

Running Head: THE WEAKEST LINK

Male Female Dynamics in Groups: A Field Study of *The Weakest Link*

Priya Raghbir

New York University

Ana Valenzuela

Baruch College

Abstract

This paper develops an integrative framework for understanding gender-based group dynamics based on sex composition. We study decisions made by male and female members of a group to eliminate or promote other male and female group members. Study 1 uses game simulations modeled on the television show, *The Weakest Link*, to show how male and female players trade off individual versus group utility in a group setting and how such trade-offs affect group dynamics. We identify four patterns of sex-based group dynamics: *old boy's club*, *queen bee*, *bounded rationality*, and *females as finalists but not as winners*. These patterns are part of a mosaic based on differences in sex composition and sex differences between male and female group members discriminating positively, negatively, or not at all against other male and female group members. We propose that differences in the degree of competition versus cooperation in the group, the initial sex composition, and the mean and variance of players' abilities by sex interactively determine which pattern will be noted. Study 2 uses observational data from *The Weakest Link* to test these predictions using actual behavior in a field setting. Implications for the motivational, cognitive, and strategic antecedents of gender effects based on sex composition of the group are discussed.

How do males and females choose whether to compete or cooperate with one another in small group settings when competitive pressures change over time? This question is an important one for organizational theorists (John & Robins, 1994), psychologists (Eagly & Karau, 2002, Karakowsky & McBey 2001, Vancouver & Ilgen 1989), and gender researchers (Kerr & Beh, 1995). The literature on gender dynamics based on sex composition has identified motivational and cognitive factors that influence the ways in which males and females evaluate one another in a group and predicted that more favorable evaluations translate into higher rewards.

We propose strategic reasons, which go beyond earlier studied cognitive and motivational reasons, to explain how males and females evaluate each other in mixed group interactions. Strategic considerations come into play when there is a tension between the rewards accruing to the overall group versus each of its individual constituents. For example, in some organizations, economic rewards are based not only on overall group performance, but also on each individual's relative performance compared to the rest of the group (e.g., employee of the month). While the overall rewards accruing to the group are based on its absolute performance, the rewards accruing to individuals within the group are proportional to their relative performance versus other group members. In such group interactions, an individual can increase their own chances of being rewarded if others in the group perform poorly. Therefore, the goal of maximizing individual (versus group) rewards could lead to group members preferring to retain poor over good performers in their group. Such situations may result in the expulsion of higher performing group members, since such members may be considered threatening to the others' future rewards. These situations are quite common in

organizations, and could lead to sub-optimal profitability and biased personnel policies. Such group dynamics are observable in a wide range of other group settings in which individual performance aids group performance but individuals compete with each other within the group, such as managerial group discussions (John & Robins, 1994), and the sports field (Kerr & Beh, 1995). We study the specific case of the different ways these strategic dynamics play out when males and females cooperate within a group, but against each other for future rewards.

We report the results of two studies based on the television show, *The Weakest Link*. The television show *The Weakest Link* has previously been used to study biases in decision-making and information processing due to the position of a player within the array (Raghubir & Valenzuela, 2006) and the sex ratio of the players (Valenzuela & Raghubir, 2007), as well as more general theories of discrimination (Levitt 2004). It is a television show that starts with eight contestants who answer trivia questions individually and sequentially. The higher the number of correct questions answered in a sequence the higher the earnings of the group (1 correct = \$1,000; maximum = \$250,000 for 8 consecutively correct answers). However, if the string is broken due to an incorrect response, all the money accumulated due to the string of correct answers is lost. To prevent this loss, at any stage a player can say the word “bank,” which leads to the amount of money accumulated up to that point in the string to be retained towards the final winnings, and a new string to begin at \$1,000. The game has eight rounds. At the end of rounds 1-6, each contestant publicly votes for a player that she or he wishes to remove from the team. The player who receives the most votes is voted out, with ties resolved by asking the strongest player in that round to cast a tie-breaking vote. The

number of players who play each round decreases from 8 to 2 players who play the seventh and eighth rounds. In the seventh round, the winnings of the duo are doubled and in the eighth round, one of the two finalists wins the game and receives the cumulative group earnings.

