	.213	.292**	.243	.107	.182*	-.349	.111	-.248**	-1.90	1.442	-.103
GPA	-.077	.118	-.042	-.106	.060	-.123	.053	.061	.059	.317	.800	.027
Black	.548	.279	.128*	-.365	.141	-.181**	.011	.145	.005	-2.82	1.90	-.102
Asian	-.085	.171	-.033	-.115	.086	-.096	.221	.089	.175**	-3.0	1.16	-.179**
Arab	-.059	.291	-.013	.026	.147	.012	-.098	.151	-.045	-4.99	1.97	-.174**
Multiracial	-.064	.599	-.007	-.056	.302	-.013	-.377	.311	-.082	-2.58	4.05	-.043
Team3	-.105	.181	-.041	-.176	.092	-.145	.110	.094	.086	-1.19	1.23	-.071
Team5	-.610	.258	-.158*	-.120	.130	-.067	.197	.134	.103	5.86	1.75	.236**
Team6	-.304	.213	-.099	-.181	.108	-.126	.209	.111	.138	2.71	1.44	.136
Team7	-.228	.440	-.034	-.077	.222	-.025	-.100	.229	-.030	3.66	2.98	.085
Team2	.496	.263	.140	-.420	.132	-.253**	.392	.137	.224**	2.24	1.78	.098
R^2		.224			.106			.145			.152	
F		4.830			1.992			2.850			3.014	
Constant	5.407	.369		3.928	.177		1.968	.173		81.95	2.75	
Sex	.174	.130	.079	.067	.062	.065	-.008	.061	-.007	-.424	.966	-.030
University	.661	.196	.233**	.321	.094	.240**	-.440	.092	-.313**	-.195	1.46	-.106
GPA	-.089	.109	-.049	-.058	.052	-.067	-.003	.051	-.004	.318	.809	.027
Black	.215	.260	.050	-.324	.125	-.161**	-.040	.122	-.019	-2.96	1.94	-.107
Asian	-.057	.158	-.022	-.032	.076	-.026	.124	.074	.098	-2.94	1.18	-.178**
Arab	-.056	.266	-.013	-.035	.128	-.017	-.027	.125	-.012	-.4.99	1.98	-.174**
Multiracial	.135	.549	.015	-.263	.263	-.060	-.135	.257	-.029	-2.52	4.09	-.042
Team3	-.128	.168	-.049	-.072	.080	-.059	-.013	.079	-.010	-1.19	1.25	-.071
Team5	-.530	.237	-.138*	-.050	.114	-.028	.115	.111	.060	5.92	1.77	.237**
Team6	-.382	.205	-.124	.071	.098	.049	-.086	.096	-.057	2.71	1.53	.136
Team7	-.017	.402	-.003	-.133	.193	-.043	-.032	.189	-.010	3.75	3.00	.087
Team2	.620	.248	.175*	-.224	.119	-.135	.165	.116	.094	2.32	1.85	.101
Feminine	-.346	.200	-.113	.787	.096	.548**	-.922	.094	-.608**	-.052	1.49	-.033
Masculine	1.159	.178	.408**		-.279	.085	-.209**	.083	.239**	.461	1.324	.025
R^2		.361			.334			.427			.153	
F		8.022			7.124			10.583			2.568	
Constant	5.60	.371		3.911	.178		1.962	.175		82.15	1.85	
Sex	.167	.130	.076	.070	.062	.067	-.007	.061	-.007	-.450	.969	-.032
University	.665	.196	.234**	.319	.094	.240**	-.441	.092	-.313**	-1.94	1.47	-.105
GPA	-.105	.109	-.057	-.052	.053	-.061	-.002	.051	-.002	.259	.817	.022
Black	.169	.263	.039	-.309	.126	-.153*	-.035	.124	-.016	-3.13	1.97	-.113
Asian	-.080	.159	-.031	-.024	.077	-.020	.126	.075	.100	-3.02	1.19	-.183**
Arab	-.061	.266	-.014	-.033	.128	-.016	-.027	.125	-.012	-5.01	1.99	-.175**
Multiracial	.153	.548	.016	-.268	.264	-.061	-.137	.258	-.030	-2.45	4.10	-.041
Team3	-.192	.177	-.074	-.050	.085	-.041	-.006	.083	-.004	-1.43	1.33	-.085
Team5	-.544	.237	-.141*	-.045	.114	-.025	.117	.112	.061	5.87	1.77	.235**
Team6	-.354	.207	-.115	.062	.099	.043	-.089	.097	-.059	2.81	1.54	.141
Team7	-.012	.402	-.002	-.135	.193	-.043	-.032	.189	-.010	3.76	3.01	.088
Team2	.597	.249	.169*	-.216	.119	-.130	.167	.117	.095	2.24	1.86	.098
Feminine	-.332	.200	-.109	.782	.096	.544**	-.923	.094	-.609**	.000	1.50	.000
Masculine	1.122	.180	.396**	-.266	.087	-.200**	.340	.085	.242**	.325	1.35	.20
InteractionFxM	.523	.466	.073	-.176	.224	-.052	-.058	.219	-.016	1.95	3.50	.042
R^2		.365			.336			.427			.154	
F		6.244			6.677			9.836			2.410	

* p < .05, **p < .01. B, unstandardized regression coefficient; SE, unstandardized standard error; β, standardized beta.

Note. "Team 2 – Team 7" is referring to team size; "InteractionFxM" is referring to the product of feminine and masculine (i.e., androgyny).

Discussion

The purpose of Study 2 was twofold. The first aim was to replicate the factor structure from Study 1 and test the convergent/divergent validity of the masculine and feminine subscales of my team androgyny instrument. The second aim was to examine androgyny at the team level of analysis in conjunction with risk taking, ethical responsibility, and performance.

I found that the team androgyny items loaded well onto two factors, replicating the pilot study, Study 1b, and confirming hypothesis 1a. However, prior to analysis, a few more items were removed due to concerns of face validity as some participants questioned how they related to their team project/presentation and/or decision-making more generally. Additionally, these items were originally inputted into the exploratory structural equation model and did not provide a good fit. Once removed, the model was an excellent fit. The team androgyny instrument also showed sufficient internal reliabilities for both the masculine and feminine scales. Moreover, they correlated weakly and positively, partially supporting hypothesis 1b. I hypothesized that the scales would be moderately and positively correlated; however, a weak positive correlation may better support my interpretation of gendered decision-making styles of teams, in that while the two scales have a slight positive relationship, they are somewhat independent, which is in line with the conceptualization of masculinity and femininity (Bem, 1974).

I found substantial evidence for the validity of my team androgyny instrument. At the *individual level of analysis*, the feminine scale correlated in the expected direction with all other measures of ethical responsibility (i.e., egoism and benevolence), ethical values, and work group participation that I used, except for risk propensity, confirming hypothesis 2a, 3a, 4a, and 6a. Regarding risk propensity, I had hypothesized a negative

relationship, however, there was a non-significant relationship. This is quite surprising given in past literature studies have shown that females are more risk averse whereas males are more risk seeking (e.g., Harris et al., 2006; Powell & Ansic, 1997). The masculine scale correlated as expected with ethical benevolence, risk propensity, and work group characteristics, supporting hypothesis 3b, 5a, and 6b. However, contrary to my hypothesis, masculinity was uncorrelated with ethical egoism, and weakly positively correlated with ethical values, I argue that this might have occurred because ethical values focus on moral aspects of a team's processes such as being responsible, fair, and respectful as opposed to a team's ethical reasoning when making a decision. Moreover, counterproductive workplace behaviors were included as an additional measure for divergent validity, assuming it would be uncorrelated with the masculine and feminine scale, partially supporting hypothesis 7. Interestingly, the feminine scale contributed negatively when engaging in counterproductive behaviors. Overall, the strongest correlations were obtained for femininity and ethical values.

Findings pertaining to the linkages between gendered decision-making styles and risk propensity, ethical responsibility, and performance at the *team level of analysis* were partially supported. While neither the main effects nor the interaction of these constructs significantly contributed to performance, the results related to risk and ethics were promising, in conjunction with past literature.

Broadly speaking, the results related to risk propensity suggest that team masculinity positively related to risk propensity when thinking about the project/presentation. In comparison, team femininity did not contribute to risk propensity. These findings support hypothesis 8a. Furthermore, while the present results regarding risk reinforce prior findings, much past literature is conducted at the individual level

and/or examines biological sex, as such, this is among the first to empirically investigate the linkage between *gender* and *risk* at the team level of analysis. Thus, by demonstrating that *teams'* high in masculinity are more likely to engage in risk while making decisions in comparison to feminine *teams*, extends previous empirical work at the individual level of analysis. Of the literature that has examined *gender* and risk taking, the findings are similar. For example, Meier-Pesti & Penz (2007) conducted a study that investigated both sex and gender as predictors of financial risk taking. Using 180 students and an experimental design with self-report measures, the authors found that masculinity was confirmed to be a predictor of financial risk taking, but femininity had no effect on risk taking. Likewise, although examining sex as opposed to gender, Nieboer (2013) found that as teams increased in the number of males, risk taking also increased (see also Karakowsky & Elangovan, 2001). Also noteworthy, the interaction between masculinity and femininity (i.e., androgyny) had no effect on risk propensity.

Regression analysis was in line with the idea that team femininity was positively related to ethical responsibility. Specifically, in relation to both ethical benevolence and ethical egoism, team femininity and masculinity supported the proposed patterns whereby femininity positively contributed to benevolence and negatively contributed to egoism and team masculinity did the opposite. Again, literature examining ethics, *gender* and *teams* is rare; However, at the individual level, as various authors have argued, in general women seem to be more ethically responsible than their male counterparts (see Bampton & Maclagan, 2009; Chan, Jamilah, & Rusinah, 2012; Valentine, Godkin, Page, & Rittenburg, 2009). One study considered the interplay of sex and gender roles on ethicality further supporting my present findings. Suar & Gochhayat (2014) examined if gender roles (i.e., femininity) were a better predictor than ethicality than being female;

specifically, considering corporate social responsibility, ethical attitudes, and behaviors, femininity was found to be a better predictor than biological sex. As such, this study extends earlier findings at the individual level of analysis. Furthermore, as found with risk propensity, the interaction effects did not significantly contribute to ethical benevolence or egoism.

There has been no research linking gendered decision-making styles of teams to performance-based outcomes. A unique aspect of this research was that I examined the link between team androgyny and performance (i.e., team grades). However, in contrast to past literature examining androgynous leadership at the individual level, the present study results, suggest that androgynous teams may not play a significant role in increasing performance outcomes (e.g., high team grade), thus, refuting hypothesis 8e. However, the regression analysis showed no gender differences at all; thus, while discouraging, I attribute this lack of findings to statistical power, as the distribution of means of the performance outcome between teams was small (team grades mean was 82). Furthermore, although non-significant, it is possible that masculine teams did perform better given the nature of the course projects. The majority of teams included in the sample were required to conduct a case study analysis and present/write up their recommendations. I question if the business cases assessed required more masculine decision-making (e.g., risk propensity) given the likelihood as the male as the protagonist. Symons & Ibarra (2014) noted that less than 10 percent of women are featured as the protagonist in business cases. Furthermore, after assessing the top 53 business case studies (according to Harvard Business Review), women were completely absent in 24 of them, and only 5 describe the women as the leader. As such, students may be socialized to see a masculine style as more conducive to successful decision-making *and* perhaps a more masculine style was

required to achieve higher performance. Additionally, educational research (e.g., Jule, 2004) has demonstrated that males tend to speak up more likely than females in the classroom, tend to be called on more and thus, get higher grades. A recent Harvard business case study also showed that while women to well on tests they fall very behind on in class participation and discussions (Kantor, 2013). Given the nature of the decision-making task being a case *and* presentation, it is not surprising that masculine teams did slightly better due to their assertive and confident style.

Interestingly, the only significant predictors of performance that emerged in the present study were team size and ethnicity of the team, with teams of five positively contributed to performance and ethnicity negatively contributing. In relation to team size, in line with previous literature, these findings suggest that size matters. Sutter (2004), for example, examined team size on decision-making using an experimental design, and found significant differences between two and four members whereby teams with four members outperformed any other number variation. Additionally, while some researchers may argue against the negative contribution of ethnicity to performance, Lee & Farh (2004) also found a negative association between diversity of teams and performance via project grade (see also Webber & Donaue, 1999).

Limitations and Future Research

The first limitation that is quite common in team-level research in the relatively small sample size (Lim & Ployhart, 2004), as my study consisted of 62 teams of various sizes, with a total of 238 participants. However, the significant results demonstrate that statistical power is not a major concern in several analyses, except for performance. For example, a power analysis (using the median split classified groups), showed that there was a 58% chance of finding population effects of r = .40 at $p < .05$. Therefore, the non-

significant findings in performance likely correspond with low observed power (see O'Keefe, 2007), and not related to the null hypothesis. I must stress, however, that this study was considered a pilot study because I was unable to test my entire conceptual model; therefore, it is one of the major reasons I included an experimental design in my final study – allowing me to have greater control of team size, team numbers, and team sex.

I also used a cross-sectional, non-experimental design in this study with a student sample. This choice may have been limiting for two reasons: First, because this study was cross-sectional and utilized some self-reporting procedures, it is possible that common method bias influenced my results. However, the use of self-report data continues to be widespread and several researchers agree that while there is the view that self-reporting *can* affect validity (Chan, 2009; Spector, 2006); Chan (2009) noted that "self-report data are not really that bad and do not deserve the negative reputation…" (p. 310). One reason being that common method bias is unlikely to be large enough to "invalidate many of our theoretical interpretations" (Doty & Glick, 1998, p. 400).

To err on the side of caution, I reduced self-report bias by not disclosing the gendered-nature of the team androgyny instrument while participants completed the questionnaire – it stated, "using the following scale below, please check whether you agree or disagree with each statement occurring in your team as your worked on your project/presentation." I also stressed to participants that despite it being a questionnaire regarding their team, they were expected to complete it individually and that their responses would be confidential and the data anonymous. I argue that this also lessened responding in a socially desirable manner given that participants were unaware of which measures had a gender component as they were worded "neutrally" (e.g., Meier-Pesti &

Penz, 2008; Vainapel, Shamir, Tenenbaum, & Gilam, 2015), used clear language, were not worded negatively and were not ambiguous (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Similar to self-report bias, it is worth mentioning the possibility of common method variance; Although, while my independent variables were measured as self-report, my dependent performance variable came from a different source, the instructor (i.e., team grades) (Podsakoff et al., 2003). Additionally, in terms of the issue of inflated observed correlations, as evident in the data, this was not a threat given that there were several non-significant correlations between many variables and several of my hypotheses were supported.

Another limitation is the generalizability of findings because all data was collected from a sample of two universities. It is not possible, thus, to demonstrate the extent to which these findings are related to organizational teams. However, the sample was quite diverse relative to sex, age, ethnicity, and focus area, representing a broad array of the population. Future research may achieve higher generalizability by conducting data using actual organizational teams from diverse industries.

Finally, despite including numerous control variables in my analyses, it is possible that other variables *may* have influenced my results in regard to risk propensity, ethical responsibility, and performance. As indicated in my conceptual model (see Appendix B), these include individual factors (e.g., values, personality traits etc.) and team-based factors (e.g., situational, behaviors, perceptions). Future research may consider analyzing the interplay of other variables that might affect the performance of teams in conjunction with gendered decision-making.

4: Study Three – Experimental Decision-Making Scenario

The objectives of Study 3 are twofold. First of all, I wanted to confirm the factor structure of my team androgyny instrument as well as replicate the validation findings from Study 2 using risk propensity and ethical responsibility.

I thought it was important to test the relationship between team masculinity and team femininity, with another measure of gender role traits. Traditionally, the conceptualization of gender role stereotypes tends to be classified as either masculine or feminine (Abele, 2003), as evident in my dissertation, *and* are often automatically assigned to the relevant sexes. To veer away from the assumption that masculine/male and feminine/female is constant, I wanted to include communality and agenticism given that these constructs are broader in their domain (Moskowitz, Suh, & Desaulniers, 1994) and tend to depend on a number of situational variables in regard to potential sex differences. For example, in organizational and leadership settings, researchers have found that sex differences are less pronounced when examining communality (considered feminine) and agenticism (considered masculine) (Eagly, 1987). Also interesting, while communality and agenticism are typically viewed as *gender role* traits, numerous studies have included the constructs without the examination of gender and or biological sex per se. Some researchers have examined the dimensions of communality and agenticism as a means for "describing and judging persons and groups (Abele & Wojciszke, 2007, p. 751)" more generally as the perspective of self and others. Agenticism most often referring to efficiency to attain one's goals and including attributes such as persistent, assertive, decisive, confident, and efficient. On the other hand, communion relating to the interests of others and includes traits like benevolence, empathetic, supportive, and accepting (Abele & Wojciszke, 2007; Berkery, 2013). Consequently, while my

dissertation does examine masculinity and femininity, testing the relationship between communion and agenticism, which are more *accepted* terms, may further support my instrument (e.g., team androgyny), specifically if they correlate as expected.

Therefore, I chose to include a measure of communality and agenticism specific for use in organizational decision-making tasks (Heilman & Okimoto, 2007). Despite using the conceptualization of androgyny, the BSRI (Bem, 1974) was not selected as a measure of convergent validity because it is considered a measure of gender roles in society more generally and several items (on both the short form and long form) are irrelevant in regard to gendered decision-making of teams. For example, the short form consists of feminine items such as "loves children," "gentle," "tender," and "affectionate," and masculine items such as "independent," "strong personality," and "forceful." Therefore, in regard to validation, I hypothesized:

Hypothesis 9a: Team androgyny will represent a two-factor solution, one factor representing masculinity, and one factor representing femininity.

Hypothesis 9b: Masculinity and femininity will be moderately and positively correlated.

Hypothesis 10a: A positive correlation between the feminine scale and communality.

Hypothesis 10b: A positive correlation between the masculine scale and agenticism.

Hypothesis 10c: A non-significant relationship between the feminine scale and agenticism.

Hypothesis 10d: A non-significant relationship between the masculine scale and communality.

Hypothesis 11a: A positive correlation between the masculine scale and ethical egoism.

Hypothesis 11b: A negative correlation between the feminine scale and ethical egoism.

Hypothesis 12a: A positive correlation between the feminine scale and ethical benevolence.

Hypothesis 12b: A negative correlation between the masculine scale and ethical benevolence.

Hypothesis 13a: A positive correlation between the masculine scale and risk propensity.

Hypothesis 13b: A negative correlation between the feminine scale and risk propensity.

The second, and main objective of Study 3 was to experimentally test the sex composition of a team, gendered decision-making styles of a team, and performance on an objective task. While Study 2 allowed me to partially explore the links of gendered decision-making styles and performance of teams, Study 3 also allowed me to vary the sex composition of each team by using an experimental design in a controlled setting. Furthermore, like Study 2, I determined if risk propensity and ethical responsibility had a gender effect at the team level of analysis. The following hypothesis were examined in this study:

Hypothesis 14a: Team masculinity and androgyny will be positively related to risk propensity in comparison to team femininity, which will be negatively related.

Hypothesis 14b: Team masculinity will be positively related to ethical egoism in comparison to team femininity or androgyny, which will be negatively related.

Hypothesis 14c: Team femininity and androgyny will be more positively related to ethical benevolence than team masculinity.

Hypothesis 14d: Team androgyny will be positively related to team performance regardless of team sex composition.

Hypothesis 15a: Female and mixed sex teams will be significantly more feminine than male teams.

Hypothesis 15b: Male teams and mixed sex teams will be significantly more masculine than female teams.

Hypothesis 15c: Team androgyny and mixed sex composition will be more positive predictors of objective performance than team femininity, masculinity, all male, or female composition.

Method

Participants

A total of 245 undergraduate business and psychology students (female *n*= 124; male *n*= 120; 1 value missing) from one university participated in this study. Participants were recruited in class or via the SONA system and volunteered to participate. Some of the student participants received extra credit points in their classes dependent on the instructor and/or course. Students ranged in age from 17 to 54 years, with a mean of 22.88 years (*SD* = 3.7). Furthermore, the ethnic distribution of the students was diverse with 52.3% Caucasian, 16.7% Asian, 9% Black, and 11% Arab, the remaining falling in the multiracial and other category.

Procedure

Students were randomly assigned to 1 of 53 three – six person teams. The sex composition of the team was also randomized into all male, all female, or mixed sex, I did my best to ensure that these sex combinations were counterbalanced in each condition (male teams n = 18; female teams n = 18*;* mixed sex teams n = 17). Students worked together in teams in either a classroom setting or in a meeting room depending on how they were recruited.

At the beginning of the 1.5-hour session, participants were told that they would be participating in a team decision-making scenario called the Winter Survival Exercise (Johnson & Johnson, 1982; 2012) then completing an individual questionnaire on the team they had worked with. Prior to the experiment, students were given 10 minutes to review the informed consent form. See Appendix J for Study 3 survey materials and Appendix K for written permission to use the Winter Survival Exercise.

The Winter Survival Exercise was developed primarily as a teaching tool to provide learning on group problem solving, decision-making, and team building (Johnson & Johnson, 1982). The Winter Survival Exercise was chosen for the purpose of this study due to it resulting in an objective performance score that could be analyzed quantitatively (Rogelberg, Barnes-Farrell, & Lowe, 1992). Prior to completing the Winter Survival Exercise students were told that they were to imagine that they were in a plane that crash-landed in Manitoba during winter, temperatures were minus 25 degrees Celsius and they were over 80 kilometers from the nearest town. They were told that 15 items had been salvaged in the plane crash and they had to rank these on importance to survival (1 being most important – 15 being least important). Sample items included a knife, compass, 30 feet of rope, and a flashlight with batteries etc. Students were then told that they had to

first rank the items on an individual basis and then as a team, ensuring that the team discussed their individual rankings together before deciding on a final team ranking. Approximately 10 minutes was allocated to ranking individual items, and 40 minutes for the team ranking. The remaining time was allocated to the questionnaire and a debriefing period.

The Winter Survival Exercise answer key consists of an "expert score" as developed by wilderness experts in the U.S. Army. The scoring procedure is the sum of the absolute difference between the expert ranks and the team ranks. A lower score indicates higher likelihood of survival whereas a high score indicates less likelihood of survival. While students were not given the expert score during the experiment, they had the option of providing their e-mail to receive the expert score upon completion of the study. Therefore, I completed all scoring post experiment. See Appendix L for Winter Survival Exercise materials

Measures

The following section describes the measures included for the purpose of data analysis in study 3.

Team Androgyny

After removal of two inconsistent items (described below), the final team androgyny measure consisted of a 10-item scale assessing the gendered decision-making styles of teams. Five of the items were perceived as feminine, and five as masculine. Sample items that measured feminine decision-making styles of teams included "our team encouraged open communication" and "we considered the rights of everyone to arrive at our final decision; masculine items included "our team was willing to take risks when completing the exercise" and "we used assertiveness to reach a solution." Each

participant was asked to rate whether they agreed or disagreed with each statement occurring in their team as they worked on the Winter Survival Exercise using a five-point Likert scale (1- strongly disagree to 5 – strongly agree). The internal reliability of the scales was acceptable (i.e., feminine scale - .71; masculine scale - .68).

Perceived Communality and Agenticism

I measured perceived communality (e.g., femininity) and agenticism (e.g., masculinity) using a 9-point bipolar adjective scale adapted from Heilman & Okimoto (2007). I modified the scale from the individual level to the team level. I included four communal items (e.g., supportive – unsupportive, understanding – not understanding, caring to not caring) and five agentic items (e.g., assertive – not assertive, strong – weak, active – passive). Each participant was asked to circle the appropriate number as it related to their team as they completed the exercise. Heilman & Okimoto (2007) reported internal reliabilities of $\alpha = .90$ (communal) and $\alpha = .82$ (masculine) for the respective scales. Likewise, for this study, the perceived communal and agentic scales had good internal reliabilities of $\alpha = .91$ and $\alpha = .90$, respectively. These scales were primarily used to further test the convergent validity of my team androgyny instrument.