The Weakest Link is characterized by changing competitive dynamics over time: Initially, the participant's goal is to maximize group performance by retaining stronger players (i.e., maximize group earnings built up through correct answers). However, as rewards become more competitive (i.e., the game nears the final round), the participant's goal changes towards maximizing their individual chances of winning a reward (i.e., reaching and winning the final). This would encourage them to retain a player who they believe they could defeat in the final. Valenzuela and Raghurir (2007) showed that when there is a single winner in the game, higher performing group members retain a token low-performing member who is perceived to be easier to defeat at the end of the game in order to increase their chance of winning the endgame. In the games studied by Valenzuela and Raghurir (2007), on average, males performed better than females. However, females were voted out less often (or later) than their performance justified, while males were voted out more often (or earlier), even when they performed well. This led to a pattern they called *females as finalists but not winners*. Valenzuela and Raghurir (2007) conjecture, and present evidence consistent with the idea, that males retained lower performing females as they were perceived to be easier to defeat in the final. Females, however, retained other lower performing females as a higher proportion of females in a group led them to perform better.

In this paper, we further examine the underlying basis for the *females as finalists but not winners* pattern and develop a comprehensive grid of how males and females compete with each other in groups. In particular, both males and females can display in-group favoritism or derogation, or no in-group bias leading to nine different group dynamics. The case of *females as finalists but not winners* is shown to be a specific case of strategic group dynamics where females display in-group favoritism and males display in-group derogation. Its opposite sex analog (males and females both favoring males) is termed *male chauvinism*. When males do not show an in-group bias but females derogate or favor other females, the dynamics that result are termed *queen bee* and *power girls* respectively. Analogously, when females do not show an in-group bias but males derogate or favor other males, the resulting dynamics are termed *alpha male* and *old boy's club* respectively. The remaining three dynamics are characterized by both sexes demonstrating in-group derogation (*eat your own*), in-group favoritism (*battle of the sexes*) or no in-group bias (*bounded rationality*). These dynamics are summarized in Figure 1.

Next, we develop the conditions that foster these different patterns of competition. We argue that competition fosters strategic game play. Given a competitive environment, strategic group dynamics become contingent on (a) sex composition, (b) the mean performance differential between males and females, and (c) the extent of variance in the performance level within each sex which together determine the need for in-group favoritism or derogation. We test this model using two different methodological approaches. Study 1 reports results of four game simulations in which people play a version of a game based on the television show *The Weakest Link*. The games are used to

identify the group dynamic that developed and are related back to game characteristics (sex composition, and the mean and variance of female and male members' performance levels). Study 2 tests the model using observational data from the television show, *The Weakest Link*. Results support framework predictions. After a brief literature review, we develop our theoretical framework, and describe our methodology and results.

Literature Review

Gender effects in performance evaluation can occur for cognitive, motivational, and/or strategic reasons. These are briefly summarized below and are followed by a conceptual framework that integrates them within a model of how people compete in mixed-sex groups and which antecedents lead to different patterns of competition.

Cognitive Biases

The cognitive bases of biases are typically due to effort-accuracy tradeoffs (Martell, 1996) and are mitigated when people have the ability, opportunity, and motivation to process all incoming cues to make a judgment (Kruglanski & Freund, 1983; Macrae, Hewstone & Griffiths, 1993). A typical manifestation of a cognitive gender bias would be to use a person's sex as a cue to assess an individual's performance level based on gender stereotypes, substituting it for individuating information about the person (Fiske & Neuberg, 1990). Stereotyping can help a perceiver to structure incoming social stimuli and reduce it to a manageable size (Taylor, Fiske, Etcoff, & Ruderman, 1978). Such stereotypes can also affect one's own performance due to people imputing their level of expertise based on the consistency of task characteristics with their gender (Vancouver & Ilgen 1989). In sum, it is possible to find situations in which neither males

nor females display systematic in-group derogation or favoritism but use sex as a cognitive cue to make decisions.

Motivational Biases

People's needs for self-enhancement can lead them to have a preference for or against members of their group (Wills, 1981). When evaluation and rewards are based not only on the sex of the person being evaluated or rewarded, but are further contingent upon the sex of the person performing the evaluation/ conferring the reward, there are two possible directions for an in-group bias: *favoritism* or *derogation*.

In-group favoritism is defined as evaluating and rewarding members of one's own sex higher than members of the opposite sex. The literature on inter-group relations has found that social categorizations on the basis of gender, race, and occupation accentuate within-group similarities and between-group differences on evaluative and behavioral dimensions (Doise & Sinclair, 1973). If the in-group is relatively successful, these social comparisons allow a person to attain a positively-valued distinctiveness (Brewer, 1979; Tajfel & Turner, 1979). Further, minority group members identify more strongly with their group and evaluate their group more positively than members of the dominant group (Bettencourt, Miller & Hume, 1999; Jackson, 1997, 1999; Otten, Mummendey & Blanz, 1996). This implies that the lower the proportion of one's own sex in the group, and the higher the relative performance of one's own sex versus the opposite sex, the higher the likelihood of in-group favoritism.