Ethical Responsibility

I included a measure of ethical responsibility (i.e., ethical reasoning) that consisted of multiple subscales (Cullen & Victor, 2008; Cullen, Victor, & Bronson, 1993; Victor & Cullen 1987). For the purpose of this dissertation, like Study 2, I used the subscales of egoism (3 items) and benevolence (3 items) because they relate to maximizing joint interests within a team, as well as maximizing one's own interest without concern for the team. Items were modified to represent team as opposed to company. Items included "in our team, people are mostly concerned for themselves," and

"in our team, our major concern is what is fair for everyone" and were rated on a five-point Likert scale from 1 – mostly false to 5 – completely true. Cronbach's alpha was $\alpha = .70$ for egoism and $\alpha = .80$ for benevolence.

Risk Propensity

To measure risk propensity, I included a subscale with one item from an instrument developed by Meertens & Lion (2008) called the Risk Propensity Scale. For the purpose of this study, I adapted the item to the team level of analysis. The item asked participants to rate the extent to which their team defined themselves according to risk while completing the Winter Survival Exercise, the Likert scale ranged from 1 – risk avoiders to 9 – risk seekers.

Team Performance

As noted previously, objective team performance was tested using the Winter Survival Exercise and calculated as the absolute sum of differences between the expert rankings and the team rankings of the 15 survival items. Lower scores indicate higher performance.

Control Variables

A number of control variables were included in the study to ensure that certain covariates were not biasing my results. First, I captured the *ethnicity* of team members within each team because culture may play a role in perceptions towards masculinity and femininity (Schein & Mueller, 1992). Ethnicity was dummy-coded. The majority of team members were Caucasian (52.3%); therefore, I argue that using Caucasian as the reference category was justified. Second, past survival training might influence results of the Winter Survival Exercise; therefore, survival training was also dummy-coded (1 – yes; 0 – no). Third, as recommended by Bliese & Halverson (1998), team size was used

as a control variable because team size has been found to influence performance. Finally, team sex was dummy coded (female, male, and mixed sex) as the influence of team sex on performance was to be tested in one of the hypotheses analysis below[5]. Like Study 2, these control variables were included primarily because they were theoretically relevant and conceptually meaningful (Becker et al., 2016).

Data Analysis

Data analysis for Study 3 included several steps: First, I confirmed the factor structure of my team androgyny instrument as well as further testing its convergent validity using another organizational gender role scale. Second, I aggregated all included measures to the team level of analysis. Finally, I conducted my hypotheses testing using multiple hierarchical regression and multiple analysis of variance approaches. All data was screened for missing values, skewness and kurtosis, and outliers prior to conducting any analyses (Meyers et al., 2013).

Factor Structure

First of all, a preliminary confirmatory factor analysis (CFA) was conducted to test the structure of my masculine and feminine scales. To test the fit of each of these scales, five indices were utilized including the chi-square statistic, the comparative fit index (CFI), the Tucker Lewis Index (TLI), the standardized root mean square residual (SRMR), and the root mean square error approximation (RMSEA). Traditionally, the chi-square statistic should result in non-significance because it examines the difference between the predicted and observed correlations (Meyers et al., 2013). However, recently,

[5] I did not include biological sex as a control variable because hypothesis 14d aimed to test the effects of gendered decision-making styles and performance of teams' regardless individual sex. While hypothesis 15c considers biological sex, it does so in regard to the team, which is why I have dummy coded team sex to be used as a control variable.

researchers caution in using the chi-square statistic as the only index to test model fit because of sample size issues. Chi-square will likely be non-significant in studies utilizing less than 200 cases but as sample size increases, power increases, and so too does the likelihood of statistical significance (Kenny & McCoach, 2003). Therefore, other fit indexes should be used in addition to the chi-square statistic.

The RMSEA is another index that tests the absolute goodness of fit and accommodates larger sample sizes (Kaplan, 2000). Values less than .08 are deemed acceptable and values less than .05 indicate good fit (Meyers et al., 2013). Likewise, the SRMR considers the difference between the observed and predicted correlation. Values can range from zero to one with a good fitting model obtaining less than .05, and an acceptable model below .08 (Byrne, 1998; Hooper, Coughlan, & Mullen, 2008).

Both the CFI and TLI are relative fit measures and compare the specified model to a null model (Meyers et al., 2013). Hu and Bentler (1999) state that a good fit is indicated by values above .95, an acceptable fit is above .90 and a marginal effect is between .80 and .89. According to Tabachnick & Fidell (2007), the CFI and the RMSEA are commonly reported fit measures. As noted previously, a chi-square value is affected by sample size whereby a small sample size (e.g., 75 cases) will likely produce non-significant results. However, as the sample size gets larger so too does the chi-square values, which also produces significant results (Jackson & Gillaspy, 2009).

Convergent and Divergent Validity

To further test the convergent and divergent validity of my masculinity and femininity scales of the team androgyny instrument, I included an agentic/communal scale, ethical egoism and benevolence, and risk propensity. Additionally, conducting the

CFA allowed me to show that the feminine and masculine scales loaded onto the hypothesized factors.

Data aggregation

I followed the same data aggregation procedures as in Study 2. Variables that had to be aggregated included the feminine and masculine scales of the team androgyny instrument, ethical egoism and benevolence, and risk propensity. I did not have to aggregate my objective team performance score as this included only one score per team. In terms of a composition model, each variable was classified as *direct consensus constructs* because the data was collected individually, but the constructs are representative of the team level.

I specifically used within-group agreement (i.e., IRA), and intra-class correlation coefficients (ICC) to justify my aggregation. To test the IRA, I used the $r_{wg(j)}$ index, which has a recommended median score from .51 - .70 for a moderate effect, and .71 - .90 to demonstrate strong within-group agreement (LeBreton & Senter, 2008). The $r_{wg(j)}$ index was calculated using an Excel tool (Biermann, Cole, & Voelpel, 2012). To test the ICC, I computed both the ICC (1) and ICC (*K*) (LeBreton & Senter, 2008), using the statistical software SPSS and inputting syntax recommended by LeBreton & Senter (2008, p. 844). Values for ICC (1) are similar to recommendations for interpreting effect sizes, whereby a value of .01 is small effect, .05 small to medium effect, .10 a medium effect, and .25 a large effect (LeBreton & Senter, 2008). However, values for ICC (*K*) are often debated, but values above .70 and significant *F* values are considered sufficient in justification (Chen, Mathieu, & Bliese, 2004).

Descriptive statistics and correlations

Once I *aggregated* the main predictor variables (i.e., masculine, feminine, ethical climate, risk propensity, and subjective team performance scales), I calculated descriptive statistics including the means and standard deviations, and the bivariate correlations for each variable to be used in the main analysis.

Hypothesis testing

To test my hypotheses, I used SPSS Version 24 to conduct a series of independent multiple hierarchical regressions, a multiple analysis of variance (MANOVA), followed by a series of analysis of variance (ANOVA) models using Tukey post hoc tests. Like Study 2, to be as thorough as possible in my data analysis, I also decided to complete the median split method. See Appendix M for results.

First, when testing the hypotheses relevant to team gender and team performance, risk propensity or ethical responsibility (i.e., Hypothesis 14a – 14d), I regressed the outcome variable (i.e., performance, ethical climate, or risk propensity) on my control variables (i.e., ethnicity, past survival training, team size) in the first step, aggregated masculinity and femininity in the second step, and the interaction between masculinity and femininity (i.e., androgyny) in the third step.

A similar approach was taken when testing the hypothesis regarding team sex, aggregated gender, and team performance. Team performance was regressed on the control variables (i.e., ethnicity, past survival training, team size, *and* team sex) in the first step, masculinity and femininity in the second step, and the interaction between masculinity and femininity in the final step.

Since I have multiplicative terms in my analysis, I centered the predictor variables (e.g., masculinity, femininity, and the interaction between masculinity and femininity) prior to conducting my regressions. I did this to prevent errors in statistical inference and

to avoid multicollinearity issues (see Kraemer & Blasey, 2004). In addition, when conducting the analysis, I calculated the variance inflation factors (VIFs) to ensure that none of my predictor variables were high (i.e., above 10 – Mason & Perreault, 1991).

Finally, to compare team sex composition and team gender, I conducted a MANOVA followed by a series of ANOVA models using Tukey post hoc test. This method allowed me to test hypothesis 15a and 15b.

Results

Factor Structure

Based on the results of the ESEM in study 2, a two-factor solution was tested to assess the construct validity of the team androgyny scale prior to further analyses[6]. Results from the CFA using Mplus7 indicate acceptable fit indices, supporting hypothesis 9a. Although the chi-square test of model fit was statistically significant – 58.19 (34, n = 245), $p < .006$, the model yielded good fit indices for RMSEA and SRMR, and acceptable fit indices for CFI and TLI. The RMSEA was .054, the CFI was .94, TLI .91, and the SRMR index was .049. As stated previously, a chi-square value is affected by sample size whereby a small sample size (e.g., 75 cases) will likely produce non-significant results; however, as the sample size gets larger so too does the chi-square values, which also produces significant results (Jackson & Gillaspy, 2009). As such, I argue that my sample size (n = 245) may have contributed to marginal significance.

Convergent and Divergent Validity

[6] I also tested a structure with one factor, which showed a poor fit to the data.

Table 4.1 represents the means, standard deviations, and bivariate correlations between the masculinity and femininity subscales of the team androgyny instrument and the validity correlates at the *individual level of analysis.*

First, the feminine scale and masculine scale demonstrated a weak but positive correlation, partially supporting hypothesis 9b and replicating the results in Study 2.

As hypothesized (hypothesis 10a and 10b), the feminine scale was shown to have a significant and positive relationship with the communal scale (r = .332, p < .01), and the masculine scale had a significant positive relationship with the agentic scale (r = .189, p < .01). Additionally, the masculine scale was uncorrelated with the communal scale (r = .105, p = .10), supporting hypothesis 10d; however, contrary to my hypothesis, the feminine scale was also weakly but positively correlated with the agentic scale (r = .164, p < .05), disconfirming hypothesis 10c.

Likewise, as hypothesized, femininity was negatively correlated with ethical egoism (r = -.371, p < .01) and positively correlated with ethical benevolence (r = .485, p < .01), as stated in hypothesis 11b and 12a. Similar to study 2, while I hypothesized that masculinity would be positively correlated to ethical egoism, the two were unrelated (r = .031, p = .631). Also contrary to my hypothesis, masculinity and ethical benevolence had a weak positive correlation (r = .212, p < .05), as such hypothesis 11a and 12b were not supported.

Finally, as found in Study 2 and supporting hypothesis 13a (but not 13b), risk propensity had a significant positive relationship with masculinity (r = .310, p < .01) and was uncorrelated with femininity (r = -.021, p = .74).

Table 4.1
Overall bivariate correlations between masculinity, femininity, and all other variables

Scale	*M*	*SD*	1	2	3	4	5	6	7
1. Team Androgyny – Feminine	4.23	.5	(.71)	.385**	.164*	.332**	-.371**	.485**	-.021
2. Team Androgyny – Masculine	3.51	.62		(.68)	.189**	.105	.031	.212**	.310**
3. Agentic	6.33	.1.9			(.90)	.677*	-.140*	.094	-.065
4. Communal	6.45	2.0				(.91)	-.288**	.215**	-.152*
5. ECQ – Egoism	1.93	.85					(.70)	-.328**	.104
6. ECQ – Benevolence	3.89	.77						(.80)	-.021
7. Risk Propensity	4.86	2.1							-

Note. Reliabilities for each scale are presented on the diagonal in parentheses. * Correlation is significant at the .05 level (2 – tailed). **Correlation is significant at the .01 level (2 – tailed).

Data Aggregation

Aggregation statistics for the masculine and feminine gendered decision-making styles of team scales, ethical egoism and ethical benevolence scales, and risk propensity are shown in Table 4.2. As demonstrated, these variables were in the acceptable range for the $r_{wg(j)}$ index, the ICC (1), and ICC (*K*). Therefore, I aggregated my data to the team level of analysis.

Regarding the feminine scale, the median $r_{wg(j)}$ was .95, indicating a high within-group agreement. ICC (1) was .27 and ICC (*K*) was .65, with the *F* statistic value being significant, as recommended, F [243, 976] = 2.82, $p < .05$. Likewise, the masculine scale had a median $r_{wg(j)}$ of .84, with an ICC (1) value of .24, ICC (*K*) value of .65, and a significant *F* value of F [242, 1215] = 2.87, $p < .05$.

Ethical egoism had a median $r_{wg(j)}$ of .79, indicating a high within-group agreement. ICC (1) was .44 and ICC (*K*) was .70, with a significant F [244, 290] = 3.31, $p < .05$. Similarly, ethical benevolence had a median $r_{wg(j)}$ of .84, showing a strong within-group agreement, an ICC (1) of .48 and an ICC (*K*) of .79, and a significant *F* value of F [240, 723] = 4.68, $p < .05$. Finally, risk propensity, which consisted of one item, required the calculation of the median *r*wg, which was .51, deemed in the moderate range.

Table 4.2
Data Aggregation Statistics

	Feminine Scale	Masculine Scale	Ethical Egoism	Ethical Benevolence	Risk Propensity
Median	.95	.84	.79	.84	.51
$r_{wg(j)}$					
ICC (1)	.27	.24	.44	.48	-
ICC (*K*)	.65	.65	.70	.79	-
F	2.82***	2.87***	3.31***	4.68***	-

Note. n = 245 participants (nested within 53 teams). ***p<.001.

Descriptive Statistics and Correlations

Table 4.3 presents the means, standard deviations, and bivariate correlations for all variables used in hypothesis testing (i.e., control variables and *aggregated* variables).

Table 4.3
Overall bivariate correlations between control variables and aggregated predictor/outcome variables

			Correlations									
	M	*SD*	1	2	3	4	5	6	7	8	9	10
1. Team Androgyny – Feminine	4.23	.24	1	.476**	-.468**	.462**	-.007	-.187**	-.044	-.035	-.202**	.108
2. Team Androgyny – Masculine	3.51	.40		1	-.086	.296**	.448**	-.177**	-.005	.024	-.134*	-.152*
3. ECQ – Egoism	1.93	.58			1	-.416**	.071	.099	.107	-.055	-.126*	-.095
4. EQC – Benevolence	3.9	.40				1	.015	-.141*	-.084	-.049	.062	-.123
5. Risk Propensity	4.85	1.08					1	-.252**	.142*	.051	.145*	.014
6. Team Performance	44.71	16.19						1	-.167**	-.085	.079	.126*
7. Ethnicity	2.21	2.83							1	.184**	-.037	-.217**
8. Survival Training	.72	.45								1	.046	.044
9. Team Size	4.75	.68									1	.220**
10. Team Sex	1.00	.83										1

**Correlation is significant at the .01 level (2 – tailed). * Correlation is significant at the .05 level (2 – tailed).

Hypothesis Testing

Hierarchical Multiple Regression Analyses[7] [8]

Risk Propensity

Results of the hypotheses regarding risk propensity are depicted in column three of Table 13. The control variables, entered on the first block were moderately significant F (8, 225) = 2.28, $p < .05$, but only accounted for 7.5 percent of predictive variance. Ethnicity (i.e., being Arabic) had a significant and positive effect on risk propensity (β = .17, $p < .01$) whereas team size (i.e., three members) had a significant negative effect on risk (β = -.149, $p < .05$). When masculinity and femininity were entered on the second block, the prediction model was statistically significant, F (10, 223) = 14.98, $p < .001$, and accounted for an additional 33 percent of the predictive variance. Namely, femininity significantly negatively (β = -.321, $p < .001$) and masculinity significantly positively (β = .712, $p < .001$) contributed to risk propensity. Finally, the third block, which contained the interaction between masculinity and femininity, did not make a significant contribution in the regression models, with a ΔR^2 of .002. Thus, hypothesis 14a was partially supported.

Ethical Responsibility

Results of the hypothesis regarding ethical egoism and team gender are depicted in the first column of Table 4.4. Inputting the control variables in step one, which included ethnicity (dummy coded), survival training, and team size (dummy coded), resulted in statistical significance, F (8, 225) = 2.48, $p < .01$, and accounted for approximately 8 percent of the predictive power. Specifically, teams size (i.e., six

[7] Please note that I also performed hierarchical multiple regression analyses without the control variable ethnicity (i.e., dummy coded). No new main effects emerged.

[8] As with Study 2, I also repeated relevant hypothesis testing using standard multiple regression whereby all variables were included in the same block. The results were largely the same.

members) resulted in lower egoism in comparison to the reference group (β = -.258, $p <$ .001). The second block, which contained the masculine and feminine scores, also demonstrated statistical significance, F (10, 223) = 11.66, $p < .001$, and increased substantially in predictive power with a ΔR^2 of .26. Specifically, beyond the covariates, masculinity was significantly and positively related to ethical egoism (β = .152, $p < .05$) and femininity was significantly and negatively related to ethical egoism (β = -.604, $p <$ .001). The final block, while significant, F (11, 222) = 10.69, $p < .001$, only attributed an addition .3 percent of predictive power to the model. More importantly, the interaction between masculinity and femininity did not significantly contribute to ethical egoism (β = -.066, p = .304), thus partially supporting hypothesis 14b.

The regression results for ethical benevolence and gender are shown in the second column of Table 4.4. The first block, which contained the control variables was not significant, F (8, 225) = .98, p = .452, R^2 = .034. However, the second block, which also included the team femininity and masculinity scores resulted in statistical significance, F (10, 223) = 8.69, $p < .001$, whereby femininity positively and significantly contributed to ethical benevolence, with a ΔR^2 of .28, adding an additional 2 percent variance to the model (team masculinity did not significantly contribute). The final block, which included the interaction between masculinity and femininity did not significantly contribute to ethical benevolence (β = .018, $p <$ .792), nor increase the predictive power, ΔR^2 = .000. However, the regression coefficient for the interaction effect was in the positive hypothesized direction. Therefore, hypothesis 14c was also partially confirmed.

Team Performance

Results of the hypothesis pertaining to team androgyny and team performance are depicted in the fourth column of Table 4.4. The dummy variables ethnicity, survival training, and team size, entered on the first step, was statistically significant, F (8, 225) = 5.522, $p < .001$, and accounted for approximately 16 percent of the variance of team performance ($R^2 = .164$). Specifically, being Arabic ($\beta = -.126$, $p < .05$) or Black ($\beta = -.13$, $p < .05$) decent resulted in poorer performance than the reference category of Caucasian. Likewise, teams with six members ($\beta = -.237$, $p < .001$) or four members ($\beta = -.288$, $p < .001$) also negatively contributed to performance in comparison to teams with five members.

The second block, which included the masculine and feminine scores, also showed statistical significance, F (10, 223) = 7.243, $p < .001$, with a ΔR^2 of .08. This significance was over and above the effects of the covariates; specifically, being a masculine team resulted in poorer performance ($\beta = -.32$, $p < .001$). While femininity was not a significant predictor of the regression, it was in the predicted positive direction.

Finally, the third block was also significant, F (11, 222) = 6.87, $p < .001$; however, it only increased minutely in predictive variance with a $\Delta R^2 = .009$. Most importantly, the interaction term between masculinity and femininity did not significantly contribute to team performance ($\beta = -.11$, $p = .104$), nor did it go in the hypothesized direction. As such, there was no support for hypothesis 14d that team androgyny would be positively related to performance.

Table 4.4
Multiple Hierarchical Regression Analysis of Team Gender and Outcomes Variables

	Ethical Egoism			Ethical Benevolence			Risk Propensity			Team Performance		
	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β
Constant	2.00	.08		3.93	.058		4.738	.154		53.21	2.166	
Asian	.195	.109	.126	-.168	.077	-.156*	.110	.206	.038	-4.077	2.893	-.094
Arab	.170	.121	.093	.004	.086	.003	.584	.230	.169**	-6.803	3.224	-.134*
Black	.062	.133	.031	.016	.094	.012	.421	.251	.112	-6.997	3.527	-.126*
Multiracial	-.06	.199	-.018	.018	.141	.008	.100	.377	.018	-7.078	5.289	-.084
Survival	-.096	.086	-.073	-.020	.061	-.023	.083	.164	.034	-3.125	2.299	-.086
Team6	-.565	.145	-.258**	.130	.103	.086	.203	.275	.049	-14.47	3.858	-237**
Team4	.010	.095	.007	-.009	.067	-.010	-.192	.179	-.073	-11.11	2.514	-.288**
Team3	-.034	.197	-.011	.011	.140	.005	-.852	.372	-.149*	-3.569	5.228	-.042
R^2		.081			.034			.075			.164	
F		2.479*			.980			2.81			5.52	
Constant	2.00	.069		3.918	.050		4.63	.125		54.02	2.074	
Asian	.176	.093	.114	-.130	.067	-.121	.305	.167	.104	-5.610	2.779	-.130*
Arab	.029	.104	.016	.103	.076	.082	.631	.188	.183**	-8.036	3.113	-.158**
Black	.050	.113	.025	.037	.082	.027	.522	.203	.138**	-7.803	3.371	-.140*
Multiracial	-.123	.169	-.041	.044	.123	.021	-.045	.305	-.008	-6.449	5.056	-.076
Survival	-.101	.073	-.078	-.025	.053	-.028	.015	.132	.006	-2.66	2.197	-.073
Team6	-.565	.145	-.258**	.100	.096	.066	1.125	.239	.272**	-19.78	3.962	-.324**
Team4	.010	.095	.007**	-.140	.061	-.145	-.002	.152	-.001	-11.03	2.515	-.286**
Team3	-.034	.197	-.011	-.017	.124	-.008	-.129	.309	-.023	-7.68	5.122	-.091
Feminine	-1.47	.164	-.604**	.792	.119	.468**	-1.50	.296	-.327**	1.611	4.908	.024
Masculine	.223	.099	.152**	.100	.072	.098	1.973	.179	.712**	-12.90	2.972	-.315**
R^2		.343			.28			.402			.245	
F		11.656**			8.69			14.98			7.243	
Constant	2.022	.073		3.914	.053		4.59	.131		55.08	2.166	
Asian	.173	.093	.112	-.130	.068	-.121	.309	.168	.106	-5.752	2.77	-.133*
Arab	.023	.104	.013	.104	.076	.083	.639	.188	.185**	-8.295	3.105	-.163**
Black	.054	.113	.027	.036	.082	.026	.518	.203	.137**	-7.65	3.359	-.137*
Multiracial	-.137	.170	-.045	.046	.123	.022	-.025	.306	-.004	-7.108	5.054	-.084
Survival	-.106	.074	-.081	-.024	.054	-.027	.021	.133	.009	-2.871	2.192	-.079
Team6	-.355	.133	-.162**	.102	.097	.067	1.139	.240	.276**	-20.24	3.957	-.332**
Team4	.228	.085	.165**	-.137	.062	-.143	.018	.154	.007	-11.68	2.536	-.303**
Team3	.124	.173	.041	-.013	.126	-.006	-.099	.311	-.017	-8.688	5.140	-.103
Feminine	-1.476	.164	-.604**	.792	.119	.568**	-1.50	.296	-.327**	1.63	4.89	.024
Masculine	.268	.108	.182**	.092	.079	.090	1.91	.195	.689**	-10.80	3.227	-.264**
InteractionMxF	-.314	.305	-.066	.059	.222	.018	.451	.550	.050	-14.84	9.088	-.111
R^2		.346			.281			.404			.254	
F		10.696			7.874			13.66			6.876	

* $p < .05$, **$p < .01$. B, unstandardized regression coefficient; SE, unstandardized standard error; β, standardized beta.

Note. "Team3 – Team 6" is referring to team size; "InteractionMxF" is the product of masculinity and femininity (i.e., androgyny).

Results of the final hypothesis surrounding *team sex*[9], team gender, and performance are depicted in Table 4.5. The control variables, entered at step one, which included team sex (dummy coded), were statistically significant F (10, 223) = 5.48, $p <$ .001, and accounted for approximately 20 percent of the variance of team performance whereby teams with six members (β = -.308, $p <$ .001) or four members (β = -.271, $p <$ .001) and male teams (β = -.227, $p <$ .01) were negatively related to performance in comparison to the reference groups of five members and mixed sex teams. Female teams, did not make a significant contribution to performance.

The second block, which included team masculinity and femininity was also statistically significant F (12, 221) = 6.973, $p <$ .001, with an ΔR^2 = .077. Most importantly, team sex (i.e., male team – β = -.216, $p <$ .01) as well as team masculinity (β = -.295, $p <$ .001) contributed negatively to team performance.