In-group derogation is defined as rewarding members of one's own sex lower than members of the opposite sex: That is, females evaluating other females worse than the average male; or males evaluating other males worse than the average female. If an

in-group's performance is worse than the performance of an individual member of that group (i.e., female performance has a lower mean than male performance and a higher variance with a single female performing better than the average female performance), then self-enhancement pressures would inhibit in-group favoritism. Consistent with this observation, Seta and Seta (1992) found that participants in a relatively unsuccessful group showed lower in-group bias, favored in-group members less, and, in some cases, even questioned group identification. This suggests that the greater the variance in the performance of one's sex, the higher the likelihood of in-group derogation.

Additionally, a player's performance affects the extent to which they use data based cues to assess another player's performance. Pyszczynski, Greenberg, and LaPrelle (1985) explore biases in information search that lead to self-serving attributions for success and failure, and for maintaining positive self-evaluations. They showed that information search about performance outcomes is biased so as to provide evidence consistent with a favorable self-evaluation. Study participants who had not performed well asked for more information only when they expected it to reveal that others had also performed poorly. Those who had performed well did not ask for additional information, regardless of their beliefs about whether it would reveal poor or strong performance of others. Thus, the motivation to maintain a positive self-evaluation in a context when others in one's group are performing at a level lower than oneself, suggests that the greater the performance variance within a sex the higher the likelihood that higher-performing members of that sex will retaliate against lower-performing members of their own sex to preserve the group's superiority.

To summarize, in-group favoritism is more likely when the sex composition is skewed and the mean performance levels of the two sexes differ, whereas in-group derogation is more likely when there is variance in the performance level within a sex.

Strategic Biases

Competition between group members is denoted as intense when an individual's chances of being rewarded become dependent on his or her performance relative to the performance of others in the group. This situation often leads to strategic game behavior aimed at increasing relative performance differential through reducing the mean performance level of other group members (threat reduction), or improving one's own performance level (performance improvement). The literature shows that a key antecedent to seeing strategic threat reduction behavior is the presence of competition, which leads to a divergence between the goals of maximizing own utility and group performance (e.g., Valenzuela and Raghurir 2007). When there is a low level of competition, evaluations are consistent with the reward structure as those who are perceived to be high performing are retained in the group. However, when competition increases, there is a growing need for group members to retain those who are a lower threat (i.e., poor performers). It is at this stage that evaluations diverge from rewards. Poor performers may be retained not because they add value to the group, but because they are perceived to be easier to eliminate as the group proceeds to identify a single winner.

While the key antecedent for strategic threat reduction behavior is competition, the key antecedent to seeing strategic performance improvement behavior is sex composition. Research analyzing group performance based on group characteristics has

found that players' performance levels are contingent upon contextual variables such as group size (Shaw, 1976), task type (Steiner, 1972), and sex composition (Bray, Kerr, & Atkin, 1978). Even when the performance differential between all-male and all-female groups was not significant (Bray et al., 1978), mixed-sex groups seemed to facilitate male achievements (Brophy, 1985) and limit female opportunities (Griffin, 1983). Lirgg (1994) found that girls perceive same-sex classes more favorably than mixed-sex classes, whereas the opposite was true for boys. Sleeper and Nigro (1987) found that before beginning work on an experimental task, both male and female subjects working with female partners expressed higher self-confidence than those working with male partners. It is possible that participants' self-confidence ratings reflected their beliefs about how competitive and aggressive a male or female partner would be (Deaux & Lewis, 1984). In a similar vein, Heilman and Kram (1978) report that both males and females took more credit for success when they worked with a female than when they worked with a male. This would suggest that females would be more likely to be retained so as to improve other players' own performance levels, particularly of those players who have lower levels of self-confidence in the given task.

In support of these two strategic routes, Valenzuela and Raghubir (2007) proposed that, in a situation characterized by competition among individuals in a group, strategic game play could affect behavior. They found that male group players retained non-threatening female players until the endgame in order to enhance individual chances of winning. Female group members also retained female players in their group, but because the presence of other females created a more supportive environment in which their performance was enhanced. As a result, female players were more likely to be

retained in a group than was warranted by their performance but were no more likely to win the final game.

To summarize, competition within the group and a skewed sex composition increases strategic considerations in a game. A skewed sex composition also increases the likelihood of in-group favoritism particularly when the mean performance of the two sexes differs. Finally, higher variance in the performance level within a sex group increases the likelihood of in-group derogation. We now propose nine different patterns of group dynamics, which are a function of the effect of competitive intensity, sex composition, and the relative mean difference and variance of male-female performance levels. These predictions are translated into a set of operational hypotheses tested in two studies.

Group Dynamics

The main and interactive effects of level of competition (round played), sex composition, and mean and variance of performance of males and female group members affect motivational and strategic considerations and lead to a mosaic of competitive male-female group dynamics (Figure 1).

When there is no in-group motivational bias or strategic aspect to a mixed-group interaction, then sex may at best be used as a shortcut to process incoming cues. Cognitive gender-based biases may occur at earlier stages of a group interaction when identifying group members' performance levels is an effortful task. This pattern is reflected in the cognitive cue-based dynamic: *bounded rationality*. When strategic considerations are low, both sexes may also just derogate (termed *eat your own*) or favor their own sex (termed *battle of the sexes*).

However, competition increases strategic considerations in game play. This suggests:

H1: In competitive scenarios (as those seen in later rounds of the game), players become more likely to vote out other players for strategic reasons instead of merely due to their perceived performance (as in earlier rounds of the game).

In competitive scenarios, strategic considerations of threat reduction or performance improvement combined with in-group derogation or favoritism lead to six possible patterns: Males and females derogating or favoring their own sex (four patterns) or males or females being favored overall (two patterns). As argued earlier, in-group derogation is a function of variance within the sex, and in-group favoritism is a function of male-female performance differential and variance. Specifically:

If females derogate other females, the *queen bee* pattern comes into play where a strong performing female member eliminates weaker female participants so that she is the only female remaining. We propose that this dynamic is more likely to occur the higher the variance in the performance of females.

If males derogate other males, we see the *alpha male* pattern, where a strong performing male member eliminates other males so that he is the strongest male remaining. This dynamic is more likely to occur the higher the variance in the performance of males.

Operationally, we test:

H2: Scenarios where there is a higher (versus lower) level of variance in the performance of males (females) will have a higher likelihood of males (females) retaining female (male) players.

If females favor other females, we see the *power girls* dynamic where female group members collude to exclude males. This dynamic is more likely to occur the higher

the relative average performance of females versus males, and the lower the variance in the performance of females.

If males favor other males, the *old boy's club* dynamic is seen where male group members collude to retain other males in the group. This dynamic is more likely to occur the higher the relative average performance of males versus females, and the lower the variance in the performance of males.

However, if the variance in the female performance is high, then this pattern can change to the *male chauvinism* one where both males and females favor males. The analog of the male chauvinism pattern is the *females as finalists but not winners pattern* where both males and females favor females (Valenzuela and Raghubir 2007). In this dynamic, males retain lower-performing female members as they expect to be able to win against them in a final round of competition (the threat-reduction mechanism), and females retain them to improve their own performance (the performance-improvement mechanism). We now propose that this dynamic is more likely to occur the higher the relative mean and variance of male versus female performance is. Operationally:

H3: Scenarios where male performance is higher than or equal to female performance will have a higher likelihood of:

- i. Males retaining male players when the variance in male performance is low.
- ii. Males retaining female players when the variance in male performance is high.
- iii. Females retaining female players when the variance in female performance is low.
- iv. Females retaining male players when the variance in female performance is high.

H4: Scenarios where female performance is higher than male performance will have a higher likelihood of females retaining female players, especially when the variance in female performance is low and the variance in male performance is high.

Note that the *old boys club* pattern comes into play when condition foster H3 i. H3 i and iv together lead to the *male chauvinism* pattern, whereas H3 ii and iii together lead to the *females as finalists but not winners* pattern, and H4 predicts the *power girls* pattern. Study 1 examines H1 to investigate the presence and increase in strategic game play as competition increases. Studies 1 and 2 test H2-H4 using different methods.

Study 1 is a set of four experiential game simulations that are analyzed at the overall game level. We measure the mean and variance in male and female performance in each game separately to test whether males and females are more likely to retain or eliminate other males or females as predicted by H2-4, and then relate this voting behavior to stated reasons for voting to assess whether there is evidence for the specific dynamic hypothesized.

Study 2 is an observational study of 20 episodes of *The Weakest Link* (Summer 2001), where we examine whether male and female players vote for their own or the opposite sex. This data set has previously been used to examine the effect of player position on winning the game (Raghubir and Valenzuela 2006), and the effect of player sex on winning the game (Valenzuela and Raghubir 2007). However, the effect of the sex of the player casting the vote has not previously been analyzed. The observational data is analyzed at three different levels of aggregation: episode level ($n = 20$), round level per episode ($n = 6$ voting rounds \times 20 episodes = 120), and individual vote per round per episode ($n = 120$ rounds \times 33 votes per round = 660).