The third block, while significant F (13, 220) = 6.484, $p <$ .001 did not increase in predictive variance, ΔR^2 = .002. Nor was the interaction between masculinity and femininity (β = -.060, p = .395), it was also in the negative direction; however, femininity was in the positive direction. As such, hypothesis 15c was rejected.

[9] I also conducted multiple hierarchical regressions using the control variables including *team sex*, and the predictor variables of team gender with the other dependent variables (i.e., risk propensity, ethical egoism, and ethical benevolence). These are not reported in the main body of the dissertation because while the models were significant, *team sex* did not make a significant contribution to risk propensity, ethical benevolence, or ethical egoism beyond the effects of the main predictor variables (i.e., masculinity and/or femininity).

Table 4.5
Multiple Hierarchical Regression Analysis of Team Sex and Performance

	Team Performance		
	B	SE B	β
Constant	57.08	2.530	
Asian	-3.55	2.86	-.082
Arab	-4.41	3.27	-.087
Black	-.5.77	3.50	-.104
Multiracial	-4.70	5.3	-.056
Survival	-3.60	2.27	-.099
Team6	-18.81	4.11	-.308**
Team4	-10.45	2.50	-.271**
Team3	-3.22	5.15	-.038
MaleTeam	-7.90	2.60	-.227**
FemaleTeam	-4.42	2.60	-.125
R^2		.197	
F		5.482	
Constant	57.69	2.42	
Asian	-5.05	2.75	-.117
Arab	-5.88	3.15	-.115
Black	-6.61	3.34	-.119*
Multiracial	-4.30	5.03	-.051
Survival	-3.13	2.17	-.086
Team6	-23.46	4.16	-.385**
Team4	-10.11	2.50	-.262**
Team3	-6.95	5.05	-.082
MaleTeam	-7.46	2.50	-.216**
FemaleTeam	-4.36	2.48	-.123
Feminine	-.187	4.94	-.003
Masculine	-12.08	2.96	-.295**
R^2		.235	
F		6.973	
Constant	58.00	2.45	
Asian	-5.15	2.75	-.119
Arab	-6.20	3.17	-.122*
Black	-6.62	3.35	-.119*
Multiracial	-4.83	5.07	-.057
Survival	-.3.20	2.18	-.088
Team6	-23.48	4.16	-.385**
Team4	-10.52	2.55	-.273**
Team3	-7.57	5.11	-.089
MaleTeam	-6.82	2.61	-.197**
FemaleTeam	-4.21	2.50	-.119
Feminine	.061	4.94	.001
Masculine	-11.04	3.20	-.270**
InteractionMxF	-8.05	9.44	-.060
R^2		.277	
F		6.484	

* p < .05, **p < .01. B, unstandardized regression coefficient; SE, unstandardized standard error; β, standardized beta.
Note. "Team3 – Team 6" is referring to team size; "InteractionMxF" is the product of masculinity and femininity (i.e., androgyny).

Multiple Analysis of Variance Approach

Team Sex and Gendered Decision-Making

First, I conducted a 2 (gender) x 3 (team sex) MANOVA of the aggregated data, followed by the single ANOVAs to test for differences between femininity/masculinity and team sex composition. The MANOVA yielded significant main effects between the dependent variables (masculinity and femininity) and team sex. Using Wilk's Lambda as the criterion, Wilk's $\Lambda = .904$, F (4, 482) = 6.17, $p < .001$. Univariate ANOVAs on each dependent variable indicated that there was a significant effect for the feminine subscale, F (2, 242) = 5.023, $p < .01$, whereby all female teams [M= 4.3, SD = .2] and mixed sex teams [M = 4.24, SD = .2] were significantly higher in femininity than all male teams [M = 4.1, SD = .30]. Additionally, there was a moderate significant univariate effect for the masculinity subscale, F (2, 242) = 2.95, p = .05. All male teams [M = 3.6, SD = .41] and female teams [M = 3.53, SD = .35] were significantly higher in masculinity than mixed sex teams [M = 3.43, SD = .42], showing partial support for hypothesis 15a and 15b. Table 4.6 presents the relevant means and standard deviations.

Table 4.6
Means and Standard Deviations of Team Sex and Gender

	Masculinity		Femininity	
Condition	*M*	*SD*	*M*	*SD*
All Male Teams	3.6*	.41	4.1	.30
All Female Teams	3.53*	.35	4.3*	.19
Mixed Sex Teams	3.43	.42	4.24*	.20

Note. The higher the mean, the higher in masculinity and femininity. Means differ significantly at $p < .05$.

Discussion

The aim of Study 3 was to confirm the factor structure of my team androgyny instrument. More importantly, I sought to examine the relationship of sex composition, team masculinity, femininity, and androgyny with risk propensity, ethical responsibility, and performance using experimental methodology. I also provided further evidence of both the reliability and validity of my measure.

More specifically, these results replicate and extend earlier findings in terms of factor structure, reliability and validation. In Study 3, I found the solution with two factors had an acceptable fit to the data from the CFA. This was after the removal of two additional items due to lack of clarity by students and did not provide a good fit. Additionally, reliabilities between the masculine and feminine scales were quite similar to Study 2, in that they were positively but weakly correlated, as such, hypothesis 9b was supported.

Specific to validation, at the individual level of analysis, the findings were promising as they were similar to the findings in Study 2. Also, noteworthy, was the inclusion of another measure of masculinity (i.e., agenticism) and femininity (i.e., communion) developed by Heilman & Okimoto (2007) for use in a study examining the gendered nature of leadership. As hypothesized, my team androgyny instrument correlated in the expected directions with communion and agenticism (confirming hypothesis 10a, 10b, and 10d), except for feminine and agenticism (hypothesis 10c), as these were weakly and positively correlated – I will return to this finding shortly.

As found in Study 2, validation in Study 3 also showed expected correlations between the feminine scale and ethical egoism and ethical benevolence, confirming hypothesis 11b and 12a. Interestingly, although disconfirming hypothesis 11a, study 3

replicated the results of study 2 in that masculinity was unrelated to ethical egoism. Likewise, masculinity had a very weak positive correlation with ethical benevolence, disconfirming hypothesis 12b. It is possible that the positive relationship existed in the current study because the Winter Survival Exercise required participants to have concern for others (i.e., fellow team members) in order to survive. Prior to beginning the exercise, teams were told that they were all survivors of a plane crash and had to collaboratively work together (i.e., in a fun way) to rank the items in order of survival. No specific instructions were given in terms of *how* they agreed on a specific ranking but they had to come to consensus in a short amount of time. As such, in this decision-making scenario, perhaps moral reasoning and maximizing the good of the team (i.e., ethical concern – Kohlberg, 1981) was important regardless of masculinity or femininity. Finally, in support of hypothesis 13a, risk propensity was positively correlated with the masculine scale, which was also found in Study 2 and replicates past literature in the realm of risk and gender (e.g., Meier-Pesti & Penz, 2008). On the contrary, while I hypothesized (hypothesis 13b) that femininity would be negatively correlated with risk, they were unrelated – this is in line with the findings in Study 2.

The present study results also offer some interesting findings and contributions with regard to *team* gendered decision-making styles, risk propensity and ethical responsibility, with the same pattern of results occurring in Study 2. Regression analysis was in line with the idea that team masculinity would be strongly and positively related to a team's risk propensity, and team femininity would negatively contribute to risk propensity. Likewise, the interaction effect between masculinity and femininity in the regression did not make a significant contribution to risk propensity, thus, hypothesis 14a

regarding masculine and androgynous teams exhibiting greater risk propensity was partially supported.

Broadly speaking, findings pertaining to the linkages between team androgyny and team performance did not support the proposed pattern (hypothesis 14d). While the regression analysis demonstrated significant main effects to explain performance, the interaction effect did not contribute. Therefore, the present study results suggest that team's gendered decision-making styles, specifically being masculine, played a relevant role in performance, showing an inverse relationship, whereby masculinity was negatively related to performance. Likewise, when adding the control variable team sex to the equation, male teams also decreased performance in comparison to mixed sex teams (female teams did not contribute to the model but were in the positive direction). Similar to studies work settings (e.g., Maccoby, 1990; Moscowitz et al., 1994) and politics (e.g., Mendelberg & Kapowitz, 2016), the present study also showed that male teams and female teams were significantly higher in masculinity than mixed sex teams, and female and mixed sex teams were significantly higher in femininity than male teams, partially supporting hypothesis 15a and 15b. Interestingly, female teams were high in both masculinity and femininity (i.e., androgyny), and mixed sex teams only high in femininity. Consequently, the interplay of the present findings surrounding gender and sex are worthwhile and warrant broader discussion.

Specific to team sex and performance, ample evidence has suggested that mixed sex teams may perform better than all male or female teams depending on the context and task. Hoogendoorn, Oosterbeek, & van Praag (2011), for example, conducted a field experiment using teams of undergraduate students randomized into all male, all female, and mixed sex teams, and found that mixed sex teams outperformed male teams but not

female-dominated teams. Likewise, Lee & Fahr (2004) examined group efficacy, group effectiveness (i.e., project grade) and gender diversity using student teams and found that mixed sex groups positively related to group effectiveness (e.g., performance), while same sex groups decreased performance. While these examples do not consider gender roles per se, their conclusions are in line with my findings regarding male and mixed sex teams.

A few studies have *indirectly* demonstrated the link between team "gender", sex, and performance, most noteworthy being research conducted by Woolley and associates. For instance, Bear & Woolley (2011) strongly suggest that team collaboration and participation (e.g., arguably feminine), increases with more women on a team, which also has a positive effect on performance (see also Woolley et al., 2010). In a similar vein, Fenwick & Neal (2001) investigated sex composition on group performance over a period of three years and found that groups with either equal number of men and women *or* a greater number of women ranked higher (i.e., performance and report marks) than all-men groups; The authors justification for these findings related to females' likelihood to incorporate co-operation and collaboration (e.g., feminine processes). Arguably, the findings of Woolley et al. (2010), Bear & Woolley (2011), and Fenwick & Neal (2001) are consistent with the present study in that mixed sex teams, *may* actually be more feminine, positively contributing to objective performance. Subsequently, high team masculinity may have related to a decline in performance in this situation because it lessened the ability to be collaborative, and thus impaired the open communication and participation required to complete the task (i.e., Winter Survival Exercise). Likewise, Rosenthal (2000) examined the gender roles of groups and found that both male and females scores on a competing style competing style (e.g., masculine) declined as the

percentage of women increased – this may also relate to why mixed sex teams in Study 3 were more likely to be feminine in their decision-making.

My results surrounding female teams as being more feminine *and* masculine (i.e., androgynous) is also consistent with individual level findings regarding female leaders adopting more masculine traits, while the same cannot be said for men in the utilization of feminine traits (Rodler, Kirchler, & Hoelzl, 2001). Likewise, Schein's (2001) studies examining the "think manager, think male" phenomenon also demonstrated that women are more likely to employ more androgynous leadership roles in comparison to men. Additionally, Berkley, Morley, & Tiernan (2013) conducted a self-report study of over 1200 students regarding the perception of gender and the managerial role. The authors found that male managers continue to be portrayed as masculine, but female managers as androgynous. For that reason, the results I found at the team level of analysis are not necessarily surprising.

Limitations and Future Research

In addition to the limitations discussed in Study 2 (e.g., common method bias, generalizability, control variables etc.), a few others warrant mentioning in Study 3. First, the internal consistency reliability of the masculine scale poses some discussion. Initially, I wondered if some items were unclear because the masculine scale corresponds with more masculine leadership styles, in turn, requiring some autocratic behaviors (e.g., Eagly & Johnson, 1990) and risk (that may have not been relevant in the WSE). The feminine scale, was also lower than in Study 2; therefore, leading me to question if the lower scale reliabilities were due to the use of the decision-making scenario the Winter Survival Exercise in conjunction with the short amount of time that the teams actually worked together. However, they were still in the acceptable range, and I argue that the

experimental design using a decision-making scenario was necessary at this point in my research program to be able to examine both sex and gender of teams in a controlled manner. Building on that, the confirmatory factor analysis showed an acceptable fit to the data, but the results indicate further exploration. Therefore, future research may be required in order to finalize the factor structure for use with organizational teams.

Second, the sample consisted of student teams, again, only working together for one hour, and completing a hypothetical scenario unrelated to organizational decision-making. This may be limiting given student's short duration with one another and because the scenario may not represent real-life; however, the majority of students *voluntarily* chose to participate in this study using the SONA system, with a high response rate. Therefore, while duration of time together may be of concern, the student teams were very involved in the exercise and they told me that it felt "real" for them as they participated. Additionally, it is unlikely that I would have been able to find an organization with enough teams performing the same task (Lee & Fahr, 2004), as such, allowing me to use an actual objective performance outcome. Future research may improve on design features by using a decision-making scenario more relatable to organizations today, and one that requires students to work together over a longer period of time making, completing multiple tasks (e.g., Fenwick & Neal, 2001 – MARKSTRAT business simulation).

Third, while every effort was taken to ensure randomization of the sex composition of the teams, it was quite challenging given students signed up independently via the SONA system, and within the classroom some students self-selected their teams. However, I was able to get a relatively equal number of all male, all female, and mixed sex teams, which has found to be difficult in other past studies

examining performance and student teams (e.g., Fenwick & Neal, 2001). Replication of this study in the future would also include a larger sample size with more teams representing each sex composition to confirm present findings.

5: Overall General Discussion and Conclusions

Overview of the Research Problem and Key Research Questions

Given the recent attention to biological sex and representation on decision-making bodies in conjunction with inconsistent findings regarding performance, sex, and organizational teams, understanding how sex *and* gender impact on team performance is important. Understanding team dynamics in regard to sex and gender may be beneficial at any organizational level – teams are the forefront of every day practice (e.g., Mathieu, Maynard, Rapp, & Gilson, 2012). There has been a great amount of research examining the sex composition of *teams*, decision-making, and performance outcomes; the same cannot be said in regard to the interplay between sex *and* gender (i.e., masculinity, femininity, androgyny, communal, agentic, etc.). Traditionally, gender has been examined at the individual level of analysis. Research has either studied gender roles relative to everyday life (i.e., well-being, self-esteem etc.) or in organizational research. Most commonly research studies have investigated gender roles of individual leaders. On the other hand, two of the most pervasive issues faced by organizations, risk and ethics, have been researched consistently in regard to *sex* at the individual level and/or team level of analysis. Although inconclusive, there is a general consensus that risk taking is associated with masculine behaviors and ethics with feminine behaviors, often assuming that "male" and "masculine" or "female" and "feminine" are synonymous. Furthermore, given that "good" decision-making would likely consider both risk *and* ethics (Disparte, 2016), I felt it was imperative to consider risk and ethics in my construction of masculinity and

femininity at the team level of analysis. Specifically, because past literature has tended to examine these constructs in connection with biological sex rather than considering gender. Accordingly, to fill this gap, my thesis introduced a new team level measure, the team androgyny instrument, for use in examining organizational decision-making and performance. The objectives of this research program were to: (1) explore and develop the measurement of masculinity and femininity at the team level of analysis for specific use in a business context; (2) examine the relationship between team masculinity, femininity, and androgyny and constructs related to risk propensity and ethical responsibility; (3) investigate if team androgyny predicts objective performance; (4) test the interplay of team sex and team gender on objective performance.

Summary of Studies 1, 2, and 3

My dissertation consisted of a pilot study and three separate studies. In study 1, a combination of scale development processes was utilized. First of all, after an extensive literature review and consultation with subject matter experts on potential past individual/team level measures (e.g., gender roles, risk propensity, ethical/moral responsibility, collaboration, group characteristics, conflict etc.), a pool of 64 items was developed. A panel of subject matter experts reviewed the item pool and classified each item as masculine or feminine. A pilot study was conducted to test the initial factor structure of the instrument as well as reduce the item pool. The pilot study provided initial support for the masculine and feminine scales; as such, a larger exploratory study was conducted with the reduced item pool prior to confirmatory analysis. Overall, this study supported content validity of my team androgyny instrument by exploring several facets of gendered decision-making styles, using subject matter experts to ensure the items

represented the broad repertoire of gendered decision-making, and through two exploratory principal components analyses.

In Study 2, undergraduate students required to work in teams for the entire semester completed a questionnaire that included the team androgyny instrument, measures related to risk propensity, ethical responsibility, work group participation, and counterproductive behaviors, as well as a performance outcome of team project grades. This study provided evidence of convergent and divergent validity for the team androgyny instrument. The team androgyny instrument factored into two hypothesized scales and the feminine and masculine scales were correlated with relevant constructs (i.e., ethical responsibility and risk propensity, respectively). Additionally, at the team level of analysis, even after controlling for specific variables, the feminine and masculine scales were related to risk and/or ethics, as hypothesized. However, neither masculinity, femininity, nor androgyny (i.e., interaction effect) significantly contributed to performance. On the contrary, when considering the ANOVA models (see Appendix I) in conjunction with performance, although not significant, masculine teams had the highest overall grades whereas androgynous, the lowest grades. This poses some interesting questions in regard to team gender *and* performance given the findings of Study 3.

Study 3 was meant to replicate findings from Study 2; however, I used an experimental design in a controlled setting which also allowed me to link team gender and sex composition to outcomes such as performance in the Winter Survival Exercise. This study incorporated a similar questionnaire as in Study 2 that respondents completed after participating in the hypothetical scenario. Indeed, like Study 2, this study showed convergent and divergent validity related to risk propensity and ethical responsibility. As well, the factor structure was confirmed, whereby a masculine and feminine scale showed

an acceptable fit to the data. This study also provided evidence to suggest that sex *and* gender of teams might have an effect on objective performance outcomes. Specifically, even after considering the control variables, all male teams *and* masculine teams had a negative effect on performance. Additionally, when examining the means of performance in the ANOVA models (see Appendix M), Study 3 revealed that feminine teams had the highest ranking, and like Study 2, androgynous teams performed the worst. Interestingly, as found in individual level of analysis leadership studies, this study also provided support in regard to sex composition and team gender in that male teams are highly masculine but feminine teams are significantly masculine and feminine. However, mixed sex teams were only significantly feminine. This study contributes to the literature on ethical responsibility, risk propensity, and performance because no other studies to date have examined the effect on *team* sex and gender.

Summary of Important Findings

According to Kachel et al. (2016) "gender research has developed many instruments to measure different aspects of self-ascriptions of gender stereotypical features, including attributes, behaviors, interests, and attitudes" (p. 15); However, to date, no instrument has been developed to measure gendered decision-making styles of organizational teams. Below are some overall important findings that show support for the team androgyny instrument.

In line with gender researchers such as Bem (1974), Spence et al. (1974), Abele (2003), Blasberg et al. (2013), and Trapnell & Paulhus (2011), the weak but positive correlations between the team androgyny femininity scale and masculinity scale (as shown in Study 2 and Study 3) suggest that they are relatively independent constructs, and masculinity and femininity *may* occur on any team regardless of the sex composition

of the team. All items loaded onto two factors using the exploratory PCA and the ESEM, and the CFA using two factors demonstrated an acceptable fit to the data. These findings combined support other two-dimensional gender constructs (i.e., Bem, 1974 – masculine and feminine; Spence et al., 1973 – Instrumental and Expressive; Blasberg et al., 2013 – agentic and communal management etc.) and shows support for separate team masculine and feminine scales. Furthermore, the team androgyny instrument demonstrated acceptable internal consistencies across the studies.

Regarding validity, the team androgyny instrument showed expected correlations with risk propensity, ethical responsibility, and communality/agenticism. Specifically, risk propensity was positively associated with the masculine scale and unrelated to the feminine scale in both Study 2 and 3. Likewise, the feminine scale was positively correlated to ethical benevolence and negatively related to ethical egoism in both studies. On the contrary, while the masculine scale was unrelated to ethical egoism in Study 2 and 3, the scale was actually positively related to ethical benevolence in Study 3. Again, in hindsight this may be due to the nature of the decision-making scenario (e.g., requiring teams to be morally responsible and have concern for the team as a whole in an effort to survive). Moreover, it was expected that the team androgyny instrument would correlate with other measures of gender. Using a measure of communality/agenticism showed that the masculine scale was positively correlated with the agenticism scale, and uncorrelated with the feminine scale. However, quite interestingly, the feminine scale was moderately and positively correlated with communal scale *and* weakly positively related to the agenticism scale.

At the team level of analysis, the overall findings regarding risk propensity and ethical responsibility (specifically ethical benevolence) have two important implications.

First, the results of my regression analysis in Study 2 and Study 3 demonstrate that there may be a *gender* component in regards risk and team masculinity, and ethical responsibility and team femininity. More specifically, in both studies while the control variables of biological sex or sex composition were not significantly related to risk or ethics, team *gender* was. In several instances throughout the literature, at the individual level and team level of analysis, women have shown to be more risk averse than their male counterparts whereas other studies argue that there are no differences. Generally speaking, risk propensity is regarded as a masculine characteristic (Meier-Pesti & Penz, 2008); However, these studies have consistently examined biological sex (female and male), as opposed to gender (feminine and masculine), and assume the constructs are identical. Likewise, past literature has also argued *sex* differences in terms of ethical responsibility, specifically as it relates to moral responsibility and ethics of care/concern for the organization, in that it is viewed as more feminine; However, other researchers have found that ethical responsibility may not correlate well with biological sex (e.g., Jaffee & Hyde, 2000). Finally, in terms of ethical egoism, the same findings emerged in Study 2 and Study 3 showing a positive contribution from team masculinity and a negative contribution from team femininity. Ethical egoism, which generally refers to maximizing self –interests (Kohlberg, 1981), could arguably tap into elements of masculinity given its more autocratic/directive nature in the realm of leadership. Overall, these findings provide support for past inconclusive literature regarding biological sex and ethics/risk by demonstrating the potential link to *gender*, specifically at the team level of analysis.

Also noteworthy, is the relationship between team sex composition and team gendered decision-making styles. In both the gender and leadership literature, studies

have demonstrated how women have become increasingly androgynous over the past few decades. For example, in a meta-analysis examining changes in masculine and feminine traits over a span of 20 years, Twenge (1997) found that while there was no change in women's femininity score (using the BSRI and PAQ), their scores in masculinity have increased. Additionally, men's scores in regard to masculinity showed an upward trend whereas their feminine scores did not change overtime. Likewise, as I discussed in my theoretical justification of this dissertation, at the individual level of analysis, the leadership literature has suggested that androgynous leadership is increasingly important (van Engen, 2012) and female leaders are more likely to employ masculine and feminine leadership styles in comparison to male leaders (Rodler, Kirchler, & Hoelzl, 2001). My results showed a similar pattern at the team level of analysis, whereby male teams were significantly more masculine and female teams were significantly higher in both femininity and masculinity. Taking these findings alone, I would be inclined to argue that team sex composition influenced the gendered decision-making styles of the team; however, my results regarding mixed sex teams showed that they were significantly higher in femininity (but not masculinity), while quite interesting, this was not expected. Evidence in support of this particular finding may be attributed to team-based research examining collaboration and sex composition, whereby mixed sex teams, specifically women, increase the social sensitivity of the team, encourage greater conversational turn taking, and more collaboration, which tends to be viewed as feminine (Bear & Woolley, 2010; Woolley et al., 2010).

In evaluating team performance, team gender showed some intriguing findings. While there were no significant effects in Study 2, the results of the ANOVA models indicated highest performance from masculine teams and lowest performance from

androgynous teams (e.g., team grades). On the contrary, in Study 3, the main effects of the regression demonstrated that team masculinity negatively contributed to team performance and when adding the control team sex to the model, all male teams also had a negative relationship with team performance in comparison to mixed sex teams. According to the regression, team femininity and androgyny did not significantly contribute to performance. When examining the ANOVA models, in Study 3, feminine teams had the highest performance, and again, androgynous teams had the lowest performance.

Consequently, while the findings do not support my hypotheses surrounding androgyny and performance, they do demonstrate that situational adaptability *may* be required depending on the decision-making task. As evident in Study 2, which involved team papers/presentations, masculine teams had higher performance than feminine/androgynous teams, whereas in Study 3, which utilized the Winter Survival Exercise, feminine teams obtained higher rankings (e.g., relative to survival) than masculine/androgynous teams. In both studies, androgynous teams actually had the lowest performance rankings. Evidently, these results show that when analyzing the sex and gender of teams' relative to performance, a variety of contextual factors come in to play (Bear & Wooley, 2010). These include other forms of demographic diversity (Pelled, Eisenhardt, & Xin, 1999) (e.g., ethnicity), type of industry (Joshi & Roh, 2009) (e.g., many participants were psychology majors), and most importantly, type of task (Bowers, Pharmer, & Salas, 2000).

If we revisit the definition of team androgyny (i.e., a team that perceives they are using both masculine *and* feminine decision-making styles), it is apparent that androgynous teams *may* have had the poorest performance because one decision-making

task (e.g., team project) required more masculine behaviors (and perhaps risk taking) and the other (e.g., Winter Survival Exercise), more feminine behaviors (and perhaps ethical responsibility). As such, they were not required to use *both* masculine and feminine decision-making styles in those particular tasks, nor were they required to *balance* ethics and risk or adapt to different tasks over time. Although further analysis is required to test this assumption, I feel that I can still argue the importance of *flexibility* in terms of femininity and masculinity decision-making styles on teams, given the depth and complexity of decisions that organizational teams are required to make. Moreover, I argue that the easiest way to achieve flexible gendered decision-making would be through sex balance.

Overall, the present findings offer some valuable insights in specific literatures on gender, leadership, decision-making and/or performance, and teams. My program of research is the first (to my knowledge) to explore the complex interplay of gender, sex composition, risk propensity, ethical responsibility, and performance at the team level of analysis. Therefore, I contributed to beginning to unravel how changing the sex and or gender of a team might impact on its decision-making at the behavioural level. While team androgyny did not lead to better performance than predominately masculine and feminine teams, I did find support in that team gendered decision-making *and* team sex composition does have an impact on performance, and that the task at hand plays an important role in gendered decision-making style employed.

Limitations and Future Research

While I have previously discussed limitations and future research for each study, in this section I outline several limitations and recommendations for future studies that are relevant to the entire dissertation.

One limitation may be the conceptualization and measurement of team androgyny (i.e., a team that perceives it is using both masculine and feminine decision-making styles) for use in an experimental scenario-based design. Mainly because it might be quite difficult to find a scenario that would allow a team to utilize both masculine and feminine gendered decision-making styles (e.g., risk and ethics). Although some may argue that the Winter Survival Exercise is more masculine in nature (e.g., Rogelberg & Rumery, 1996) due to its focus on survival tactics, I speculate that team masculinity negatively contributed to performance because of its emphasis on "competition," "risk," and "assertiveness." Moreover, despite using teams that were required to work together for an entire semester in Study 2, their project/presentation may not have allowed the opportunity to develop as a team. Likewise, in Study 3, they only worked together for one 1.5 hours. Therefore, future research could employ a different design strategy to allow teams to work together over an extended period of time using a more complex business simulation that would require different types of decision-making over time. For instance, if using university students, a pre and post experimental design may provide interesting results; specifically ensuring that the chosen simulation requires students to work together consistently for a sufficient amount of time, collecting data shortly after team formation and then again when the team dissolves. Another recommendation would be to incorporate multiple simulations/scenarios asking teams to fill out the team androgyny instrument more generally as opposed to thinking of it in terms of one scenario, this may better capture their gendered decision-making style as a team overall. A final suggested future study related to an experimental design would be to incorporate some form of videotaped recording of team decision-making processes and analyzing their use of risk propensity, ethical responsibility, and other variables of interest while completing a

scenario/simulation. Therefore, employing both a qualitative and quantitative investigation to develop more reliable conclusions.

Additionally, I previously mentioned that using a student sample in my program of research might be a limitation in terms of generalizability of findings. Students are readily accessible and allow for more control of context, as I was able to use pre-existing teams within the classroom (Study 2) and the SONA system (Study 3). Given my intriguing findings, replication using actual teams within organizations would be ideal, but it would be quite difficult to find a representative sample based on sex composition (i.e., all male, all female, and mixed-sex teams). However, it may be practical to further explore the factor structure of the team androgyny instrument using individuals who are members of organizational teams to see if the internal consistency of Study 3 was an issue related to the Winter Survival Exercise or the measure itself. It is possible that items may need to be revisited or revised to ensure the construct represents all facets of masculine and feminine team decision-making. Once the measure is finalized, future research should examine longitudinal data to understand how team gendered decision-making may change overtime and/or to explore test-retest reliabilities. Again, given that in my studies students may have not had a sufficient amount of time to work together, employing longitudinal experimentation may be needed to uncover the relationship between team gender and performance.

Future research should also focus on the interconnected relationship that seems to exist between risk propensity, ethical responsibility, and gendered decision-making at the team level of analysis. While outside the scope of this program of research, future studies might use moderation analysis, structural equation modeling (e.g., path analysis) or

multilevel modeling (i.e., incorporating both individual level attributes and team level factors) to more rigorously statistically analyze the effects.

Third, a dilemma faced throughout this dissertation related to scoring procedures for team androgyny. Traditionally, and most commonly, researchers such as Bem (1974) and Spence et al. (1974) have recommended median split classification procedures in analyzing the gender roles of their representative measures. Hypothetically speaking, dichotomizing individuals (or teams) into appropriate gendered categorical groups seems logical and simplistic (Hoffman & Borders, 2001); However, using median splits has some psychometric concerns as it often results in a loss of information, in addition to the assumption that femininity and masculinity are independent dimensions (Burn, O'Neil, & Nederend, 1996). With the use of median splits, those who score below the mean in both masculinity and femininity are often excluded from further analysis, resulting in loss of data. Most importantly, median cutoffs will likely be different with every sample used, thus, androgyny would also differ, making it difficult to generalize across the population (Sedney, 1981). To ameliorate this issue, several researchers have recommended using multiple regression analysis whereby androgyny is the product of the masculinity and femininity scores (Holmbeck, 1989) entered in a stepwise fashion. This is particularly useful for studies interested in predicting dependent variables on the basis of masculinity and femininity because variables remain continuous. Therefore, while I included the median split classification to be thorough in my analysis (see Appendix I and M), my discussion of findings are based off of the results of the regression technique. In an ideal world, future research would include the development of a unique scoring procedure that eliminates psychometric issues (although I realize that that is not likely possible, as no statistical procedure is perfect).

Finally, given the critiques Bem faced regarding the use of the terms "masculinity" and "femininity," and the more fundamental issues regarding the changing perception of gender roles overtime, I see this as potential limiting in my dissertation as well. However, this may be a relatively simple fix because my team androgyny instrument is not based on gender stereotypical traits in the broader cultural sphere, rather relate specifically to decision-making on organizational teams. Additionally, in Study 3, I incorporated a measure of communality and agenticism to test the links to my team androgyny measure. The findings showed that they were correlated. I originally included the constructs of communion and agenticism to potentially begin to veer away from more stereotypical terms such as femininity and masculinity. As such, future research may investigate changing the language to avoid the issue of gender-stereotypical thinking. As Hoffman & Borders (2001) stated, "naming is a powerful phenomenon that can serve to maintain rather than ameliorate a dichotomy, and with it, a status quo" (p. 53).

Practical Implications

From a practical perspective, different implications can be drawn from this dissertation for organizations trying to implement sex/gender balance of their respective teams, as well as those aiming to balance decision-making as it relates to risk propensity and ethical responsibility.

Importantly, organizations should acknowledge that sex and gender are two distinct concepts, and both may play a factor in terms of performance at the team level of analysis. For example, in Study 3 while only two female teams were classified as masculine, the majority were classified as feminine or androgynous. On the contrary, male teams were highly concentrated as masculine or undifferentiated, and mixed sex as feminine or androgynous. Consequently, awareness of these differences may assist in

team development of specific organizational projects, specifically if certain projects require greater risk propensity or ethical responsibility, versus when a balance of the two is needed.

Additionally, demonstrating the links between team gender, risk propensity, and ethical responsibility not only supports federal legislation and/or recommendations surrounding gender quotas, indirectly, it shows how gender balance may also create smarter decision-making. Several examples of "bad" decisions apparent over the last decade not only resulted in extreme financial loss, but also eventual collapse and/or acquisition of the respective companies (e.g., Lehman Brothers, Firestone, Enron, etc.). Although examined from a more practical perspective, many of these practitioners argue some of the worst business decisions were caused by exaggerated risk without consideration of ethical codes of conduct (Disparte, 2016). If we revisit the Firestone example from earlier in the dissertation, we see that the company failed to acknowledge warning signs that their decision-making regarding new tires was too risky, eventually leading to their acquisition. Likewise, Volkswagen is now in a similar situation in terms of the company's environmental emissions scandal to increase profits (Bansal et al., 2015). Examples like these, although unrelated to gender/sex, demonstrate the potential benefits of being able to balance/adapt risk and ethics through sex balanced teams, that may also be more gender balanced.

Perhaps organizations may consider standardizing decision-making practices relevant to team gender. Therefore, not only encouraging sex/gender balanced teams but also implementing a comprehensive decision-making framework inclusive of masculine/feminine aspects of decision-making. As evident in the literature on gender and leadership, at the individual level particularly, decision-making style continues to be

biased, and favor masculinity (i.e., think-manager, think-male – Schein & Davidson, 1993). Likewise, at the team level, findings surrounding sex composition, risk propensity, and ethical responsibility seem to be mixed, this may be in part due to its lack of consideration for gender (as opposed to sex). By considering gender, I show the need to not only encourage sex balance, but also the importance of teams to be able to situationally adapt their gendered decision-making styles depending on the task in question. Neal & Spetzler (2015) noted that "almost three-quarters of companies have no formal corporate-wide approach to making major, complex decisions." By defining a "good" decision as consisting of elements of masculinity and femininity, and building a comprehensive decision-making process that is aligned with risk and value systems may indirectly assist in terms of equity in organizations (as it relates to sex).

In terms of my team androgyny measure itself, organizations aiming to be more inclusive regarding gender equity and/or valuing gender within their overall culture, could also use this for training and development purposes. Intact teams could take the team androgyny measure and it could be used as a diagnostic tool to identify teams current gendered decision-making styles. The results of the measure could then create a tailored training program aimed at improving their perceptions towards masculine (e.g., agentic) or feminine (e.g., communal) decision-making. Essentially, by focusing on the gendered style they were lowest in, and working towards improving on that style they might be able to make better decisions. Alternatively, organizational teams could complete the team androgyny measure and then undergo a comprehensive training program focused on awareness of gendered decision-making styles relative to risk and ethics. Additionally, this test, and training, may precede the introduction of a comprehensive decision-making framework (discussed above) relative to risk and ethical value systems as well as gender.

Doing so, will give teams the awareness to reflect and ask questions such as, "is this an appropriate context to be using this style of decision-making?"

Finally, returning to general support for sex based targets in organizational teams, the business case is obvious. As anticipated, my results did show performance benefits relevant to having more heterogeneous behaviours. In Study 2, masculine teams had the highest performance whereas in Study 3, feminine teams, had the highest performance. Likewise, in Study 3, results surrounding team sex composition (i.e., mixed sex as opposed to all male or female) demonstrated superior performance. Consequently, for teams to perform well, they may not need to be androgynous (i.e., using both styles in the same situation), but be able to flexibly adapt their style (i.e., being feminine in certain situations, and masculine in other situations). Overall, this may be an opportunity for organizations to fully utilize the decision-making capacity they have at hand – whether it comes in a male or female body.

Conclusion

In the first paragraph of this thesis, like many others, I asked "What if Lehman brothers had been Lehman Sisters?" with the hope of demonstrating the value of both the Lehman brothers *and* Lehman sisters. While I have come to realize that much research still needs to be done, I feel that my dissertation has begun to lay the groundwork for important research in the realm of sex, gender, and decision-making (i.e., risk and ethics) of organizational teams. Recall the Lehman Brothers case study from earlier in this dissertation, I demonstrated how from a masculine standpoint the team may have engaged in risk leading to unethical decisions, and from a feminine standpoint, avoidance of risky decisions might have hindered initial success but ensured ethical responsibility.

In sum, I assert this thesis has taken important steps towards a better understanding of team-gendered decision-making relative to risk propensity and ethical responsibility. It provided empirical evidence at the team level of analysis; thus, expanding the sex-based view and opening the doors for new lines of research that further advance these initial findings. Furthermore, given the relationship between team gender, risk propensity and ethical responsibility, utilizing them together as a mechanism to improve decision-making processes may not only create an awareness of the value of masculine *and* feminine decision-making styles, but also improve organizational outcomes such as performance.

I still have questions regarding team masculinity, femininity, androgyny and performance given that my results were non-significant in Study 2 (although masculine teams did have the highest objective performance in comparison to feminine and androgynous teams), and partially significant in my final study, whereby male and masculine teams obtained the lowest performance rankings on the Winter Survival Exercise (feminine teams ranking higher, androgynous ranking lowest). Therefore, this dissertation created new questions and illustrated important directions for future research. Consequently, although I cannot answer what would have happened if Lehman brothers had been Lehman brothers and Lehman sisters, I can argue that team gender *and* sex does impact on decision-making and performance, but how, depends on the task/situation in question. I hope that by beginning the conversation, both theoretically and empirically, future researchers will be able to further contribute and expand the link between performance, sex composition, and gendered decision-making styles of organizational teams.

References

Abele, A.E. (2003). The dynamics of masculine-agentic and feminine-communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology, 85*(4), 768-776.

Abele, A.E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93*(5), 751 – 763.

Adams, R.B. (2016). *Women on boards: the superheroes of tomorrow?* (Report No. 466/2016). Retrieved from European Corporate Governance Institute Website: http://www.ecgi.global/sites/default/files/working_papers/documents/SSRN-id2696804.pdf

Adams, R.B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics, 94*(2), 291-309.

Ahern, K., & Dittmar, A. (2012). The changing of the boards: The impact on firm valuation of mandated female board representation. *Quarterly Journal of Economics, 127*(1), 137-197. Doi: 10.1093/qje/qjr049.

Alonso-Almeida, M., & Bremser, K. (2015). Does gender specific decision making exist? *EuroMed Journal of Business, 10*(1), 1450 – 2194.

Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quarterly Journal of Economics, 116*, 293-312.

Apesteguia, J., Azmat, G., & Iriberri, N. (2012). The impact of gender composition on team performance and decision making: Evidence from the field. *Management Science, 58*(1), 78-93.

Arnold, J.A., Arad, S., Rhoades, J.A., & Drasgow, F. (2000). The empowering leadership questionnaire: the construction and validation of a new scale for measuring leader behaviors. *Journal of Organizational Behavior, 21*(3), 249-269.

Asparouhov, T., & Muthén, B. (2008). *Exploratory structural equation modeling.* Retrieved from Mplus website: https://www.statmodel.com/download/EFACFA810.pdf

Bampton, R., & Maclagan, P. (2009). Does a 'care orientation' explain gender difference in decision making? A critical analysis and fresh findings. *Business Ethics: A European Review, 18*(2), 180-191. Doi: 10.1111/j.1467-8608.2009.01556.x.

Bansal, T., King, M., & Seijts, G. (2015). The Volkswagen emissions scandal: A case study in corporate misbehavior. Retrieved from *the Globe and Mail* website: http://www.theglobeandmail.com/report-on- business/rob-commentary/the-vw-emissions-scandala-case-study-in-what-not-to-do/article26550100/

Barber, B.M., & Odean, T. (2001). Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics, 1*(1), 261 – 292.

Barry, B., & Stewart, G.L. (1997). Composition, process, and performance in self-managed groups: The role of personality. *Journal of Applied Psychology, 82 (1),* 62 – 78.

Bass, B.M., & Avolio, B. (1996). *Multifactor Leadership Questionnaire for Teams.* Redwood City, CA: Mind Garden.

Bauer, D.J., & Curran, P.J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research, 40 (3).* 373 – 400. Doi: 10.1207/s15327906mbr4003_5.

Bear, S., Rahman, N., & Post, C. (2010). The impact of board diversity and gender composition on corporate social responsibility and firm reputation. *Journal of Business Ethics, 97*(2), 207-221. Doi: 10.1007/s10551-010-0505-2.

Bear, J.B., & Woolley, A.W. (2010). The role of gender in team collaboration and performance. *Interdisciplinary Science Review, 36* (2), 146 – 153.

Becker, T.E., Atinc, G., Breaugh, J.A., Carlson, K.D., Edwards, J.R., & Spector, P.E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior, 37*, 157 – 167.

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

Bem, S.L. (1977). On the utility of alternative procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology, 43*(2), 196-205.

Bem, S.L. (1979). Theory and measurement of androgyny: A reply to the Pedhazur-Tetenbaum and Locksley-Cohen critiques. J*ournal of Personality and Social Psychology, 37*(6), 1047-1054.

Bem, S.L. (1985). Androgyny and gender scheme theory: A conceptual and empirical integration. In T.B. Sonderegger (Ed.), *Nebraska symposium on motivation 1984: Psychology and gender (Vol. 32, 179 – 226).* Lincoln: University of Nebraska Press.

Bem, S.L., & Lewis, S.A. (1975). Sex role adaptability: One consequence of psychological androgyny. *Journal of Personality and Social Psychology, 31*(4), 634 – 643.

Berdahl, J.L. (1996). Gender and leadership in work groups: Six alternative models. *Leadership Quarterly, 7*(1), 21-40. Doi: 10.1016/S1048-9843(96)90033-8.

Berkery, E., Morley, M., & Tiernan, S. (2013). Beyond gender role stereotypes and requisite management characteristics: From communal to androgynous, the changing views of women. *Gender in Management: An International Journal, 28*(5), 278 – 298.

Biermann, T., Cole, M.S., & Voelpel, S. (2012). Within-group agreement: On the use (and misuse) of rwg and rwg(j) in leadership research and some best practice guidelines. *The Leadership Quarterly, 23*(1), 66 – 80.

Blasberg, S.A., Rogers, K.H., & Paulhus, D.L. (2013). The Bidimensional impression management index (BIMI): Measuring agentic and communal forms of impression management. *Journal of Personality Assessment, 96*(5), 523 – 531. Doi: 10.1080/00223891.2013.862252

Bliese, P.D. (2000). Within-group agreement, non-independence, and reliability. Implications for data aggregation and analysis. In K.J. Klien, & S.W.J. Kozlowsky (Eds.). *Multilevel theory, research and methods in organizations. Foundations, extensions, and new directions,* 349 – 381. San Francisco, CA: Jossey – Bass.

Bliese, P. D., & Halverson, R. R. (1998). Group consensus and psychological wellbeing: A large field study. *Journal of Applied Social Psychology, 28*(7), 563 – 580. Doi: 10.1111/j.1559-1816.1998.tb01720.x

Bogan, V.L., & Just, D.R. (2013). Team gender diversity and investment decision-making behavior. *Review of Behavioral Finance, 5*(2), 134 – 152.

Booth, A.L. & Nolen, P. (2012). Gender differences in risk behaviour: Does nurture matter? *The Economic Journal, 122*, F56-F78. Doi: 10.1111/j.1468-0297.2011.02480.x.

Bowers, C. A., Pharmer, J., and Eduardo, S. (2000). When member homogeneity is

needed in work teams: A meta-analysis. *Small Group Research 31*(3), 305 – 327.

Brenner, O.C., Tomkiewicz, J., & Schein, V. (1989). The relationship between sex-role stereotypes and requisite management characteristics revisited. *Academy of Management Journal, 32*(3), 662-669.

Bukowski, W.M., Panarello, B., & Santo, J.B. (2016). Androgyny in liking and in being liked are antecedent to well-being in pre-adolescent boys and girls. *Sex Roles, 76*(11), 1-12. Doi: 10.1007/s11199-016-0638-6.

Bren, S.M., O'Neil, K., & Nederend, S. (1996). Childhood tomboyism and adult androgyny. *Sex Roles, 34* (5 – 6*)*, 419 – 428.

Brink, A.G., Hobson, J.L., & Stevens, D.E. (2017). The effect of high power financial incentives on excessive risk-taking behavior: An experimental examination. *Journal of Management Accounting Research, 29*(1), 13-29.

Browne, M.W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioral Research, 36*(1), 111 – 150.

Byrne, B.M. (1998). *Structural Equation Modeling with LISREL, PRELIS and SIMPLIS: Basic Concepts, Applications and Programming*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Cadsby, C.B., Servatka, M., & Song, F. (2010). Gender and generosity: does degree of anonymity or group gender composition matter? *Experimental Economics, 13*(3), 299-308.

Campion, M.A., Medsker, G.J., & Higgs, A.C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*(4), 823 – 850. Doi: 10.1111/j.1744-6570.1993.tb01571.x.

Carmeli, A., Tishler, A., & Edmondson, A.C. (2011). CEO relational leadership and strategic decision quality in top management teams: The role of team trust and learning from failure. *Strategic Organization, 10*(1), 31-54. Doi: 10.1177/1476127011434797

Carter, A., & Yeqing, B. (2005). Exploring the concept and measurement of general risk aversion. In NA-Advances in Consumer Research Volume 32. (Eds). Geeta Menon & Akshay R. Rao, Duluth, MN: Association for Consumer Research, pg. 531-539.

Catalyst. (2003). *The Catalyst Census of Women Board Directors of the Fortune 1000*. New York, NY: Catalyst.

Catalyst. (2007). *2007 Catalyst census of women board directors of the Fortune 1000*. New York, NY: Catalyst.

Castillo, M., Leo, G., & Petrie, R. (2012). Room effects, working paper, George Mason University, Fairfax, VA.

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: a typology of composition models. *Journal of Applied Psychology*, *83*(2), 234 – 246.

Chan, D. (2009). So why ask me? Are self-report data really that bad? In C.E. Lance & R.J. Vandenberg (Ed). *Statistical and Methodological Myths and Urban Legends* (pp. 309 – 335). New York, NY: Routledge.

Chan, L.M., Jamilah, O., & Rusinah, J. (2012). Gender differences of human resources practitioners in their ethical decision making. *International Journal of Academic Research*, *4*(4), 38 – 44.

Chen, G., Mathieu, J.E., & Bliese, P.D. (2004). A framework for conducting multi-level construct validation. *Research in Multi-Level Issues, 3*(1), 273 – 303.

Chiu, C.J. (2014). Development and validation of performance assessment tools for interprofessional communication and teamwork (PACT). *A dissertation - University of Washington.*

Churchill, G.A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*(1), 65 – 73.

Clark, C. (2013, July 25). Research shows profitable companies have more women on the board. *Financial Times.* Retrieved from https://www.ft.com/content/2d320562-f52d-11e2-94e9-00144feabdc0

Clark, L.A., & Watson, D. (1995). Constructing Validity: basic issues in objective scale development. *Psychological Assessment, 7*(3), 309-319.

Cook, J.D., Hepworth, S.J., Wall, T.D., & Warr, P.B. (1981). *The experience of work.* San Diego: CA: Academic Press.

Costa, A.C. & Anderson, N. (2011). Measuring trust in teams: Development and validation of a multifaceted measure of formative and reflective indicators of team trust. *European Journal of Work and Organizational Psychology, 20*(1), 119-154.

Cullen, J.B., Victor, B., & Bronson, J.W. (1993). The ethical climate questionnaire: an assessment of its development and validity. *Psychological Reports, 73,* 667 – 674.

Cullen, J.B., & Victor, B.I. (2008). The ethical climate questionnaire: An assessment of its development and validity. *Psychological Reports, 73,* 667-674. Doi: 10.2466/pr0.1993.73.2.667.

Deo, M., & Sundar, V. (2015). Gender difference: investment behavior and risk taking.

SCMS Journal of Indian Management,12(3), 74-81.

Donnelly, K., & Twenge, J.M. (2016). Masculine and feminine traits on the Bem sex role inventory, 1993 – 2012: a cross-temporal meta-analysis. *Sex Roles, 76*(9-10), 1-10. Doi: 10.1007/s11199-016-0625-y.

Doty, D. H., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? *Organizational Research Methods, 1*(4), 374-406. Doi: http://dx.doi.org/10.1177/109442819814002

Dufwenberg, M. & Muren, A. (2006). Gender composition in teams. *Journal of Economic Behavior and Organization, 61*(1), 50-54.

Eagly, A.H. (2005). Achieving relational authenticity in leadership: Does gender matter? *Leadership Quarterly, 16*(3), 459-474. Doi: 10.1016/j.leaqua.2005.03.007.

Eagly, A. H, & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin, 108*(2), 233-256.