Study 1: Experiential Game Simulations: *Give Me the Money*

Methodology

Each session involved a simulated game called *Give Me the Money*, based on the television show, *The Weakest Link*. In the simulated game, the moderator was blind to the hypotheses and repeated the instructions of the original game. Questions were chosen from the trivia game, *Who Wants to be a Millionaire?* Participants could win a maximum of \$1,000 overall in 8 rounds, with increments based on the original game (\$1, \$2, \$4, \$8, \$16, \$32, \$64, and \$125). All participants played the game with the following variation: At the end of each round, each player wrote on one side of an index card the name of the person whom they were voting out of the game, and described, on the reverse side, the reasons why they voted the way they did. Players were not allowed to talk to each other during the game.

Two researchers observed the game under the guise of helpers. One kept track of the link of correct answers and the other recorded performance and money banked. Their observations, along with participants' stated reasons, and a comparison of the stated reasons with each player's estimations of other players' accuracy levels, were used to characterize game dynamics. Each game took one hour and was videotaped.

At the end of the game participants completed a debrief questionnaire in which they estimated each player's accuracy throughout the game, described why they believed each person was voted out in each round, why they believed the winner and the finalist made it to the final rounds of the game, and what advice they would give to other players.

Experimental participants were undergraduates who undertook the study for partial course credit. There were eight participants per game, and four game simulations ($n = 32$, males = 17, females = 15). Students could sign up for the different sessions, allowing for different starting compositions. Male: female starting ratios were: 3:5, 3:5,

5:3, and 6:2, respectively. There was one male-female final, two all-male finals, and one all-female final. The participant who won each game simulation was paid his or her winnings (\$41, \$22, \$47, and \$47). Males won three of the four game simulations.

Results

Hypothesis 1. To examine whether there were any strategic reasons stated, we content analyzed the reasons respondents mentioned for voting out a person themselves while playing. To test whether players were more likely to vote out other players for strategic reasons versus due to their perceived poor performance in later rounds of the game, as argued by H1, we content analyzed the reasons stated by all players at the end of the game as to why a player was voted out in round 1-6. Two judges coded reasons as being performance related (e.g., “She did not answer any questions correctly”), strategic (e.g., “He was a strong opponent”), or other (e.g., “He was next clockwise not including myself”). Inter-rater agreement was high (90.9%), with differences resolved through discussion. Participants could mention more than one reason.

The analysis of the reasons provided at the time of voting out a player ($n = 131$, non-response = 1) support the presence of strategic reasons for voting. While overall, the majority of the reasons provided were poor performance-related (78%), as many as 9.2% ($n = 12$) were strategy-related. Seven different players of whom five were male mentioned strategic reasons (or 22% of all players).

To test whether the incidence of strategic reasons increases in later rounds, we examined the percentage of occasions when performance and strategic reasons were cited in each round of the game. Consistent with hypothesis 1, in earlier rounds of the game respondents cited poor performance as the primary reason why a player was voted out

(87.5%, 96.87%, 84.37%, 68.75%, and 62.5% for rounds 1 through 5 across all four games, all binomial tests $p < .05$). This percentage reduced to chance levels in the last voting round, round 6 (46.87%). In round 6, most respondents stated that players were being voted out because they were perceived as a threat (75%, $p < .05$). Threat related reasons had barely been mentioned in the first three rounds of the game (0%, 0%, and 3.12% respectively). Thus, as predicted by H1, as the game progresses, players change from eliminating weak players for performance related reasons to eliminating strong players for strategic reasons.

To test H2-H4, we first examine the characteristics of the four experiential simulations in terms of their starting composition, and the mean and variance of males and female performance (summarized in Table 1). These characteristics set the pre-conditions each game satisfies to test the specific predictions of H2-H4 in terms of which sex males and females will vote out. Finally, the stated reasons for voting are analyzed to examine if the game as a whole can be characterized in terms of one of the nine overall predicted group dynamics.

Game 1

The initial sex composition was 3:5 male: female players. The first three players to be voted out were female followed by a male leading to a balanced male: female (2:2) composition for the fifth round. The two female players were voted out in rounds 5 and 6, leading to an all-male final. The male: female ratio across the 6 rounds of the game proceeded as such: 3:5, 3:4, 3:3, 3:2, 2:2, 2:1, and 2:0. On average, males and females answered the same number of questions correctly per round ($M = 2.83$ vs. 2.17 for males versus females, n.s.), and both had a low level of variance in their performance ($SD =$

1.47 for both). The lower levels of variance should have inhibited in-group derogation, and promoted in-group favoritism, and, indeed, leading to either the *old boys' club* or *battle of the sexes* pattern.