Ertac, S., & Gurdal, M. (2012). Deciding to decide: gender, leadership and risk-taking in groups. *Journal of Economic Behavior & Organization, 83*, 24-30. Doi: 10.1016/j.jebo.2011.06.009.

Everitt, S. (1975). Multivariate analysis: The need for data and other problems. *British Journal of Psychiatry, 126,* 237 – 240.

Farrell, M. (2013, September 17). *5 years after Lehman: Where are they now?* Retrived from CNN: http://money.cnn.com/gallery/investing/2013/09/13/financial-crisis-executives-now/index.html

Fenwick, G.D., & Neal, D.J. (2001). Effect of gender composition on group performance. *Gender, Work, and Organization, 8* (2), 205 – 225.

Foushee, H.C., Helmreich, R.L., & Spence, J.T. (1979). Implicit theory of masculinity

and femininity: dualistic or bipolar? *Psychology of Women Quarterly, 3*(3), 259 – 269.

Galbreath, J. (2011). Are there gender-related influences on corporate sustainability? A study of women on boards of directors. *Journal of Management & Organization, 17*(1), 17- 38.

Gershenoff, A.B., & Foti, R.J. (2003). Leader emergence and gender roles in all-female groups: A contextual examination. *Small Group Research, 34*(2), 170 –196.

Gervais, S., & Odean, T. (2001). Learning to be overconfident. *The Review of Financial Studies, 14*(1), 1-27.

Glover, S.H., Bumpus, M.A., Sharp, G.F., & Munchus, G.A. (2002). Gender differences in ethical decision making. *Women in Management Review, 17*(5), 217 – 227.

Goktepe, J.R., & Hall, D.T. (1976). Inheriting a career: The influence of sex, values, and parents. *Journal of Vocational Behavior, 8*(1), 19-30.

Groysberg, B., & Bell, D. (2013). Dysfunction in the boardroom. *Harvard Business Review.* Retrieved from https://hbr.org/2013/06/dysfunction-in-the-boardroom

Guth, W., Schmidt, C., & Sutter, M. (2007). Bargaining outside the lab – a newspaper experiment of a three-person ultimatum game. *The Economic Journal, 117*(518), 449 – 469. Doi: 10.1111/j.1468-0297.2007.02025.x

Hall, J.A., & Taylor, M.C. (1985). Psychological androgyny and the masculinity x femininity interaction. *Journal of Personality and Social Psychology, 49* (2), 429 – 435.

Harris, C.R., Jenkins, M., & Glaser, D. (2006). Gender differences in risk assessment: Why do women take fewer risks than men? *Judgment and Decision Making, 1*(1), 48-63.

Heilbrun, A.B. (1976). Measurement of masculine and feminine sex role identities as independent dimensions. *Journal of Consulting and Clinical Psychology, 44*(2), 183-190. Doi: 10.1037/0022-006X.44.2.183.

Heilman, M.E., & Okimoto, T.G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology, 92* (1), 81 – 92.

Hinkin, T.R. (1995). A review of scale development practices in the study of organizations. *Journal of Management, 21*(5), 967 – 988.

Hinkin, T.R., Tracey, J.B., & Enz, C.A. (1997). Scale construction: Developing reliable and valid measurement instruments. *School of Hotel Administration Collection.* Retrieved from http://scholarship.sha.cornell.edu/cgi/viewcontent.cgi?article=1618&context=articles

Ho, S.S.M., Li., Y. Tam, K., & Zhang, F. (2015). CEO gender, ethical leadership and accounting conservatism. *Journal of Business Ethics, 127*(2), 351-370. Doi: 10.1007/s10551-013-2044-0.

Hoffman, R.M., & Borders, L.D. (2001). Twenty-five years after the bem sex role inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counseling and Development, 34*(1), 39-55.

Hoogendoorn, S., Oosterbeek, H., van Praag, M. (2011). The impact of gender diversity on the performance of business teams: evidence from a field experiment. *Tinbergen Institute*, 1 – 30.

Holmbeck, G. N. (1989). Masculinity, femininity, and multiple regression: Comment on

Zeldow, Daugherty, and Clark's "masculinity, femininity, and psychosocial adjustment in medication students: A 2-year follow up. *Journal of Personality Assessment, 53* (3), 583- 599.

Holt, C., & Ellis, J.B. (1998). Assessing the current validity of the bem sex role inventory. *Sex Roles, 39,* (11/12), 929- 941. Doi: 10.1023/A:1018836923919

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods, 6(1),* 53 – 60.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1 – 55.

Hugelschafer, S., & Achtziger, A. (2014). On confident men and rational women: it's all on your mind(set). *Journal of Economic Psychology, 41*(1), 31-44.

Isidro, H., & Sobrel, M. (2015). The effects of women on corporate boards on firm value, financial performance, and ethical and social compliance. *Journal of Business Ethics, 132*(1), 1-19. Doi: 10.1007/s10551-014-2302-9.

Ivanova-Stenzel, R., & Kübler, D. (2009). *Gender differences in team work and team competition.* Retrieved from https://www.wzb.eu/sites/default/files/+wzb/mp/vam/genderdifferences.pdf

Jackson, D.L., & Gillaspy, J.A. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods, 14*(1), 6 – 23.

Jaffee, S., & Hyde, J.S. (2000). Gender differences in moral orientation: A meta-analysis. *Psychological Bulletin, 126*(5), 703 – 726.

James, R.J., Demaree, R.G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69*(1), 85 – 98. Doi: 10.1037/0021-9010.69.1.85

James, R.J., Demaree, R.G., & Wolf, G. (1993). Rwg: An assessment of within-group interrater agreement. *Journal of Applied Psychology, 78*(1), 306 – 309. Doi: 10.1037/0021-9010.78.2.306

Jehn, K. A. & Mannix, E. A. (2001). The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *The Academy of Management Journal, 44*(2), 238-251.

Jeong, S.H., & Harrison, D. (2016). Glass breaking, strategy making, and value creating: meta-analytic outcomes of females as CEOs and TMT members. *Academy of Management Journal*, DOI:10.5465/amj.2014.0716

Johnson, M.E. (1988). Influences of gender and sex role orientation on help-seeking attitudes. *The Journal of Psychology, 122*(3), 237-241. Doi: 10.1080/00223980.1988.9915510

Johnson, D. W., & Johnson, F. P. 1982. *Joining together (2nd ed.)*. Englewood Cliffs, NJ: Prentice Hall.

Johnson, D. H., & Johnson, F.P. (Ed.). (2012). *Joining Together: Group Theory and Group Skills* (Vol. 11). Essex, UK: Pearson.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta analytic review. *Academy of Management Journal, 52* (3), 599 – 627.

Jule, A. (2004). *Gender, participation, and silence in the language classroom sh-shushing the girls.* New York, NY: Palgrave MacMillan.

Kachel, S., Steffens, M.C., & Niedlich, C. (2016). Traditional masculinity and femininity:

validation of a new scale for assessing gender roles. *Frontiers in Psychology, 7* (956), 1 – 19.

Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23* (3), 187 – 200.

Kantor, J. (2013, September 7). Harvard business school case study: Gender equity. *The New York Times.* Retrieved from http://www.nytimes.com/2013/09/08/education/harvard-case-study-gender-equity.html?pagewanted=all

Kaplan, D.W. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage Publications.

Karakowsky, L., & Elangovan, A.R. (2001). Risky decision making in mixed-gender teams: Whose risk tolerance matters? *Small Group Research, 32*(1), 94-111.

Karakowsky, L., McBey, K., Miller, D.L. (2004). Gender, perceived competence, and power displays: Examining verbal interruptions in a group context. *Small Group Research, 35*(4), 407-439.

Kark, R., Waismel-Manor, R., & Shamir, B. (2012). Does valuing androgyny and femininity lead to a female advantage? The relationship between gender-role, transformational leadership and identification. *The Leadership Quarterly, 23(3),* 620-640.

Kelloway, K. (2015). *Using Mplus for structural equation modeling: a researchers guide.* Thousand Oaks, CA: Sage Publications.

Kenny, D.A. (1979). *Correlation and causality.* New York, NY: Wiley.

Kenny, D.A., & McCoach, D.B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling, 10*(3), 333 –

351.

Kent, R., & Moss, S.E. (1994). Effects of sex and gender role on leader emergence. *Academy of Management Journal, 37*(3), 1335-1346. Doi: 10.2307/256675.

Kirchmeyer, C. (1996). Gender roles and decision-making in demographically diverse groups: A case for reviving androgyny. *Sex Roles, 34* (9/10), 649-663. Doi: 10.1007/BF01551500.

Kohlberg, L.K. (1981). *The Philosophy of Moral Development.* New York, NY: Harper & Row.

Kolb, J.A. (1997). Are we still stereotyping leadership: A look at gender and other predictors of leader emergence. *Small Group Research, 28*(3), 370-393.

Korabik, K. (1990). Androgyny and leadership style. *Journal of Business Ethics, 9*(4-5), 283-292.

Kraemer, H.C., & Blasey, C.M. (2004). Centering in regression analyses: a strategy to prevent errors in statistical inference. *Internal Journal of Methods in Psychiatry Research, 13* (3), 141 – 151.

Kunce, J.T., Cook, W.D., & Miller, D.E. (1975). Random variables and correlational overkill. *Educational and Psychological Measurement, 35*, 529 – 532.

Leaper, C. (2015). Gender and social-cognitive development. In R. M. Lerner (Series Ed.), L. S. Liben & U. Muller (Vol. Eds.), *Handbook of child psychology and developmental science* (7th ed.), Vol. 2: *Cognitive processes* (pp. 806-853). New York: Wiley.

LeBreton, J.M., & Senter, J.L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11* (4), 815 – 852.

Lee, C., & Fahr, J.L. (2004). Joint effects of group efficacy and gender diversity on group

cohesion and performance. *Applied Psychology: An International Review, 53* (1), 136 – 154.

Leftkowitz, E.S., & Zeldow, P.B. (2006). Masculinity and femininity predict optimal mental health: A belated test of the androgyny hypothesis. *Journal of Personality Assessment, 87*(1), 95-101.

LePine, J.A., Hollenbeck, J.R., Ilgen, D.R., Colquitt, J.A., & Ellis, A. (2002). Gender composition, situational strength, and team decision-making accuracy: A criterion decomposition approach. *Organizational Behavior and Human Decision Processes, 88*(1), 445-475.

Lichenstein, B.B., Uhl-Bien, M., Marion, R., Seers, A., Orton, J.D. (2006). Complexity leadership theory: An interactive perspective on leading in complex adaptive systems. *Emergence: Complexity and Organization, 8, 4*, 2 – 12.

Lim, B. C., & Ployhart, R. E. (2004). Transformational leadership: Relations to the five-factor model and team performance in typical and maximum contexts. *Journal of Applied Psychology, 89*(4), 610–621.

Lubinski, D., Tellegen, A., & Butcher, J.N. (1981). The relationship between androgyny and subjective indicators of emotional well-being. *Journal of Personality and Social Psychology, 40*(4), 722 – 730.

Lund, D.B. (2008). Gender differences in ethics judgment of marking professionals in the United States. *Journal of Business Ethics, 77*(4), 501-515. Doi: 10.1007/s10551-007-9362-z.

Maccoby, E.E. (1990). Gender and relationships: a developmental account. *American Psychologist, 45*(4), 513 – 520.

Mahapatra, L. (2013, September 13). *The bros of Lehman Brothers: Where are they now?*

Retrieved from IBT: http://www.ibtimes.com/bros-lehman-brothers-where-are-they-now-infographic-1405744

Mathieu, J.E., Maynard, M.T., Rapp, T.L. & Gilson, L.L. (2008). Team effectiveness 1997 – 2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34*(3), 410 – 476.

Marsh, H.W., Kit-Tai, H., & Wen, Z. (2009). In search of golden rules: comment on hypothesis testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling: A Multidisciplinary Journal, 1* (3), 320 – 341.

Marsh, H.W., Muthen, B., Asparouhov, T. Ludtke, O., Robitzsch, A., Morin, A.J.S., & Trautwein, U. (2004). Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluation of university teaching. *Structural Equation Modeling, 16,* 439 – 476.

Martin, C.L., Cook, R.E., & Andrews, N.C.Z. (2016). Reviving androgyny: A modern day perspective on flexibility of gender identity and behavior. *Sex Roles, 76*(9-10), 1-12. Doi: 10.1007/s11199-016-0602

McGregor, J., & Tweed, D. (2001). Gender and managerial competence: support for theories of androgyny? *Women in Management Review, 16*(6), 279 – 287.

Meertens, R.M., & Lion, R. (2008). Measuring an individual's tendency to take risks: The risk propensity scale. *Journal of Applied Social Psychology, 38* (6), 1506-1520.

Meier-Pesti, K., & Penz, E. (2008). Sex or gender? Expanding the sex-based view by introducing masculinity and femininity as predictors of financial risk taking. *Journal of Economic Psychology, 29*(2), 180 – 196. Doi: https://doi.org/10.1016/j.joep.2007.05.002.

Meyers, L.S., Gamst, G., & Guarino, A.J. (2013). *Applied multivariate research: Design and interpretation.* Thousand Oaks, CA: Sage Publications.

Miller, L., & Ubeda, P. (2012). Are women more sensitive to the decision-making context? *Journal of Economic Behavior & Organization, 83*(1), 98-104. Doi: 10.1016/j.jebo.2011.06.014.

Morgeson, F.P., & DeRue, D.S., & Karam, E.P. (2010). Leadership in teams: A functional approach to understanding leadership structures and processes. *Journal of Management,* 5-39.

Moscowitz, D., Suh, E.J., & Desaulniers, J. (1994). Situational influences on gender differences in agency and communion. *Journal of Personality and Social Psychology*, *66*(4), 753 – 761.

Moss-Kanter, R. (2010, October 25). What if Lehman brothers had been Lehman sisters? *Harvard Business Review*. Retrieved from https://hbr.org/2010/10/what-if-lehman-brothers-had-be

Neill, J. (2008). Writing up Factor Analysis. *Creative Commons.* Retrieved from http://www.bwgriffin.com/gsu/courses/edur9131/content/Neill2008_WritingUpAFactorAnalysis.pdf

Newton-Small, J. (2016, January 5). How more women on Wall Street could have prevented the financial crises. *Fortune.* Retrieved from http://fortune.com/2016/01/05/wall-street-women-financial-crisis/

Nieboer, J. (2013). Essays on group decision making under risk (Doctoral thesis, University of Nottingham, United Kingdom). Retrieved from

http://eprints.nottingham.ac.uk/13385/1/Jeroen_Nieboer_-_PhD_thesis_%281%29.pdf

Norlander, T., & Erixon, A., & Archer, T. (2000). Psychological androgyny and creativity: Dynamics of gender-role and personality trait. *Social Behavior and Personality: an International Journal, 28*(5), 423-435.

O'Keefe, D.J. (2007). Post hoc power, observed power, a prioi power, retrospective power, prospective power, achieved power: sorting out appropriate uses of statistical power analyses. *Communication Methods and Measures, 1*(4), 291 – 299.

Orchard, C.A., King, G.A., Khalili, H., & Bezzina, M.B. (2012). Assessment of Interprofessional Team Collaboration Scale (AITCS): Development and Testing of the Instrument. *Journal of Continuing Education in the Health Professional, 32* (1), 58 – 67.

Padilla-Walker, L.M., & Jensen, L.A. (2016). Validation of the long and short-form of the ethical values assessment (EVA): A questionnaire measuring the three ethics approach to moral psychology, *International Journal of Behavioral Development, 40*(2), 181-192.

Pelled, L.H., Eisenhardt, K., & Xin, K. (1999). Exploring the black box: An analysis of workgroup diversity, conflict, and performance. *Administrative Science Quarterly, 44*(1), 1 -28.

Perryman, A.A., Fernando, G.D., & Tripathy, A. (2016). Do gender differences persist? An examination of gender diversity on firm performance, risk, and executive compensation? *Journal of Business Research, 69*(2), 579-586.

Peterson, D.K. (2002). Deviant workplace behavior and the organization's ethical

climate. *Journal of Business & Psychology, 17*(1), 47-61.

Plunkett-Tost, L., Gino, F., & Larrick, R.P. (2013). When power makes others speechless: The negative impact of leader power on team performance. *Academy of Management Journal,* 56(5), 1465-1486.

Podsakoff, P.M., MacKenzie, S.B., Lee, J.Y., & Podsakoff, N.P. (2003). Common method bias in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88 (5),* 879 – 903.

Post, C. & Byron, K. (2014). Women on boards and firm financial performance: a meta-analysis. *Academy of Management Journal, 58*(5), 1546-1571.

Powell, G.N. (2012). Six ways of seeing the elephant: the intersection of sex, gender, and leadership. *Gender in Management: An International Journal, 27*(2), 119-141.

Powell, M., & Ansic, D. (1997). Gender differences in risk behavior in financial decision-making: An experimental analysis. *Journal of Economic Psychology, 18*(6), 605-628.

Powell, G.N., & Butterfield, D.A. (2015). The role of androgyny in leader prototypes over four decades. Gender and Management: *An International Journal*, *30*(1), 69-86. doi: 10.1108/GM-07-2013-0082

Radecki, C.M., & Jaccard, J. (1996). Gender-role differences in decision-making Orientations and decision-making skills. *Journal of Applied Social Psychology, 26*(1), *76-94.*

Rahim, M.A. (1983). A measure of styles of handling interpersonal conflict. *The Academy of Management Journal, 26*(2), 368-376.

Rao, K., & Tilt, C. (2016). Board composition and corporate social responsibility: The

role of diversity, gender, strategy and decision making. *Journal of Business Ethics, 138*(2), 327 – 347.

Richardson, J. (2008, September 16). *America's most wanted: The Lehman Brothers Board of Directors.* Retrieved from Global Investment Watch: http://globalinvestmentwatch.com/americas-most-wanted-the-lehman-brothers-board-of-directors-2/

Rodler, C., Kirchler, E., & Holzl, E. (2001). Gender stereotypes of leaders: An analysis of the contents of obituaries from 1974 to 1998. *Sex Roles, 45*(11/12), 827 – 843.

Rogelberg, S.G., Barnes-Farrell, J.L., & Lowe, C.A. (1992). The stepladder technique: An alternative group structure facilitating effective group decision making. *Journal of Applied Psychology, 77* (5), 730 – 737.

Rogelberg, S.G., & Rumery, S.M. (1996). Gender diversity, team decision quality, time of task, and interpersonal cohesion. *Small Group Research, 27*(1), 79 – 90.

Rohner, U., & Dougan, B.W. (2012). *Gender diversity and corporate leadership*. Retrieved from the Credit Suisse Website: https://www.calstrs.com/sites/main/files/fileattachments/csri_gender_diversity_and_corporate_performance.pdf

Rohrman, B. (1997). *Risk orientation questionnaire: Attitudes towards risk decisions.* Retrieved from: http://www.rohrmannresearch.net/pdfs/rohrmann-rac-roq.pdf

Rohrmann, B. (2005). Risk Attitude Scales: Concepts, questionnaires, utilizations. *Project Report,* 1-21.

Rosenthal, C. S. (2000). Gender styles in state legislative committees: Raising their voices in resolving conflict. *Women & Politics, 21*(2), 21–45. Doi:10.1300/J014v21n02_02.

Rothschild, J., & Miethe, T.D. (1999). Whistle-blower disclosures and management retaliation: The battle to control information about organizational corruption. *Work & Occupations, 26*(1), 107 – 128.

Suar, D., & Gochhayat, J. (2014). Influence of biological sex and gender roles on ethicality. *Journal of Business Ethics, 134*(2), 199 – 208.

Schaffer, B.S., & Riordan, C.M. (2003). A review of cross-cultural methodologies for organizational research: a best-practices approach. *Organizational Research Methods, 6* (2), 169 – 215.

Schein, V.E. (2001). A global look at psychological barriers to women's progress in management. *Journal of Social Issues, 57*(4), 675 – 688.

Schein, V.E., & Davidson, M.J. (1993). Think manager, think male. *Management Development Review, 6*(3), 24-28.

Schein, V.E., & Mueller, R. (1992). Sex role stereotyping and requisite management characteristics: A cross cultural look. *Journal of Organizational Behavior, 13*(1), 439-447.

Sealy, R., & Vinnicombe, S. (2010). Boardroom balance. *International Centre for Women Leaders.* 24.

Sedney, M.A. (1981). Comments on median split procedures for scoring androgyny measures. *Sex Roles, 7* (2), 217 – 222.

Singh, V., Terjesen, S., & Vinnicombe, S. (2008). Newly appointed directors in the boardroom: How do women and men differ? *European Management Journal, 26*, 48-58. doi:10.1016/j.emj.2007.10.002

Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior, 68*, 446-460.

Spence, J.T. (1985). Gender identity and its implications for the concepts of masculinity and femininity. In T.B. Sonderegger (Ed.), *Nebraska Symposium on Motivation: Psychology of gender* (pp. 59-96). Lincoln, NE: University of Nebraska Press.

Spence. J.T., Helmreich, R.L. & Stapp, J. (1974). The personal attributes questionnaire: A measure of sex role stereotypes and masculinity-femininity. *JSAS Catalog of Selected Documents in Psychology, 4*, 43.

Spence, J.T., Helmreich, R.L., & Stapp, J. (1975). Ratings of self and peers on sex attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology, 32*(1), 29-39.

Spector, P.E. (1992). *Summated rating scale construction: an introduction.* California: Sage University.

Status of Women Canada. (2014). *Good for business: A plan to promote the participation of more women on Canadian boards.* Retrieved from http://www.swccfc.gc.ca/initiatives/wldp/wb-ca/wob-fca-eng.html

Streiner, D.L. (2002). Breaking up is hard to do: the heartbreak of dichotomizing continuous data. *Canadian Journal of Psychiatry, 47* (3), 262 – 266.

Stephens, H.E., (2004). *Gender-role orientation and leadership effectiveness* (Masters thesis). Retrieved from Library and Archives Canada. (0-494-01864-X)

Sutter, M. (2004). Are four heads better than two? An experimental beauty-contest game with teams of different size. *Max Planck Institute for Research into Economic Systems, 88*(1), 41 – 46.

Symons, L., & Ibarra, H. (2014, April 28). What the scarcity of women in business case studies really looks like. *Harvard Business Review.* Retrieved from http://web.b.ebscohost.com.library.smu.ca:2048/ehost/pdfviewer/pdfviewer?vid=3&sid=66544773-1db6-42d1-ab92-032ff8795c23%40sessionmgr104

Tabachnick, B.G., & Fidell, L.S. (1996). *Using Multivariate Statistics,* 3rd ed. New York, NY: HarperCollins.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using Multivariate Statistics*, 5th ed. Boston, MA: Allyn & Bacon.

Trapnell, P.D., & Paulhus, D.L. (2011). Agentic and communal values: Their scope and measurement. *Journal of Personality Assessment, 94*(1), 39-52.

Turner, C. (2012, May 7). A balance of both masculine and feminine strengths: The bottom-line benefit. *Forbes.* Retrieved from https://www.forbes.com/sites/womensmedia/2012/05/07/a-balance-of-both-masculine-and-feminine-strengths-the-bottom-line-benefit/

Twenge, J.M. (1997). Changes in masculine and feminine traits over time: a meta-analysis. *Sex Roles, 36* (5/6). 305 – 324.

Vainapel, S., Shamir, O., Tenenbaum, Y., & Gilam, G. (2015). The dark side of gendered language: The masculine-generic form as a cause for self-report bias. *Psychological Assessment, 27*(4). 1513 – 1519.

Valentine, S.R., Godkin, L., Page, K., & Rittenburg, T. (2008). Gender and ethics: ethical judgments, ethical intentions, and altruism among healthcare professionals. *Women in Management Review, 24*(2), 112 – 130.

Valentine, S.R., & Rittenburg, T.L. (2007). The ethical making of men and women

executives in international business situations. *Journal of Business Ethics, 71* (2), 125- 134.

Van Engen, G.M. (2012). Are (male) leaders "feminine" enough? Gendered traits of identity as mediators of sex differences in leadership styles. *Gender im Management: An International Journal, 27*(5), 296 – 314.

Van Mierlo, H., Vermunt, J.K., & Rutte, C.G. (2008). Composing group-level constructs from individual-level survey data. *Organizational Research Methods, 12*(2), 368 – 392.

Velicer, W.F., & Fava, J.L. (1998). Effects of variable and subject sampling on factor pattern recovery. *Psychological Methods, 3* (2), 231 – 251.

Victor, B. & Cullen, J.B. (1987). A theory and measure of ethical climates in organizations. *Research in Corporate Social Performance and Policy, 9,* 51 – 71.

Weber, E.U., Blais, A.R., & Betz, N.E. (2002). A domain-specific risk attitude scale: Measuring risk perceptions and risk behaviors. *Journal of Behavioral Decision Making, 15,* 263-290. Doi: 10.1002/bdm.414.

Webber, S.S., & Donahue, L.M. (2001). Examining the "double-edged sword" of diversity in work groups: A meta-analysis. *Journal of Management, 27*(2), 141-162.

Wei, X., Liu, Y, & Allen, N.J.. (2016). Measuring team emotional intelligence: A multidimensional comparison. *Group Dynamics: Theory, research, and practice, 20*(1), 34-50.

Weiss, M., Kolbe, M., Grote, G., Dambach, M., Marty, A., Spahn, D.R., & Bastian, G. (2014). Agency and communion predict speaking up in acute care teams. *Small Group Research, 45*(3), 290-313.

Whetten, D.A. (1989). What constitutes a theoretical contribution? *Academy of Management Review, 14*, 490-495.

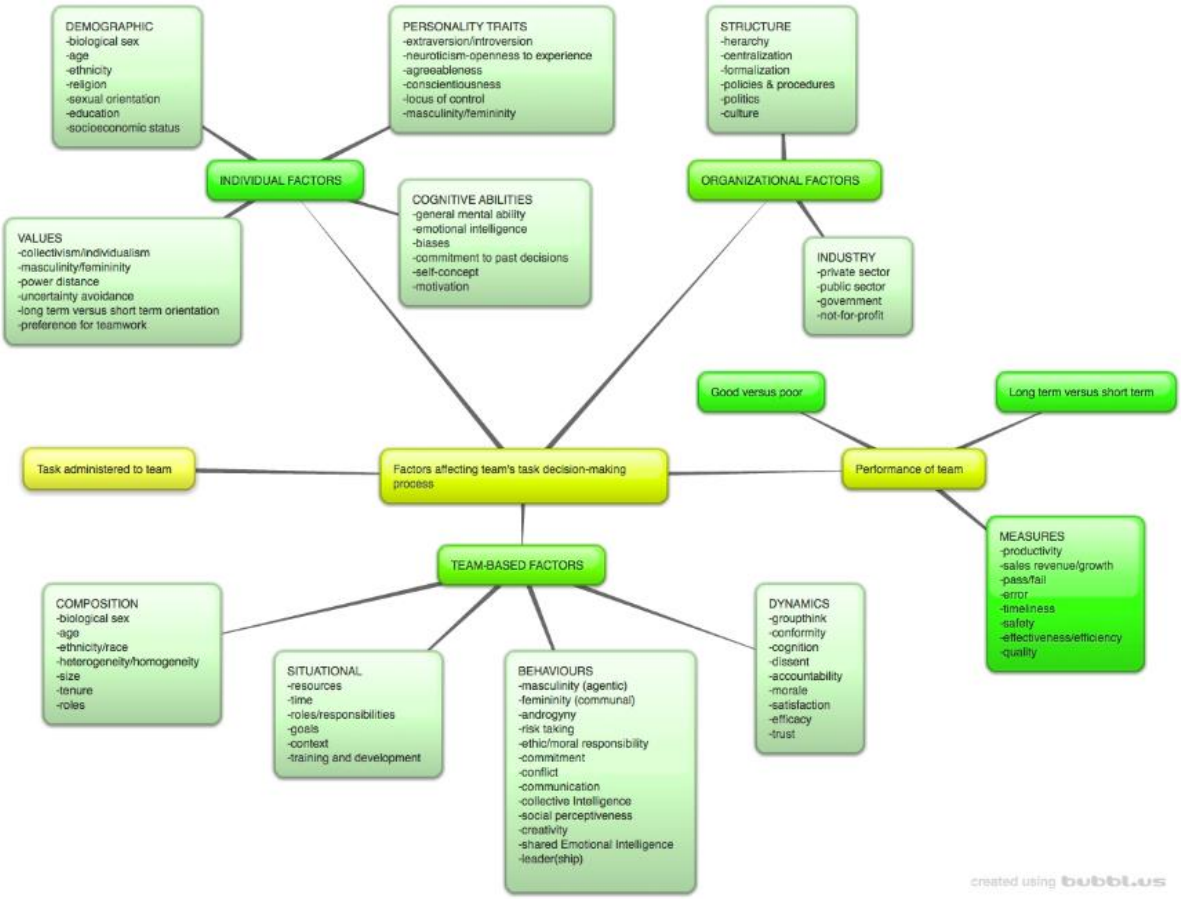
Wood, W. (1987). Meta-analytic review of sex differences in group performance. *Psychological Bulletin, 102* (1), 53 – 71.

Woodhill, B.M., & Samuels, C.A. (2003). Positive and negative androgyny and their relationship with psychological health and well-being. *Sex Roles, 48*(11-12), 555-565.

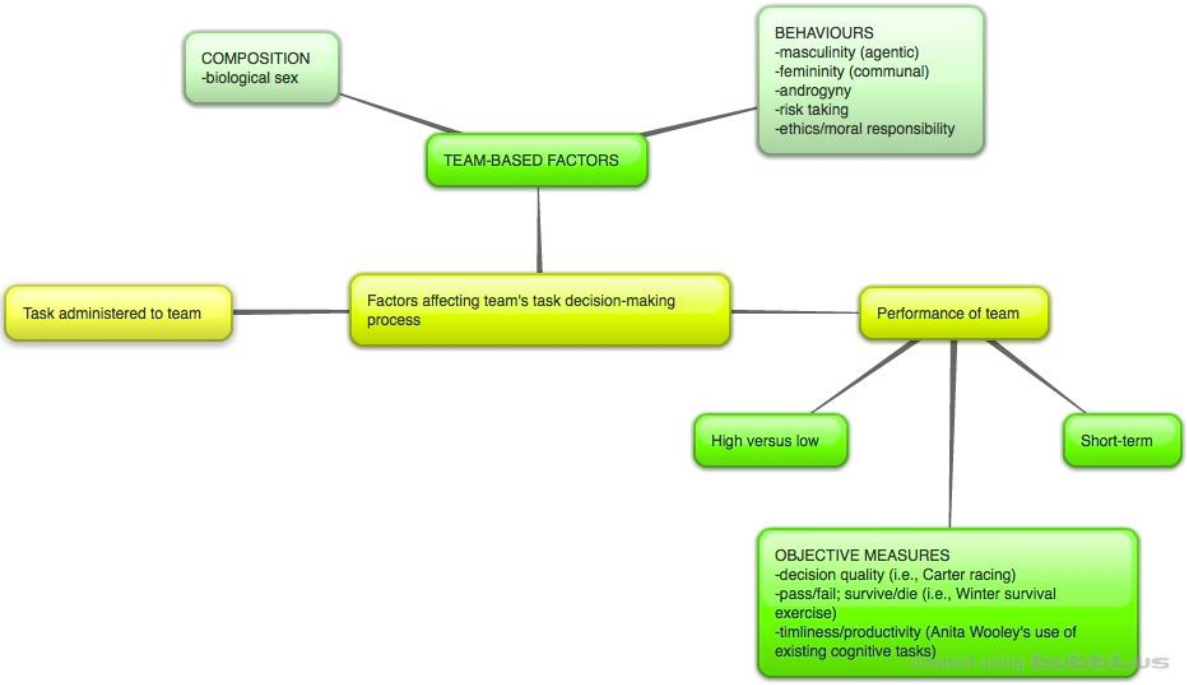
Worthington, R.L., & Whittaker, T.A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34 (6),* 806 – 838.

Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., & Malone, T.W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science, 330,* 686- 688.

Appendix A
Overall Conceptual Team Decision-Making Process Model



Appendix B
Conceptual Model of this Dissertation



Appendix C
List of Items used for Item Confirmation

Below is a list of statements that describe organizational team behaviors. Thinking about each statement, give your <u>first impression</u> of whether the statement would be apparent in a masculine team or a feminine team. <u>Don't spend too long on deciding what your answer should be</u>.

Statement	**Masculine or feminine? (Please fill in)**
1. Our team is cautious when making a decision	
2. When making a decision, our team follows the motto 'nothing ventured, nothing gained'	
3. Our team enjoys risky decisions	
4. When making a decision, our team prefers to stay on the safe side	
5. Success makes our team take more risks	
6. Our team makes decisions carefully	
7. Even when our team is unsure about how to complete a task, we try our luck	
8. Before making a decision, our team considers all possible alternatives	
9. Our team can make decisions easily	
10. Our team stands up well under pressure	
11.We ensure that everyone on our team participates in decision making	
12. Our team is confident when making a decision	
13. Our team is willing to take risks	
14. Our team is willing to take a stand	
15. As a team, we express trust in each other	
16. As a team, we collectively work together to reach a decision	
17. Our team openly listens to each other's issues	
18. As a team, we care about each other's well-being	

19. Our team attentively listens to each other's suggestions	
20. As a team, we assert our opinions when making a decision	
21. Our team tends to use power in competitive situations	
22. The glue that holds our team together is our task accomplishments	
23. The glue that holds our team together is our loyalty	
24. Our team encourages open communication	
25. Our team has a lot of team spirit	
26. Our team is always focused on the task at hand	
27. Our team uses logical thinking when making a decision	
28. We are perceptive to each other's points of view	
29. As a team, we use cooperation to gain agreement on a decision	
30. If there is an issue, our team takes the time to resolve it	
31. Our team is task oriented	
32. Our team is focused on building positive relationships	
33. We believe that coming out on top is the key to success	
34. Our team gathers multiple inputs before reaching a decision	
35. Our team takes turns providing input when completing a task	
36. As a team, we share the power	
37. Our team sees any task as a game to be won	
38. A spirit of teamwork exists	
39. Our team will do anything to get ahead	
40. The key to a successful team is building relationships	
41. We tend to avoid conflict	
42. We use concrete facts to resolve a conflict	
43. We resolve conflict indirectly	
44. We resolve conflict directly	
45. We focus on the overall process to	

reach consensus	
46. Our team uses competition when completing a task	
47. Kindness is important to the success of our team	
48. To reach a decision, we focus on the goal at hand	
49. Our team will do whatever it takes to perform well	
50. We avoid using our emotions when finalizing a decision	
51. We spend a lot of time finding an ideal solution to a problem	
52. Once we connect to a solution, we stop discussing other possibilities	
53. Using collaboration helps us reach sound decisions	
54. We consider the rights of others to arrive at a fair/moral decision	
55. Our team tends to generate a solution as quickly as possible	
56. Our team spends a lot of time exploring possible solutions	
57. We make decisions based on tradition	
58. We make decisions based on the interests of all parties involved	
59. To our team, success means winning	
60. We rate success as being valued	
61. Our team uses intuition when making a decision	
62. We don't let distracting conversations get in the way of completing a task	
63. We use our assertiveness to our advantage	
64. We follow the motto, 'If it ain't broke, don't fix it'	

Appendix D
Pilot Test Study Materials

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Letter of Information about this Research Study

Have you ever-wondered how high performing teams behave when making decisions? As organizations become more complex and decentralized, researchers and practitioners alike are intrigued in determining the factors that lead to effective team-based decision-making. This study does just that. You are invited to complete a survey based on the gendered behaviors of organizational teams.

If you decide to participate, it will take approximately 10-15 minutes of your time. Your responses are completely anonymous and only the researchers will ever see your evaluation. Your decision to participate or not participate in this study will in no way affect your enrollment as a student at Saint Mary's University.

Participation in this survey is voluntary and participants are free to withdraw at any time without penalty. You are not obliged to answer questions that you feel are objectionable or which make you feel uncomfortable. There are no known physical, economic, or social risks to your participating in this study. Your decision to take part will be indicated by submitting the survey to the researcher.

We plan to publish papers reporting the aggregate findings of this study in academic journals. Only group level data will be reported in any publications. There will be no way that the researchers can tell who has completed which survey. The data will be saved in a password-encrypted file. The data will be destroyed after a period of five years.

Should you wish to receive a report with the details of this study or should you have any questions about this research please contact one of the researchers either by email or by telephone: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293; or Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328. We anticipate that a report will be available in the fall of 2017.

The proposal for this research (File number -16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board at ethics@smu.ca or by telephone at (902) 420-5728.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

Using the following rating scale, please rate the likelihood of each statement occurring in a *FEMININE* organizational decision-making team. The rating scale goes from 1 – 'strongly disagree' to 5 – 'strongly agree'.

Here is an example: 'Our team does a thorough job together'. Decide if 'our team does a thorough job together' is characteristic of a feminine team. If you decide 'our team does a thorough job together' is unlikely in a feminine team, than you should choose number 1 which corresponds to 'strongly disagree'.

Please reply to all statements. Give your first impression of whether or not the statement would be apparent in a feminine team. Don't spend too long on deciding what your answer should be. **Mark an X in the appropriate box.**

Remember: Not all statements are apparent in a feminine team so please spread your ratings out.

Statement	1 – Strongly disagree	2 – Disagree	3 – Neutral	4 – Agree	5 – Strongly Agree
1. Our team is cautious when making a decision					
2. We use our assertiveness to our advantage					
3. Our team follows the motto 'nothing ventured, nothing gained' when making a decision					
4. We don't let distracting conversations get in the way of completing a task					
5. Our team encourages open communication					
6. We resolve conflict directly					
7. Our team uses intuition when making a decision					
8. Our team enjoys risky decisions					
9. We rate success as being valued					
10. Our team prefers to say on the safe side when making a decision					
11.We make decisions based on the interests of all parties involved					
12. The glue that holds our team together is our task accomplishments					
13. Our team stands up well under					

pressure					
14. Our team makes decision carefully					
15. To our team, success means winning					
16. We ensure that everyone on our team participates in decision making					
17. We follow the motto, 'If it ain't broke, don't fix it'					
18. We spend a lot of time finding an ideal solution to a problem					
19. Even when our team is unsure about a decision, we try our luck					
20. Our team attentively listens to each other's suggestions					
21. We make decisions based on tradition					
22. Success makes our team take more risks					
23. As a team, we express trust in each other					
24. Our team is task oriented					
25. The key to a successful team is building relationships					
26. Our team tends to generate a solution as quickly as possible					
27. Before making a decision, our team considers all possible alternatives					
28. Our team openly listens to each other's issues					
29. Our team is always focuses on the task at hand					
30. Kindness is important to the success of our team					
31. Our team will do whatever it takes to perform well					
32. We avoid using our emotions when finalizing a decision					
33. Our team spends a lot of time exploring possible solutions					
34. Our team uses logical thinking when making a decision					
35. As a team, we use cooperation to gain agreement on a decision					

36. Our team is willing to take risks					
37. Our team cane make decisions easily					
38. As a team, we share the power					
39. We tend to avoid conflict					
40. We consider the rights of others to arrive at a fair/moral decision					
41. As a team, we assert our opinions when making a decision					
42. The glue that holds our team together is our loyalty					
43. Our team gathers multiple inputs before reaching a decision					
44. Our team uses competition when completing a task					
45. Using collaboration helps us reach sound decisions					
46. Our team is confident when making a decision					
47. As a team, we collectively work together to reach a decision					
48. We use facts to resolve a conflict					
49. Our team sees any task as a game to be won					
50. As a team, we care about each other's well being					
51. Our team will do anything to get ahead					
52. Once we connect to a solution, we stop discussing other possibilities					
53. We resolve conflicts indirectly					
54. If there is an issue, our team takes the time to resolve it					
55. Our team is willing to take a stand					
56. Our team tends to use power in competitive situations					
57. We are perceptive to each other's point of view					
58. To reach a decision, we only focus at the goal at hand					
59. A spirit of teamwork exists					
60. We focus on the overall process to reach consensus					

61. Our team is focused on building positive relationships					
62. Our team takes turn providing input when completing a task					
63. We believe that coming our on top is the key to success					
64. Our team has a lot of team spirit					

Using the following rating scale, please rate the likelihood of each statement occurring in a *MASCULINE* organizational decision-making team. The rating scale goes from 1 – 'strongly disagree' to 5 – 'strongly agree'.

Here is an example: 'Our team does a thorough job together'. Decide if 'our team does a thorough job together' is characteristic of a masculine team. If you decide 'our team does a thorough job together' is unlikely in a masculine team, than you should choose number 1 which corresponds to 'strongly disagree'.

Please reply to all statements. Give your first impression of whether or not the statement would be apparent in a masculine team. Don't spend too long on deciding what your answer should be. **Mark an X in the appropriate box.**

Remember: Not all statements are apparent in a masculine team so please spread your ratings out.

Statement	1 – Strongly disagree	2 – Disagree	3 – Neutral	4 – Agree	5 – Strongly Agree
1. Our team is cautious when making a decision					
2. We use our assertiveness to our advantage					
3. Our team follows the motto 'nothing ventured, nothing gained' when making a decision					
4. We don't let distracting conversations get in the way of completing a task					
5. Our team encourages open communication					
6. We resolve conflict directly					
7. Our team uses intuition when making a decision					
8. Our team enjoys risky decisions					
9. We rate success as being valued					

10. Our team prefers to say on the safe side when making a decision					
11.We make decisions based on the interests of all parties involved					
12. The glue that holds our team together is our task accomplishments					
13. Our team stands up well under pressure					
14. Our team makes decision carefully					
15. To our team, success means winning					
16. We ensure that everyone on our team participates in decision making					
17. We follow the motto, 'If it ain't broke, don't fix it'					
18. We spend a lot of time finding an ideal solution to a problem					
19. Even when our team is unsure about a decision, we try our luck					
20. Our team attentively listens to each other's suggestions					
21. We make decisions based on tradition					
22. Success makes our team take more risks					
23. As a team, we express trust in each other					
24. Our team is task oriented					
25. The key to a successful team is building relationships					
26. Our team tends to generate a solution as quickly as possible					
27. Before making a decision, our team considers all possible alternatives					
28. Our team openly listens to each other's issues					
29. Our team is always focuses on the task at hand					
30. Kindness is important to the success of our team					
31. Our team will do whatever it takes to perform well					

32. We avoid using our emotions when finalizing a decision					
33. Our team spends a lot of time exploring possible solutions					
34. Our team uses logical thinking when making a decision					
35. As a team, we use cooperation to gain agreement on a decision					
36. Our team is willing to take risks					
37. Our team cane make decisions easily					
38. As a team, we share the power					
39. We tend to avoid conflict					
40. We consider the rights of others to arrive at a fair/moral decision					
41. As a team, we assert our opinions when making a decision					
42. The glue that holds our team together is our loyalty					
43. Our team gathers multiple inputs before reaching a decision					
44. Our team uses competition when completing a task					
45. Using collaboration helps us reach sound decisions					
46. Our team is confident when making a decision					
47. As a team, we collectively work together to reach a decision					
48. We use facts to resolve a conflict					
49. Our team sees any task as a game to be won					
50. As a team, we care about each other's well being					
51. Our team will do anything to get ahead					
52. Once we connect to a solution, we stop discussing other possibilities					
53. We resolve conflicts indirectly					
54. If there is an issue, our team takes the time to resolve it					
55. Our team is willing to take a stand					
56. Our team tends to use power in					

competitive situations					
57. We are perceptive to each other's point of view					
58. To reach a decision, we only focus at the goal at hand					
59. We focus on the overall process to reach consensus					
60. Our team is focused on building positive relationships					
61. Our team takes turn providing input when completing a task					
62. We believe that coming our on top is the key to success					
63. Our team has a lot of team spirit					

DEMOGRAPHIC QUESTIONS

1. Age in years: _____

2. What is your gender?

___Male
___Female
___Other

3. What is your ethnicity (Please circle)?

___Caucasian
___Black
___Hispanic
___Latino
___Asian
___Aboriginal
___Arab
___Multiracial
___Would rather not say
___Other

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Thank you for taking the time to participate in our study about the decision-making styles and behaviors of organizational teams. The data collected from these surveys will help to develop an instrument based on gendered decision-making behaviours of teams (i.e., teams being masculine, feminine, or androgynous). The findings of this study will allow us to conduct future research testing the gendered behaviors of teams relative to performance in team-based tasks.

Please remember that any data pertaining to you, as an individual participant will be anonymous. Once all the data is collected, it will be saved in a password-encrypted file. The data will be destroyed after a period of five years. Any papers that we plan to publish will only report the aggregate findings and group level data – no identifiable information will be used.

If you would like to follow up with the researchers to express any concerns, receive more information on the study, or have the final research report shared with you directly, please contact one of the researchers by e-mail or telephone: Danielle Mercer, danielle.mercer@smu.ca or (902) 210-4293; or Dr. Catherine Loughlin, Catherine.loughlin@smu.ca or (902) 491-6328. Once the study is complete, anticipated by fall 2017, we will happily e-mail a summary of the results to you.

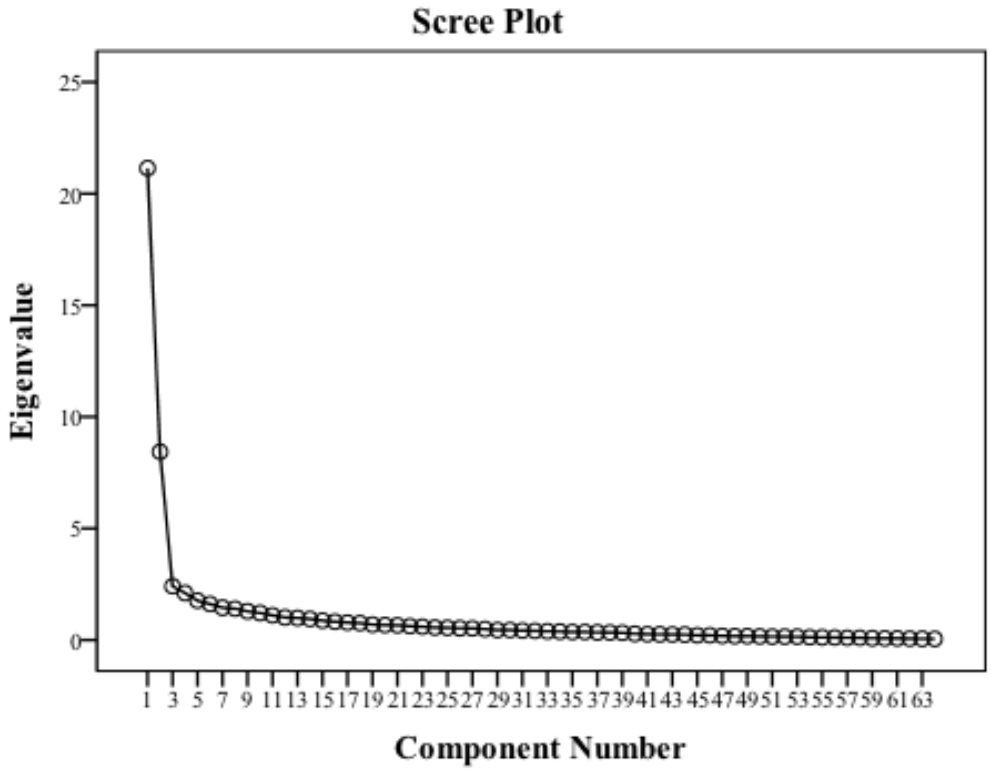
Again, this research project (File number – 16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board by e-mail ethics@smu.ca or by telephone (902) 420-5728.

Thank you for your participation.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

Appendix E
Scree Plot Results from the Pilot Study Principal Components Analysis



Appendix F
Study 1b Study Materials

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Letter of Information about this Research Study

Have you ever-wondered how high performing teams behave when making decisions? As organizations become more complex and decentralized, researchers and practitioners alike are intrigued in determining the factors that lead to effective team-based decision-making. This study does just that. You are invited to complete a survey based on the gendered behaviors of organizational teams.

If you decide to participate, it will take approximately 10-15 minutes of your time. Your responses are completely anonymous and only the researchers will ever see your evaluation. Your decision to participate or not participate in this study will in no way affect your enrollment as a student at Saint Mary's University.

Participation in this survey is voluntary and participants are free to withdraw at any time without penalty. You are not obliged to answer questions that you feel are objectionable or which make you feel uncomfortable. There are no known physical, economic, or social risks to your participating in this study. Your decision to take part will be indicated by submitting the survey to the researcher.