The verbatim responses of the players show that in rounds 5 and 6 the two male players, John and Cruz, strategically colluded to eliminate the two remaining female players (Aimee and Melody). The female players were unaware of this. They naively believed that they were eliminated due to their poor performance. Specifically, in the fifth (penultimate voting) round, John was retained although he was the weakest link in the round, while Melody, a female player was voted out. John wrote: "Cruz can win and split with me." Melody, the female player voted out in round 5, wrote: "Out of everyone left, I was the weakest link." Other female players observing the game believed that she was voted out because she was a threat. Given that the female players appeared less confident than the males players, the dynamic became the *old boys' club*.

In the final voting round, Aimee (female) and John and Cruz (both male) remained and Aimee was voted out. Her accuracy was comparable to John's in that round (66.67%) as well as cumulatively throughout the game (54.54% vs. 45.45%). John, the weakest link from round 4 who had been retained in the game, wrote that he voted for Aimee because her opponent, Cruz, "... knows lots of questions, he is most likely to win, but hey, I knew, if he win the money [sic] he might treat me to a candy." Aimee believed she was voted off for her performance. She wrote: "I was the weaker of the three left. John had voted for me before. Maybe Cruz knew this and voted for me, too." However, two of the four female players observing the final believed that Aimee was voted out because she was a threat.

The two male finalists showed strategic behavior but differed in the reasons they stated for voting. John's behavior appeared to be the most strategic, and he was aware of it. All six of his votes were for female players. He voted three times for Aimee. The verbatim advice he gave to others was: "First half of game, rid [sic] of weakest link. 2nd half of game, rid [sic] of strongest link." When asked why he won, he said, "Made sure [I] got rid of strong opponents towards the end."

The other male finalist, Cruz, also showed strategic behavior, but his stated reasons all reflect performance (e.g., "She got an easy wrong") and he attributed his and John's presence in the final to luck. Five of Cruz's six votes were for female players. In the fifth round, he voted for Aimee, who was the strongest link, and in the sixth round he voted for her again.

To examine whether Cruz underestimated Aimee's performance in fact, or merely stated that she had performed poorly, we examined Cruz's estimate of Aimee's versus John's performance. Cruz estimated that Aimee performed better than John in five of the rounds played and at the same level in one round (overall = 55% vs. 42.10% for Aimee versus John). Therefore, Cruz identified Aimee as the strongest performing player, even though he stated that she was a poor performer. Cruz lost in the final to John, despite being the strongest player overall (81.82% correct responses at the end of round six).

To summarize, the written reasons that players gave suggests that the two male finalists colluded to be in the final together. In such a situation, there is no need to retain a weaker player until the end of the game, as the two finalists expect to share the total winnings. There was no mean performance differential between males and females, and the variance in performance of each group was low. These conditions foster in-group

favoritism by male players (as they know each other), and no in-group derogation by female players (due to low variance in their performance). The strategy points to *old boy's club* (Figure 1) as an optimal strategy. The best game plan for the two colluders, both male, and not necessarily the top performers, is to eliminate other players, especially the two females in later rounds, and keep the final as an *old boys' club*.

Game 2.

Table 1 shows that in Game 2 women perform better overall (M questions answered correctly per round = 2.50 vs. 6.00 for males versus females, paired $t(5) = 10.25, p < .05$) but have a high level of variance within their performance level ($SD = 1.67$). This high variance among female players should lead to in-group derogation by females. If males were to demonstrate in-group favoritism then the pattern we would expect would be that of *male chauvinism*, but if they do not, one would expect to see the *queen bee* dynamic (Figure 1, H2). The *queen bee* dynamic is characterized by a stronger performing female voting out other females. The conditions fostering male players exhibiting in-group favoritism do not exist (e.g., higher male performance level), so a *queen bee* dynamic should manifest.

Game 2 had a 3:5 male: female start and culminated in an all-female final. The sex of the players voted out by round was: male, female, male, female, female, and male. The male: female ratios changed from 3:5, 2:5, 2:4, 1:4, 1:3, 1:2, to 0:2. In round 1 of the game, the weakest link, a male, was voted out, leaving only two males remaining. In round 2, the weakest link was a male player again. However, this male player was retained, while a female player was voted out, signaling the start of the *queen bee* dynamic. A female player mentioned her reason for voting in round 2 as: "In the show's

competition, vote off the strongest.” This player emerged as one of the finalists in the game, and continued to vote off other female players. For example, her reason for voting off a female in round 6 was: “She is the strongest in this round.”

In round 3, the weakest link from round 2 was voted out, and a female player who was the weakest link for the round was retained in his place, and voted out in the next round. A female player in round 4 said of the reason she was voting off another female: “Just ‘cause she’s strongest.” Therefore, in the first four rounds stronger female players began voting off other players they saw as a threat; but, overall, weaker players were eliminated either in the round in which they were the worst performers, or soon after.