We plan to publish papers reporting the aggregate findings of this study in academic journals. Only group level data will be reported in any publications. There will be no way that the researchers can tell who has completed which survey. The data will be saved in a password-encrypted file. The data will be destroyed after a period of five years.

Should you wish to receive a report with the details of this study or should you have any questions about this research please contact one of the researchers either by email or by telephone: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293; or Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328. We anticipate that a report will be available in the fall of 2017.

The proposal for this research (File number -16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board at ethics@smu.ca or by telephone at (902) 420-5728.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

PART A: TEAM MASCULINITY QUESTIONAIRE[10]

Directions: Using the following scale below, please rate the likelihood of each statement occurring in a **masculine** organizational team. The rating scale goes from 1 – strongly disagree to 5 – strongly agree.

Please reply to all statements. Give your **first impression** of whether or not the statement would be likely to occur in a **masculine** team. Don't spend too long on deciding what your answer should be.

Remember: Not all statements occur in a **masculine** team so please spread your ratings out.

1 **Strongly Disagree**	**2** **Disagree**	**3** **Neutral**	**4** **Agree**	**5** **Strongly Agree**

STATEMENT	**1**	**2**	**3**	**4**	**5**
1. Our team openly listens to each others issues					
2. To our team, success means winning					
3. Our team tends to use power to reach a solution					
4. We are perceptive to each others suggestions					
5. As a team, we care about each other's well-being					
6. Our team will do whatever is takes to perform well					
7. Our team encourages open communication					
8. We consider the rights of others to arrive at a fair decision					
9. Our team is willing to take risks when making decisions					

[10] Order of team masculinity scale and team femininity scale was randomized

	1 Strongly Disagree	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
10. Our team prefers to stay on the safe side when completing a task					
11. We use our competitiveness to successfully complete a task					
12. Our team follows the motto 'nothing ventured, nothing gained' when making a decision					
13. We ensure that everyone on our team participates in decision-making					
14. Our team sees any task as a game to be won					
15. We make decisions based on the interests of all parties involved					
16. Our team tends to generate a solution as quickly as possible					
17. Before reaching a solution, our team considers all possible alternatives					
18. We avoid using emotions when finalizing our decision					
19. We use assertiveness to reach a solution					
20. Kindness is important to the success of the team					

PART B: TEAM FEMININITY QUESTIONAIRE

Directions: Using the following scale below, please rate the likelihood of each statement occurring in a **feminine** organizational team. The rating scale goes from 1 – strongly disagree to 5 – strongly agree.

Please reply to all statements. Give your **first impression** of whether or not the statement would be likely to occur in a **feminine** team. Don't spend too long on deciding what your answer should be.

Remember: Not all statements occur in a **feminine** team so please spread your ratings out.

1 Strongly Disagree	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**

STATEMENT	**1**	**2**	**3**	**4**	**5**
1. Our team openly listens to each others issues					
2. To our team, success means winning					
3. Our team tends to use power to reach a solution					
4. We are perceptive to each others suggestions					
5. As a team, we care about each other's well-being					
6. Our team will do whatever is takes to perform well					
7. Our team encourages open communication					
8. We consider the rights of others to arrive at a fair decision					
9. Our team is willing to take risks when making decisions					

	1 Strongly Disagree	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
10. Our team prefers to stay on the safe side when completing a task					
11. We use our competitiveness to successfully complete a task					
12. Our team follows the motto 'nothing ventured, nothing gained' when making a decision					
13. We ensure that everyone on our team participates in decision-making					
14. Our team sees any task as a game to be won					
15. We make decisions based on the interests of all parties involved					
16. Our team tends to generate a solution as quickly as possible					
17. Before reaching a solution, our team considers all possible alternatives					
18. We avoid using emotions when finalizing our decision					
19. We use assertiveness to reach a solution					
20. Kindness is important to the success of the team					

PART C: DEMOGRAPHIC QUESTIONS

1. Age in years: _____

2. What is your gender?

___Male
___Female
___Other

3. What is your ethnicity (Please circle)?

___Caucasian
___Black
___Hispanic
___Latino
___Asian
___Aboriginal
___Arab
___Multiracial
___Would rather not say
___Other

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Thank you for taking the time to participate in our study about the decision-making styles and behaviors of organizational teams. The data collected from these surveys will help to develop an instrument based on gendered decision-making behaviours of teams (i.e., teams being masculine, feminine, or androgynous). The findings of this study will allow us to conduct future research testing the gendered behaviors of teams relative to performance in team-based tasks.

Please remember that any data pertaining to you, as an individual participant will be anonymous. Once all the data is collected, it will be saved in a password-encrypted file. The data will be destroyed after a period of five years. Any papers that we plan to publish will only report the aggregate findings and group level data – no identifiable information will be used.

If you would like to follow up with the researchers to express any concerns, receive more information on the study, or have the final research report shared with you directly, please contact one of the researchers by e-mail or telephone: Danielle Mercer, danielle.mercer@smu.ca or (902) 210-4293; or Dr. Catherine Loughlin, Catherine.loughlin@smu.ca or (902) 491-6328. Once the study is complete, anticipated by fall 2017, we will happily e-mail a summary of the results to you.

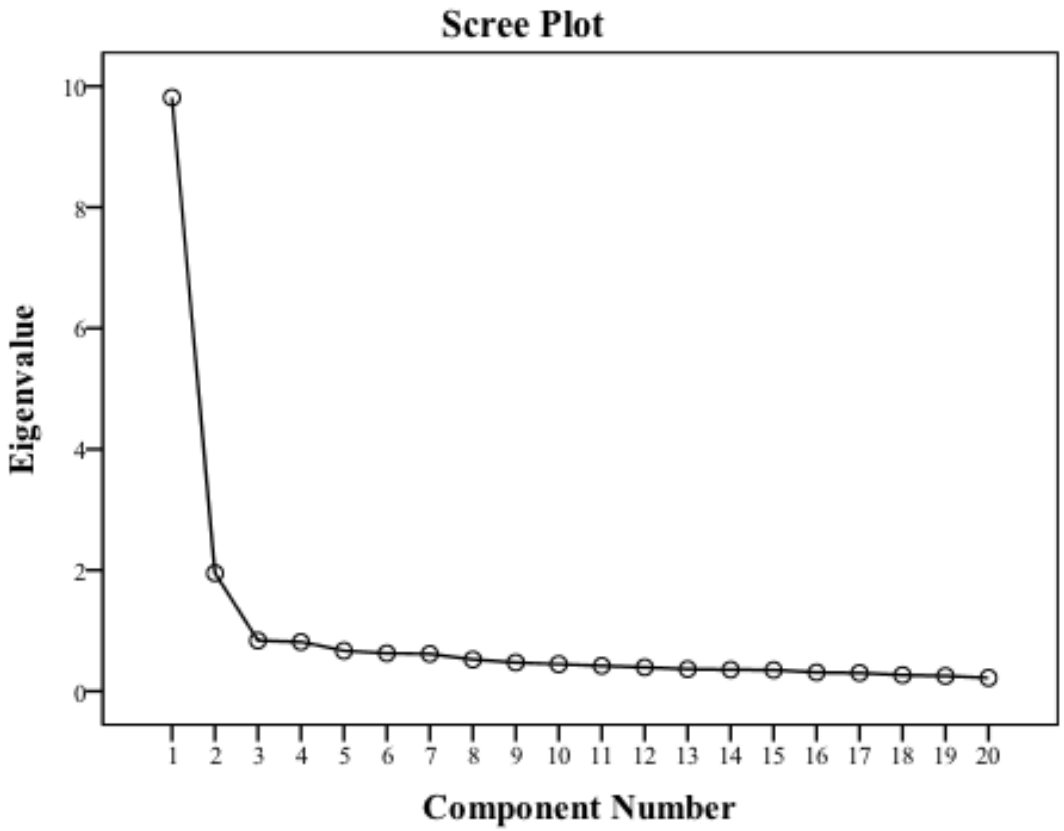
Again, this research project (File number – 16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board by e-mail ethics@smu.ca or by telephone (902) 420-5728.

Thank you for your participation.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

Appendix G
Scree Plot Results from Study 1b Principal Components Analysis



Appendix H
Study 2 Survey Materials

INFORMED CONSENT FORM
Women in leadership: Could more androgynous teams better navigate the paradoxes in complex decision-making? (REB File #06-082)

Researchers: Danielle Mercer, PhD Candidate, & Dr. Catherine Loughlin
Department of Management, Saint Mary's University, Halifax, NS

Contact Information: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293
Dr. Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328

PURPOSE OF THIS RESEARCH

Have you ever-wondered how high performing teams behave when making decisions? As organizations become more complex and decentralized, researchers and practitioners alike are intrigued in determining the factors that lead to effective team-based decision-making. This study does just that. **You are invited to complete a survey based on the gendered behaviors of business teams, such as the team you are participating in for your university class.**

IF YOU CHOOSE TO PARTICIPATE

As a participant, you will be asked to complete the pencil and paper survey by answering various statements using a rating scale. The survey will take approximately 10-15 minutes of your time. We will also request access to your *team's* final grade through your professor. Please be ensured that **we will not** receive this grade until all grades have been submitted to the Registrar's Office at the end of the semester. As well, each team will be given a unique code when completing the in class survey so we, as the researchers, will have no identifiable information (i.e., we will not have access to your name or student number). Therefore, your responses to the survey and your team grade will be anonymous and only the researchers will ever see the anonymous responses/grades.

Remember you are under no obligation to participate. Participation in this survey is voluntary and you a free to withdraw without prejudice or penalty. Your decision to participate or not participate will in no way influence your relationship with the researchers of Saint Mary's University. However, please note that once the surveys have been submitted, you will no longer be able to withdraw since there is no identifiable information connected to your evaluation. You are not obligated to answer questions that you feel are objectionable or which make you feel uncomfortable. There are no known physical, economic, or social risks to your participating in this study. Your decision to participate will be indicated by submitting the completed survey to the researcher.

HOW THE RESEARCH WILL BE USED

The research will initially be used for Danielle Mercer's PhD dissertation. Please note that we will only write and publish papers reporting the aggregate findings of this study for academic conferences and journals. Only group level data will ever be reported. Again, there will be no way that the researchers can tell who has completed which survey. All paper responses will be locked in the researcher's office in accordance with the American Psychological Association guidelines for data storage. For data analysis, we will record the survey responses and grades on a password-protected computer in a password-protected file. The computer-transferred data will be retained for five years after publication, after which it will be securely destroyed.

POTENTIAL BENEFITS OF THE RESEARCH

There are many potential benefits from the results of this study. Our research will not only produce publications in academic and practitioner journals, but we believe that our results may assist in future policy change in both organizations and the government. As a participant, you are helping us in make a positive change in the role of gender, decision-making styles, and teams such as boards of directors.

QUESTIONS ABOUT THE RESEARCH

Should you wish to receive a report with the details of this study or should you have any questions about this research please contact one of the researchers either by email or by telephone: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293; or Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328. We anticipate that a report will be available in the fall of 2017.

Certification:

The proposal for this research (File number -16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board at ethics@smu.ca or by telephone at (902) 420-5728.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

PART A: Team Questionnaire[11]

Directions: Using the following scale below, please **check** whether you agree or disagree with each statement **occurring in your team as your worked on your team case presentation**. The rating scale goes from 1 – **strongly disagree** to 5 – **strongly agree**.

Please reply to all statements. Give your **first impression** of whether or not the statement would be likely to occur in **your team.** Don't spend too long on deciding what your answer should be.

STATEMENT	**1 Strongly Disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
1. Our team openly listened to each others issues					
2. To our team, success meant winning					
3. Our team used power to reach a solution					
4. We were perceptive to each others suggestions					
5. As a team, we cared about each other's well-being					
6. Our team was willing do whatever it took to perform well					
7. Our team encouraged open communication					
8. We considered the rights of others to arrive at a fair decision					
9. Our team was willing to take risks when completing our project					
10. Our team preferred to stay on the safe side when completing any team tasks					
11. We used our competitiveness to successfully complete any team tasks					
12. Our team followed the motto 'nothing ventured, nothing					

[11] The wording of the instructions/items changed slightly depending on if the team was required to complete a project and/or presentation

gained' when making a decision					
13. We ensured that everyone on our team participated in decision-making					
14. Our team saw our course project as a game to be won					
15. We made decisions based on the interests of all parties involved					
16. Our team tended to generate a solution as quickly as possible					
17. Before reaching a solution, our team considered all possible alternatives					
18. We avoided using emotions when finalizing our decision					
19. We used assertiveness to reach a solution					
20. Kindness was important to the success of the team					

2. Directions: Please indicate whether you agree or disagree with each of the following statements **about your team.** The rating scale goes from 1 – **mostly false** to 5 – **completely true.**

STATEMENT	**1 Mostly False**	**2 Somewhat False**	**3 Somewhat True**	**4 Mostly True**	**5 Completely True**
1. When completing our case presentation, certain team members were only concerned for themselves					
2. When working together on our case presentation, there was no room for personal opinions					
3. When completing our case presentation, certain team members cared about their own interests instead of the team as a whole					
4. During our time as team, we looked out for each other					

5. Throughout the semester, our major concern was what was fair for all team members					
6. During our time as a team, we cared for all members					

3. Directions: Please indicate whether you agree or disagree with each of the following statements **about your team.** The rating scale goes from 1 – **not at all important** to 5 – **completely important.**

STATEMENT	**1 Not at all Important**	**2 Of Little Importance**	**3 Of Average Importance**	**4 Very Important**	**5 Completely Important**
1.Our team took responsibility for each other					
2. We tried to achieve our overall goal					
3. Our team was fair to all members					
4. Our team respected everyone's rights					

4. Directions: Please indicate the extent to which you agree or disagree with the following statements **about your team** by circling the appropriate number**.** The rating scale goes from 1 – **totally disagree** to 9 – **totally agree.**

1. Our team avoided taking risks (e.g., strictly followed the instructor's guidelines) when completing our case presentation

Totally disagree 1 2 3 4 5 6 7 8 9 Totally agree

2. Our team took risks (e.g., did something creative) when completing our case presentation

Totally disagree 1 2 3 4 5 6 7 8 9 Totally agree

3. Our team disliked not knowing exactly how to proceed with the case presentation

Totally disagree 1 2 3 4 5 6 7 8 9 Totally agree

4. When completing our case presentation, we took a risk because we saw it as a challenge

Totally disagree	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Totally agree**

5. Our team was defined as…

Risk Avoiders	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Risk Seekers**

5. Directions: Please read each statement and then rate to what extent the statement is true for **your team.** The rating scale goes from 1 – **strongly disagree** to 7 – **strongly agree**.

STATEMENT	**1 Strongly Disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
1. Our team was cautious when completing our case presentation					
2. We followed the belief that 'one must take risks to achieve something' when completing our presentation					
3. Our team did not enjoy risky decisions					
4. We stayed on the safe side in completing our presentation					
5. Success in group work made our team take more risks					
6. Each decision related to our case presentation was made carefully and accurately					
7. Members of our team were not afraid to express their opinions					

6. Directions: This question consists of statements about your team, and how your team functions as a group. Please indicate the extent to which each statement **describes your team.** The rating scale goes from 1 – **Strongly disagree** to 5 – **Strongly agree.**

STATEMENT	1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
1. As a member of this team, I had a real say in how we carried out our work					
2. Most members of my team got a chance to participate in decision-making					
3. My team was designed to let everyone participate in decision-making					
4. Members of my team were very willing to share information with other team members about our work.					
5. Members of my team cooperated to get our work done.					
6. Members of my team varied widely in their skill expertise.					
7. Members of my team had a variety of different backgrounds.					
8. Members of my team had skills and abilities that complemented each other.					

7. Directions: This question consists of statements about your team, and how your team functions as a group. Please indicate the extent to which each statement **describes your team.** The rating scale goes from 1 – **Strongly disagree** to 5 – **Strongly agree**

STATEMENT	1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree

1. Team members always knew others' emotions from their behavior					
2. Team members were good observers of others' emotions					
3. Team members were sensitive to the feelings and emotions of others					
4. Team members had good understanding of the emotions of people around them					

8. Directions: How often has your team engaged in the following activities when completing your case presentation together? The rating scale goes from 1 – **Never** to 5 – **Everyday**

STATEMENT	**1 Never**	**2 Once or Twice**	**3 Once or twice a month**	**4 Once or twice per week**	**5 Everyday**
1. Started harmful rumors					
2. Been nasty or rude to each other					
3. Insulted each other about your performance					
4. Made fun of each other's personal life					
5. Ignored each other					
6. Blamed each other for mistakes					
7. Started arguments with each other					
8. Verbally abused each other					
9. Insulted each other					

PART B: Demographic Questions

1. **Age in years: _____**

2. Please select your gender:

___Male
___Female
___Other

3. Please select your ethnicity:

___Caucasian
___Black
___Hispanic
___Latino
___Asian
___Aboriginal
___Arab
___Multiracial
___Would rather not say
___Other

4. Please select your GPA (Please remember that your selection is anonymous):

___4.3
___4.0
___3.7
___3.3
___3.0
___2.7
___2.3
___2.0
___1.7
___1.0
___0.0

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Thank you for taking the time to participate in our study about the decision-making styles and behaviors of organizational teams. The data collected from these surveys will help to develop and test an instrument based on gendered decision-making behaviours of teams (i.e., teams being masculine, feminine, or androgynous). The findings of this study will allow us to conduct future research testing the gendered behaviors of teams relative to performance in team-based tasks.

Please remember that any data pertaining to you, as an individual participant will be anonymous. Once all the data is collected, it will be saved in a password-encrypted file. The data will be destroyed after a period of five years. Any papers that we plan to publish will only report the aggregate findings and group level data – no identifiable information will be used.

If you would like to follow up with the researchers to express any concerns, receive more information on the study, or have the final research report shared with you directly, please contact one of the researchers by e-mail or telephone: Danielle Mercer, danielle.mercer@smu.ca or (902) 210-4293; or Dr. Catherine Loughlin, Catherine.loughlin@smu.ca or (902) 491-6328. Once the study is complete, anticipated by fall 2017, we will happily e-mail a summary of the results to you.

Again, this research project (File number – 16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board by e-mail ethics@smu.ca or by telephone (902) 420-5728.

Thank you for your participation.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

Appendix I
Study 2 Median Split Results

Multiple Analysis of Variance Approach

Median Split Classification and Dependent Variables

Following the recommendations of Spence et al. (1975) and Bem (1977), I also computed median splits and conducted a MANOVA followed by a series of ANOVAs with each of my dependent variables (i.e., aggregated risk propensity, aggregated ethical benevolence and ethical egoism, and team performance). Each team was coded into four categories of feminine, masculine, androgynous, and undifferentiated based on the median scores of the feminine and masculine scales. Across the sample, the median score for the aggregated masculine scale was 3.35, and the median score for the aggregated feminine scale was 4.22. Therefore, any team that scored above the masculine and feminine median were classified as androgynous (n = 63 participants, clustered into 17 teams*)*, below the masculine and feminine median were classified as undifferentiated (n = 66 participants, aggregated into 18 teams), only above the feminine median were classified as feminine (n = 52 participants, clustered into 13 teams*)*, and only above the masculine median were classified as masculine (n = 49 participants, aggregated into 13 teams).

Between-subjects MANOVA was performed on the four dependent variables – risk propensity, ethical benevolence, ethical egoism, and team performance, with the independent variable being gender classification (i.e., masculine, feminine, androgynous, or undifferentiated). As suggested by Meyers et al. (2013), Bartlett's test of sphericity was used to ensure that the dependent variables had sufficient correlation, which is indicated by statistical significance (Approximate Chi-Square = 46.33, $p < .001$).

Furthermore, Box's *M* was also significant (Box's $M = 167.45$, $p < .001$), which demonstrates that my dependent covariance matrices are not equal across my gender classification groups.

Using Wilk's Lambda as the criterion, gender classification, Wilk's $\Lambda = .627$, $F(12, 590) = 9.485$, $p < .001$, was significantly associated with the dependent variables. A series of univariate ANOVAs with Tukey's post hoc tests are conducted below on each of the dependent variables. See Table 3.6 for relevant means and standard deviations.

Risk Propensity

To further explore the links between risk propensity and gender, I conducted a univariate ANOVA. The results of the univariate ANOVA indicated that there was a significant effect between risk propensity and the classified groups $F(3, 226) = 11.442$, $p < .001$. More specifically, a Tukey post hoc tested revealed that masculine [M = 5.7, SD = .924] and androgynous teams [M = 5.6, SD = 1.15] were significantly higher in risk propensity than feminine [M = 5.0, SD = .74] and undifferentiated teams [M = 4.8, SD = 1.14]. As such, this result shows further support for hypothesis 8a.

Ethical Responsibility

First, I explored the links between ethical benevolence and gender by conducting a univariate ANOVA. The results showed a significant main effect, $F(3, 226) = 13.403$, $p < .001$. As hypothesized, the Tukey post hoc test demonstrated that feminine [M = 4.0, SD = .42] and androgynous teams [M = 4.0, SD = .56] were significantly higher in ethical benevolence than masculine [M = 3.6, SD = .44] and undifferentiated teams [M = 3.8, SD = .48] – further confirming hypothesis 8b.

Second, I examined ethical egoism and gender, which also resulted in statistical significance, $F(3, 226) = 17.504$, $p < .001$, whereby masculine [M = 2.2, SD = .52] and

undifferentiated teams [M = 2.08, SD = .54] were significantly higher in egoism than feminine [M = 1.73, SD = .43] or androgynous teams [M = 1.62, SD = .45]. Based on these results, hypothesis 8c was further supported.

Team Performance

When examining team performance and the classified gender groups, there was no statistical significance on the post hoc test, F (3, 226) = 2.324, p = .076. As such, as also evident in the regression analysis, hypothesis 8e was rejected.

Table 3.5
Means and Standard Deviations for Risk Propensity, Ethical Benevolence, Ethical Egoism, and Team Performance. *(N = 230)*

	Risk Propensity		Ethical Benevolence		Ethical Egoism		Team Performance	
Classification	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Feminine	4.95	.74	4.0*	.42	1.73	.43	82	5.0
Masculine	5.69*	.92	3.6	.44	2.2	.52	85	5.0
Androgynous	5.58*	1.15	4.0*	.56	1.61	.45	81	8.0
Undifferentiated	4.78	1.14	3.53	.48	2.08	.54	82	9.1

Note. The higher the mean, the higher the level of risk propensity, ethical benevolence, ethical egoism, and team performance where. *p < .01.

Appendix J
Study 3 Survey Materials

INFORMED CONSENT FORM
Women in leadership: Could more androgynous teams better navigate the paradoxes in complex decision-making? (REB File #16-082)

Researchers: Danielle Mercer, PhD Candidate, & Dr. Catherine Loughlin
Department of Management, Saint Mary's University, Halifax, NS

Contact Information: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293
Dr. Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328

PURPOSE OF THIS RESEARCH

Have you ever-wondered how high performing teams behave when making decisions? As organizations become more complex and decentralized, researchers and practitioners alike are intrigued in determining the factors that lead to effective team-based decision-making. This study does just that. **You are invited to participate in a team decision-making task called the Winter Survival Exercise and complete a questionnaire that assesses gendered behaviors of business teams.**

IF YOU CHOOSE TO PARTICIPATE

As a participant, you will be asked to participate in a decision-making scenario called the "Winter Survival Exercise" then complete a pencil and paper survey by answering various statements about your team using a rating scale. The decision-making scenario and survey will take approximately 60 minutes of your time. As the researchers, we will not collect any identifiable information; therefore, your answers to the Winter Survival Exercise and responses to the survey will be anonymous and only the researchers will ever see the anonymous responses. However, you have the option to submit your name and e-mail for a chance to win a Cineplex movie ticket package.

Remember you are under no obligation to participate. Participation in this survey is voluntary and you a free to withdraw without prejudice or penalty. Your decision to participate or not participate will in no way influence your relationship with the researchers of Saint Mary's University. However, please note that once the surveys have been submitted, you will no longer be able to withdraw since there is no identifiable information connected to your evaluation. You are not obligated to answer questions that you feel are objectionable or which make you feel uncomfortable. There are no known physical, economic, or social risks to your participating in this study. Your decision to participate will be indicated by submitting the completed survey to the researcher.