Round 5 started with one male and three female players. In that round a female who was not the weakest link was voted out and a male player was retained instead, with the male player voted out in round 6. Both players voted out in rounds 5 and 6 believe that they were voted out for performance related reasons (“I did not bank soon, because I was trying to earn higher points”). Observers of the game, however, mentioned strategic reasons for their being eliminated. For example, of the elimination of the male player in round 6, two observers wrote: “He was the strongest player. The deciding vote was Leslie who wanted less competition,” and “Because they saw him as a threat and because he was the only guy left.” This is similar to the pattern we saw in Game 1 where players seem to be less aware of the strategic dynamics of the game and inaccurately attribute their elimination to their performance (Vancouver & Ilgen 1989).

The female winner gave performance-related reasons for voting off other players, and attributed her winning the game to her superior performance. Reasons for voting off other players included: “Missed an easy question,” “Missed all of her questions,” and

“Because she did not answer [the] first question correctly.” These reasons are also reflected in the advice she gave. The winner advised others: “Answer the questions right. Be sure of your answers. Remember to collect money. Pay attention to who gets questions right so you know when to bank. Make smart guesses.” The winner attributed her win to performance: “I answered most of my questions right (only missed two in the whole game) and tried to play smart about the money. I was trying to be confident.”

The other female finalist attributed her being in the final to luck. She advised: “If the pattern of questions gets easier, don’t bank as often. If the pattern of questions gets harder, bank often. In each round, you can stay in the game by luck.” She attributed her presence in the final round to luck: “I knew most of the answers. I’ve made some lucky guesses and banked money at appropriate times.”

Neither female finalist mentioned that they had strategically reduced the level of threat that other players represented by voting them off. The male player voted out in round 6, however, believed that there was more to explain regarding the two female finalists. Of the winner, he wrote: “She knew the answers to most questions and people were targeting who they thought was weak, not who they thought might challenge them later.” Of the losing finalist, he wrote: “She was low-key. She got her answers right and did not bring attention to herself.” Overall, he proffered the following advice: “Try not to bring attention to yourself. Don’t antagonize the others.”

Participants’ written strategy did not reflect their behavior. The winning female casted a vote for the strongest link in round 1 after inaccurately stating that the player had missed a question. Since that player had not been voted out in round 1, she voted for her again in round 3 despite her having banked the most money in that round and not being

the weakest link. This player made it to the final round against her and then lost. Five of her six votes were for females. Analogously, the other finalist (a female) also voted for female players four times in the six rounds.

To summarize, this game is characterized as a *queen bee* dynamic, in which a strong performing female player identifies competition in the early rounds of the game and starts to vote off players who represent a threat: particularly other females.

Game 3

Game 3 started off with six men and two women, and ended with a male-female final, and a male winner. This game also showed strategic content. The means in Table 1 show that the game is characterized by a skewed sex composition (6:2) that favors females being retained in the game (1 of the 2 females was a finalist but not a winner). Further, males correctly answered more questions per round than females ($M = 5.38$ vs. 1.38 for males versus females, $t = 3.17$, $p < .05$), with high variance within that performance (SD for males = 2.72), both conditions favoring males displaying in-group derogation. Under conditions where females demonstrate in-group favoritism, and males demonstrate in-group derogation, the resulting competitive dynamic should be *females as finalists but not winners* (Figure 1, H3). This was, in fact, the pattern observed.

While one female player was eliminated in round 2 after having incorrectly answered her question, the second female was retained until the final even though she was the weakest link in rounds 1, 4, and 6. Most players recognized this dynamic, including a player who was voted out in her stead. This player wrote: “Everybody feared his dominance. He was far superior to the rest of the table. David was sleazy. He did not want true competition and tried to feign he was voting for Michelle, changing his vote at

the last minute.” The female finalist also recognized why she was retained: “David thought that he could win more easily against me in the final round.”

David wrote of his reason for voting out the last player: “He is too smart.” David believes he won because: “I was able to answer questions well and played the strategy right by seeing who was likely to get voted off.” It is particularly telling that the reason he believed Michelle made it to the final round was: “She played the game well and was not threatening.” His key pieces of advice were: “Don’t make yourself stand out in the beginning. Try to establish yourself as smart. Figure out the voting pattern and play that. Get rid of competition later on in the game.” This player reported never having seen the television show, *The Weakest Link*.

In sum, the *females as finalists but not winners* dynamic represents a case in which as competition increases within the game, weaker-performing female players (difference in performance mean by sex) are kept until the end by stronger-performing male players (in-group derogation by males) as they are easy to beat in the final rounds. This male performance differential with high competition was the overall pattern in the *The Weakest Link* games analyzed by Valenzuela and Raghurir (2007) who first reported this special dynamic.