HOW THE RESEARCH WILL BE USED

The research will initially be used for Danielle Mercer's PhD dissertation. Please note that we will only write and publish papers reporting the aggregate findings of this study for academic conferences and journals. Only group level data will ever be reported. Again, there will be no way that the researchers can tell who has completed which survey. All paper responses will be locked in the researcher's office in accordance with the American Psychological Association guidelines for data storage. For data analysis, we will record the survey responses and grades on a password-protected computer in a password-protected file. The computer-transferred data will be retained for five years after publication, after which it will be securely destroyed.

POTENTIAL BENEFITS OF THE RESEARCH

There are many potential benefits from the results of this study. Our research will not only produce publications in academic and practitioner journals, but we believe that our results may assist in future policy change in both organizations and the government. As a participant, you are helping us in make a positive change in the role of gender, decision-making styles, and teams such as boards of directors.

QUESTIONS ABOUT THE RESEARCH

Should you wish to receive a report with the details of this study or should you have any questions about this research please contact one of the researchers either by email or by telephone: Danielle Mercer – danielle.mercer@smu.ca or (902) 210-4293; or Catherine Loughlin – Catherine.loughlin@smu.ca or (902) 491-6328. We anticipate that a report will be available in the fall of 2017.

Certification:

The proposal for this research (File number – 16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board at ethics@smu.ca or by telephone at (902) 420-5728.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

PART A: Team Questionnaire

1. **Directions:** Using the following scale below, please **check** whether you agree or disagree with each statement **occurring in your team as your worked on the Winter Survival Exercise**. The rating scale goes from 1 – **strongly disagree** to 5 – **strongly agree**.

STATEMENT	**1 Strongly Disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
1. Our team openly listened to each others issues when completing the exercise					
2. Our team generated a solution to the exercise as quickly as possible					
3. We were perceptive to each others suggestions					
4. Our team followed the motto 'nothing ventured, nothing gained' when making our decision					
5. We considered the rights of everyone to arrive at a final decision					
6. Our team was willing to take risks when completing the exercise					
7. We used our competitiveness to successfully complete the exercise					
8. We ensured that everyone on our team participated in the exercise					
9. Our team saw the exercise as a game to be won					
10. Our team encouraged open communication					
11. We used assertiveness to reach a solution					
12. Kindness was important to the success of the team					

2. Directions: Please indicate the extent to which you agree or disagree with the following statements **about your team** by circling the appropriate number**.** The rating scale goes from 1 – **totally disagree** to 9 – **totally agree.**

1. Our team **avoided** taking risks (e.g., ranked the most obvious survival items first) when completing the exercise

Totally disagree	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Totally agree**

2. Our team **took** risks (e.g., when ranking the survival items) when completing the exercise

Totally disagree	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Totally agree**

3. Our team was defined as…

Risk Avoiders	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Risk Seekers**

3. Directions: Please answer the following in terms of **how it really is in your team**, ***not*** how you would prefer it to be. Be as candid as possible. The rating scale goes from 1 – **mostly false** to 5 – **completely true.**

STATEMENT	**1 Mostly False**	**2 Somewhat False**	**3 Somewhat True**	**4 Mostly True**	**5 Completely True**
1. When completing the exercise, certain team members were only concerned for themselves					
2. When working together on the exercise, there was no room for personal views					
3. When completing the exercise, certain team members cared about their own interests instead of the team as a whole					
4. Each team member was included before making the final decision					
5. We looked out for each other when completing the exercise					
6. Our major concern was what was fair for all team members					
7. People on the team were concerned about what was best for everyone					

4. Directions: Please read each statement and then rate to what extent the statement is true for **your team.** The rating scale goes from 1 – **strongly disagree** to 7 – **strongly agree**.

STATEMENT	1 Strongly Disagree	2 Disagree	3 Neutral	4 Agree	5 Strongly Agree
1. Our team was cautious when completing the exercise					
2. We did not take any risks when ranking the items					
3. We stayed on the safe side when completing the exercise					
4. Each decision related to ranking items was made carefully					
5. We were concerned of our outcome when completing the exercise					

5. Directions: Please indicate whether you agree or disagree with each of the following statements **about your team.** The rating scale goes from 1 – **not at all important** to 5 – **completely important.**

STATEMENT	1 Not at all Important	2 Of Little Importance	3 Of Average Importance	4 Very Important	5 Completely Important
1.Our team took responsibility for each other					
2. We tried to achieve our overall goal					
3. Our team was fair to all members					
4. Our team respected					

everyone's rights					

6. Directions: This question consists of statements about your team, and how your team functioned as a group. Please indicate the extent to which each statement **describes your team.** The rating scale goes from 1 – **Strongly disagree** to 5 – **Strongly agree**

STATEMENT	**1 Strongly Disagree**	**2 Disagree**	**3 Somewhat Disagree**	**4 Neither Agree or Disagree**	**5 Somewhat Agree**	**6 Agree**	**7 Strongly Agree**
1. All team members had a chance to express opinions							
2. Team members listened to each other's input							
3. Members held back in fear of what others thought							
4. Members were free to make positive and negative comments							

7. Directions: This question consists of statements about your team, and how your team functions as a group. Please indicate the extent to which each statement **describes your team.** The rating scale goes from 1 – **Strongly disagree** to 5 – **Strongly agree**

STATEMENT	**1 Strongly Disagree**	**2 Disagree**	**3 Neutral**	**4 Agree**	**5 Strongly Agree**
1. As a member of this team, I had a real say in how we completed the exercise					

2. Most members in my team got a chance to participate in ranking the items					
3. My team allowed everyone to participate in decision making					

8. Directions: This question consists of statements about your team, and how satisfied you are with your team. Please indicate the extent to which each statement **describes your team.** The rating scale goes from 1 –**Very Dissatisfied** to 5 – **Very Satisfied**

STATEMENT	**1 Very Dissatisfied**	**2 Somewhat Dissatisfied**	**3 Neutral**	**4 Somewhat Satisfied**	**5 Very Satisfied**
1. All in all, how satisfied are you with your members of your team?					
2. All in all, how satisfied are you with your team's performance on the exercise?					
3. How satisfied are you with the progress you made in the exercise?					
4. Considering the effort your put into the exercise, how satisfied are you with your team's performance?					

9. Directions: This question consists of statements about your team's performance. Please indicate the extent to which each statement **describes your team.**

STATEMENT	**1 Not Productive at all**	**2 Somewhat Productive**	**3 Fairly Productive**	**4 Productive**	**5 Very Productive**
1. How productive do you think your					

team was during the exercise?					

STATEMENT	1 Very Poor	2 Poor	3 Acceptable	4 Good	5 Very Good
2. How well do you think you worked together as a team?					

10. Directions: This question consists of several adjectives describing your team. Please circle the appropriate number as it relates **to your team as you completed this exercise.**

Supportive	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Not Supportive**
Understanding	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Not Understanding**
Sensitive	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Insensitive**
Caring	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Not Caring**
Strong	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Weak**
Assertive	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Not Assertive**
Bold	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Timid**
Active	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Passive**
Dominant	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**Submissive**

PART B: Demographic Questions

1. Age in years: _____

2. Please select your gender:

___Male
___Female
___Other

3. Please select your ethnicity:

___Caucasian
___Black
___Hispanic
___Latino
___Asian
___Aboriginal
___Arab
___Multiracial
___Would rather not say
___Other

4. Please select your GPA (Please remember that your selection is anonymous):

___4.3
___4.0
___3.7
___3.3
___3.0
___2.7
___2.3
___2.0
___1.7
___1.0
___0.0

5. Have you ever had any survival training?

___Yes
___No
___Unsure

6. Have you completed the Winter Survival Exercise before?

___Yes
___No

SOBEY SCHOOL OF BUSINESS
SAINT MARY'S UNIVERSITY

Thank you for taking the time to participate in our study about the decision-making styles and behaviors of organizational teams. The findings of this study will allow us to conduct further future research testing the gendered behaviors of teams relative to performance in team-based tasks.

Please remember that any data pertaining to you, as an individual participant will be anonymous. Once all the data is collected, it will be saved in a password-encrypted file. The data will be destroyed after a period of five years. Any papers that we plan to publish will only report the aggregate findings and group level data – no identifiable information will be used.

If you would like to follow up with the researchers to express any concerns, receive more information on the study, or have the final research report shared with you directly, please contact one of the researchers by e-mail or telephone: Danielle Mercer, danielle.mercer@smu.ca or (902) 210-4293; or Dr. Catherine Loughlin, Catherine.loughlin@smu.ca or (902) 491-6328. Once the study is complete, anticipated by fall 2017, we will happily e-mail a summary of the results to you.

Again, this research project (File number – 16-082) has been reviewed and cleared by Saint Mary's University Research Ethics Board. If you have any ethical concerns about the research (such as the way you have been treated or your rights as a participant), you may contact the Chair of the Research Board by e-mail ethics@smu.ca or by telephone (902) 420-5728.

Thank you for your participation.

Sincerely,

Danielle Mercer, MBA, PhD Candidate
Catherine Loughlin, PhD

Answer Key

Mid-January is the coldest time of the year in Minnesota and Manitoba. The first problem the survivors' face, therefore, is to preserve their body heat and protect themselves against its loss. This problem can be met by building a fire, minimizing movement and exertion, and using as much insulation as possible. The participants have just crash-landed. Many individuals tend to overlook the enormous shock reaction this has upon the human body, and the death of the pilot and co-pilot increases the shock. Decision making under such conditions is extremely difficult. Such a situation requires a strong emphasis upon the use of reasoning not only to make decisions, but also to reduce the fear and panic every person would naturally feel. Along with fear, shock reaction is manifested in the feelings of helplessness, loneliness, and hopelessness. These feelings have brought about more fatalities than perhaps any other cause in survival situations. Through the use of reasoning, hope can be generated. Certainly the state of shock means that movement of individuals should be at a minimum and that an attempt to calm them should be made.

Before taking off a pilot always has to file a flight plan. The flight plan contains the vital information regarding the flight, such as the course, speed, estimated time of arrival, type of aircraft, number of people on board, and so on. Search-and-rescue operations would begin shortly after the plane failed to arrive at its destination at its estimated time of arrival. The eighty miles to the nearest known town is a very long walk even under ideal conditions, particularly if one is not used to walking such distances. Under the circumstances of being in shock, dressed in city clothes, having deep snow in the woods, and a variety of water barriers to cross, to attempt to walk out would mean almost certain death from freezing and exhaustion. At the temperatures given, the loss of body heat through exertion is a very serious matter. Once the survivors have found ways in which to keep warm, their most immediate problem is to provide signaling methods to attract the attention of search planes and search parties. Thus, all the items the group has must be assessed according to their value in signaling the group's whereabouts.

Items	**Ranking**
Compress kit (with 28 feet of 2 inch gauze)	11
Sectional air map made of plastic	14
Ball of steel wool	2
30 feet of rope	7
Cigarette lighter (without fluid)	1
Family-size chocolate bar (one per person)	4
Loaded .45 caliber pistol	9
Flashlight with batteries	6
Newspaper (one per person)	8
Quart of 85 proof whiskey	13
Compass	15
Extra shirt and pants for each survivor	3
Two ski poles	12
Knife	10
Can of shortening	5

Rank 1: Cigarette lighter (without fluid). The gravest danger facing the group is exposure to the cold. The greatest need is for a source of warmth and the second greatest need is for signaling devices. This makes building a fire the first order of business. Without matches something is needed to produce sparks to start a fire. Even without fluid the cigarette lighter can be used to produce sparks. The fire will not only provide warmth, it will also provide smoke for daytime signaling and firelight for nighttime signaling.

Rank 2: Ball of steel wool. To make a fire, a means of catching the sparks made by the cigarette lighter is needed. Steel wool is the best substance with which to catch a spark and support a flame, even if it is a little bit wet.

Rank 3: Extra shirt and pants for each survivor. Clothes are probably the most versatile items one can have in a situation like this. Besides adding warmth to the body they can be used for shelter, signaling, bedding, bandages, string when unraveled, and tinder to make fires. Even maps can be drawn on them. The versatility of clothes and the need for fires, signaling devices, and warmth make this item number three in importance.

Rank 4: Family-size chocolate bar (one per person). To gather wood for the fire and to set up signals, energy is needed. The Hershey bars would supply the energy to sustain the survivors for quite some time. Because they contain basically carbohydrates, they would supply energy without making digestive demands upon the body.

Rank 5: Can of shortening. This item has many uses – the most important being that a mirror-like signaling device can be made from the lid. After shining the lid with the steel wool, the survivors can use it to produce an effective reflector of sunlight. A mirror is the most powerful tool they have for communicating their presence. In sunlight, a simple mirror can generate 5 to 7 million candlepower. The reflected sunbeam can be seen beyond the horizon. Its effectiveness is somewhat limited by the trees but one member of the group could climb a tree and use the mirror to signal search planes. If the survivors have no other means of signaling, they would still have better than 80 percent chance of being rescued within the first twenty-four hours. Other uses for the item are as follows: The shortening can be rubbed on the body to protect exposed areas, such as the face, lips, and hands, from the cold. In desperation it could be eaten in small amounts. When melted into an oil the shortening is helpful in starting fires. Melted shortening, when soaked into a piece of cloth, will produce an effective candlewick. The can is useful in melting snow to produce drinking water. Even in the wintertime water is important as the body loses water in many ways, such as through perspiration, respiration, shock reactions, and so on. This water must be replenished because dehydration affects the ability to make clear decisions. The can is also useful as a cup.

Rank 6: Flashlight. Inasmuch as the group has little hope of survival if it decides to walk out, its major hope is to catch the attention of search planes. During the day the lid-mirror, smoke, and flags made from clothing represent the best devices. During the night the flashlight is the best signaling device. It is the only effective night-signaling devices beside the fire. In the cold, however, a flashlight loses the power in its battery very quickly. It must therefore, be kept warm if it is to work, which means that it must be kept close to someone's body. The value of the flashlight lies in the fact that if the fire burns low or inadvertently goes out, the flashlight could be immediately turned on the moment a plane is heard.

Rank 7: Piece of rope. The rope is another versatile piece of equipment. It could be used to pull dead limbs of trees for firewood. When cut into pieces, the rope will help in constructing shelters. It can be burned. When frayed it can be used as tinder to start fires. When unraveled it will make good insulation from the cold if it is stuffed inside clothing.

Rank 8: Newspaper (one per person). The newspaper could be used for starting a fire much the same as the rope. It will also serve as an insulator; when rolled up and placed under the clothes around a person's legs or arms, it provides dead-air space for extra protection from the cold. The paper can be used for recreation by reading it, memorizing it, folding it, or tearing it. It could be rolled into a cone and yelled through as a signal device. It could also be spread around an area to help signal a rescue party.

Rank 9: .45-caliber pistol. This pistol provides a sound-signaling device. (The international distress signal is three shots fired in rapid succession.) There have been numerous cases of survivors going undetected because by the time the rescue party arrived in the area the survivors were too weak to make a loud enough noise to attract attention. The butt of the pistol could be used as a hammer. The powder from the shells will assist in fire building. By placing a small bit of cloth in a cartridge, emptied of its bullet, a fire can be started by firing the gun at dry wood on the ground. At night the muzzle blast of the gun is visible, which also makes it useful as a signaling device. The pistol's advantages are counterbalanced by its dangerous disadvantages. Anger, frustration, impatience, irritability, and lapses of rationality may increase as the group waits to be rescued. The availability of a lethal weapon is a real danger to the group under these conditions. Although it could be used for hunting, it would take a highly skilled marksman to kill an animal and then the animal would have to be transported through the snow to the crash area, probably taking more energy than would be advisable.

Rank 10: Knife. A knife is a versatile tool, but it is not too important in the winter setting. It could be used for cutting the rope into desired lengths, making shavings from pieces of wood for tinder, and many other uses could be thought up.

Rank 11: Compress kit (with gauze). The best use of this item is to wrap the gauze around exposed areas of the body for insulation. Feet and hands are probably the most vulnerable to frostbite, and the gauze can be used to keep them warm. The gauze can be used as a candlewick when dipped into melted shortening. It would also make effective tinder. The small supply of the gauze is the reason this item is ranked so low.

Rank 12: Ski poles. Although they are not very important, the poles are useful as a flagpole or staff for signaling. They can be used to stabilize a person walking through the snow to collect wood, and to test the thickness of the ice on a lakeshore or stream. Probably their most useful function would be as supports for a shelter or by the fire as a heat reflector.

Rank 13: Quart of 85-proof whiskey. The only useful function of the whiskey is to aid in fire building or as a fuel. A torch could be made from a piece of clothing soaked in the whiskey and attached to an upright ski pole. The danger of the whiskey is that someone might try to drink it when it is cold. Whiskey takes on the temperature it is exposed to, and a drink of it at minus thirty degrees would freeze a person's esophagus and stomach and do considerable damage to the mouth. Drinking it warm will cause dehydration. The bottle, kept warm, would be useful for storing drinking water.

Rank 14: Sectional air map made of plastic. This item is dangerous because it will encourage individuals to attempt to walk to the nearest town – thereby condemning them to almost certain death.

Rank 15: Compass. Because the compass may also encourage some survivors to try to walk to the nearest town, it too is a dangerous item. The only redeeming feature of the compass is the possible use of its glass top as a reflector of sunlight to signal search planes, but it is the least effective of the potential signaling devices available. That it might tempt survivors to walk away from the crash site makes it the least desirable of the fifteen items.

Appendix K
Written Permission to use the Winter Survival Exercise

David W. Johnson <dwj@visi.com> Jan 16 at 11:12 PM

To Danielle Mercer

You have my permission to use the Winter Survival Exercise. I think it is in Joining Together: Group theory and group skills, Allyn and Bacon.

David W. Johnson

On Jan 16, 2017, at 5:45 PM, Danielle Mercer <Danielle.Mercer@smu.ca> wrote:

Dear Dr. Johnson:

I am a PhD in Management student at Saint Mary's University in Halifax, NS, Canada - I am hoping to use the Winter Survival Exercise in my dissertation to study gendered decision-making styles of teams. It is readily available online but I either wanted to obtain permission to use it or wonder where I might be able to purchase it? Thank you in advance.

Best,

Danielle Mercer

Appendix L
Winter Survival Exercise Materials

WINTER SURVIVAL EXERCISE

The Situation

You have just crash-landed in the woods of North Minnesota and Southern Manitoba. It is 11:32 a.m. in mid-January. The small plane in which you were traveling has been completely destroyed except for the frame. The pilot and co-pilot have been killed, but no one else is seriously injured. The crash came suddenly before the pilot had time to radio for help or inform anyone of your position. Since your pilot was trying to avoid a storm, you know the plane was considerably off course. The pilot announced shortly before the crash that you were eighty miles northwest of a small town that is the nearest known habitation. You are in a wilderness area made up of thick woods broken by many lakes and rivers. The last weather report indicated that the temperature would reach minus twenty-five degrees in the daytime and minus forty at night. You are dressed in winter clothing appropriate for city wear – suits, pantsuits, street shoes, and overcoats.

The Challenge

While escaping from the plane your group salvaged the fifteen items listed below. Your task is to rank these items according to their importance to your survival. Imagine that:

- The number of people in your group here today is the same number in the plane crash
- You are the actual people who were in the plane crash
- Your group has decided to stick together

The task:

Step 1: **Individually** review the list of 15 items. Without discussing the list of items with your team, rank the items in order of importance to your survival. "1" being the most important and proceeding to "15" for the least important. **You have 10 minutes to complete this step.**

Step 2: **Now as team**, reconsider the 15 items and agree on a new set of ranks together. Again, you will tank these items in order of importance of your survival. "1" being the most important and proceeding to "15" for the least important. **Your team has 25 minutes to complete this step.**

Items	***Step 1:* Individual Ranking**	***Step 2:* Team Ranking**
Compress kit (with 28 feet of 2 inch gauze)		
Sectional air map made of plastic		
Ball of steel wool		
30 feet of rope		
Cigarette lighter (without fluid)		
Family-size chocolate bar (one per person)		
Loaded .45 caliber pistol		
Flashlight with batteries		
Newspaper (one per person)		
Quart of 85 proof whiskey		
Compass		
Extra shirt and pants for each survivor		
Two ski poles		
Knife		
Can of shortening		

Appendix M
Study 3 Median Split Results

Multiple Analysis of Variance Approach

Median Split Classification and Dependent Variables

To further explore links between team gender and the dependent variables (i.e., aggregated ethical egoism and ethical benevolence, aggregated risk propensity, and team performance), I conducted a MANOVA followed by a series of ANOVAs. The median split technique was used to classify teams as feminine, masculine, androgynous, and undifferentiated based on the median scores of the feminine and masculine scales. The median score for the aggregated feminine scale was 4.24 and for the aggregated masculine scale was 3.52. Any team that scored above both the feminine and masculine median were coded as androgynous (n = 80 clustered in 18 teams), below both were coded as undifferentiated (n = 53 clustered in 11 teams), above the feminine median were coded as feminine (n = 59 clustered in 13 teams), and only above the masculine (n = 53 clustered in 11 teams) median were coded as masculine.

Similar to Study 2, between-subjects MANOVA was performed on the four dependent variables – ethical egoism, ethical benevolence, risk propensity, and team performance, with the independent variable being gender classification. Bartlett's test of sphericity was computed to ensure that the dependent variables were correlated, which is demonstrated by statistical significance (Approximate Chi-Square = 69.5, $p < .001$). Additionally, Box's M provided support regarding differences across gender classification and my dependent covariance matrices (Box's M = 248.12, $p < .001$).

Using Wilk's Lambda as the criterion when conducing a MANOVA, gender classification was significant in regard to my four dependent variables, Wilk's $\Lambda = .451$,

F (12, 629) = 18.43, $p < .001$. Therefore, I conducted univariate ANOVAs on each dependent variable using Tukey's Post Hoc tests. Table 4.9 presents the relevant means and standard deviations.

Ethical Responsibility

First, I conducted a univariate ANOVA to examine the differences between gender classification and ethical egoism. As expected, the results demonstrated statistical significance, F (3, 241) = 17.37, $p < .001$. More specifically, masculine [M = 2.3, SD = .61] and undifferentiated teams [M = 2.2, SD = .57] were significantly higher in egoism than feminine [M = 1.71, SD = .55] or androgynous teams [M = 1.73, SD = .41].

Second, I examined the differences between ethical benevolence and gender classification. Again, the results showed a significant main effect, F (3, 241) = 22.32, $p < .001$, whereby undifferentiated teams [M = 3.53, SD = .26] were significantly lower in benevolence than feminine [M = 4.01, SD = .33], masculine [3.96, SD = .28], or androgynous teams [M = 3.99, SD = .46].

Risk Propensity

The results of the univariate ANOVA using Tukey's Post Hoc test of risk propensity and gender classification revealed a significant main effect, F (3, 241) = 33. 1, $p < .001$. As expected, masculine [M = 5.34, SD = .93] and androgynous teams [M = 5.4, SD = .82] were significantly higher in risk propensity when completing the exercise than feminine [M = 3.95, SD = .92] or undifferentiated teams [M = 4.6, SD = .93]. Additionally, there was also significance between feminine and undifferentiated teams whereby feminine teams were the least likely to be pro risk.

Team Performance

When examining the difference between team performance and gender classification, the univariate ANOVA revealed interesting results, $F(3, 241) = 3.896$, $p < .01$. While there were no significant differences between feminine teams [M = 49.20, SD = 23.5], masculine teams [M = 46. 09, SD = 12.76], and undifferentiated teams [M = 45.18, SD = 6.79], androgynous teams [M= 40.16, SD = 15.12] performed significantly poorer.

Table 4.7
Means and Standard Deviations for gender classification and ethical climate, risk propensity, and team performance

	Ethical Egoism		Ethical Benevolence		Risk Propensity		Team Performance	
Classification	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Feminine	1.71	.55	4.00	.33	3.95*	.92	49.2	23.51
Masculine	2.3*	.61	3.96	.28	5.34	.93	46.09	12.76
Androgynous	1.73	.41	3.99	.46	5.4	.82	40.16	15.12
Undifferentiated	2.12*	.57	3.53*	.26	4.62*	1.05	45.18	6.79

Note.The higher the mean, the higher the level of ethical egoism, ethical benevolence, risk propensity, and team performance where. *$p < .05$.