Game 4

This game started off with five men and three women and ended with an all-men final. The highest performing player won. The sex of the players voted out by rounds was: male, male, female, male, female, and female. The written comments of the players do not reveal strategic voting behavior. Male players showed more accuracy in their answers than female players (average number of questions answered correctly per round

= 5.33 vs. 2.83 for males versus females, paired $t(5) = 3.10, p < .05$). The key distinguishing feature differentiating Games 3 (6:2 male to female ratio) and 4 (5:3 male to female ratio), which were both characterized by males performing better than females, was the difference in the initial starting sex composition. Game 4 was less skewed towards males, reducing the pressure for females to display in-group favoritism. Further, the variance in the performance of males players was lower in Game 4 ($SD = 1.86$) versus Game 3, which reduced the incentive for male players to demonstrate in-group derogation. In Game 3, the conjunction of females demonstrating in-group favoritism, and males demonstrating in-group derogation led to the *females as finalists* pattern. Alternately in Game 4, with neither driver present, neither sex displayed an in-group bias, leading to *bounded rationality* (Figure 1).

In this pattern, voting behavior was predominantly performance-based. In the first two rounds that are characterized by low levels of competition and high levels of cognitive load, female weakest links were retained instead of male players who were voted out. In the last four rounds, weakest links were voted out. Except for one player who believed that gender considerations were at play, all reasons to vote out players reflected poor performance, citing overall inaccuracy (“She answered the most questions wrong”), missing an easy question (“She didn’t know the answer to Ireland, her home town”), time-wasting (“She took too long to answer her questions”), or not banking (“She lost the most points”).

The finalist who lost attributed the reason he made it to the final round to a combination of performance and luck. He advised others: “Answer questions quickly. Answer confidently. Try to vote with the majority.” The winner’s advice reflected a mix

of actual performance, minimizing the likelihood of errors being noticed, and cueing ability: “Answer correctly. Bank because people remember when you lose points. Don’t take a lot of time. Smile. Be happy.”

Players were aware that their earlier votes may have been based on inaccurate information integration. For example, a strong performing female who was the strongest link in rounds 2 and 4, but who was voted out in round 5, advised: “Vote consistently. Who you voted for on a hunch the first time will probably still be weak later.” Many players stressed the importance of appearing confident--a cue for assessing ability.

This game can be characterized as one with *bounded rationality* in which players and observers are aware of the difficulty of assessing performance and, therefore, use and recommend the use of heuristics such as speed, confidence, and demeanor.

Discussion

To summarize, four game simulations showed empirical support for three sex-based group dynamics with high strategic content: *queen bee*, *old boy’s club*, and *females as finalists but not as winners*, and one game showed support for the cognitive dynamic of *bounded rationality*. Table 1 summarizes the group characteristics for each of these games in terms of sex composition, and the mean and variance of the performance of men and women. Specifically, in Game 2 in which female players outperform male players, there is high variance in the performance of women, which is associated with in-group derogation, specifically, the queen bee dynamic. In contrast, in Game 3, where male players outperform female players and there is low variance among the performance of women, men retain weak-performing female players in order to reduce the threat that

other stronger players could pose in the final round. Thus, at the overall game level, the results of this experiential study support H2-H4.

The primary limitation of this study was that four game simulations that vary in terms of male-female performance differential, male-female performance variance, and male-female ratios do not allow for a rigorous test of the proposed model mapped out in Figure 1, and operationalized as H1-H4. To address this limitation, Study 2 examines the robustness of the dynamics found in the experiential study by examining the effects of group characteristics on voting behavior by using observational data from the actual television show.

Study 2: Observational Data: *The Weakest Link*

Measures

The observational study was based on 20 episodes of *The Weakest Link* (Summer 2001; Raghubir and Valenzuela, 2006, Valenzuela and Raghubir 2007). The data were coded at the individual level per round per episode, and aggregated up to round level, and episode level. At the round level, we coded *sex* of every player (male = 79; female = 81), *sex composition* (the percentage of females in every round of the game), and *competitive environment* (rounds 1 through 4 = lower, 5 & 6 = higher). At the individual level, we recorded *performance* (the number of correct answers given by and the total number of question asked of each player in each round). The performance level was also used to compute mean and standard deviations of the percentage of correct responses, separately for males and females by round. This was aggregated up to compute the overall performance of females and males in the episode. There were a total number of 729 observations across the 20 episodes. This data was used to compute the overall mean and

standard deviation of male and female performance levels in each of the 20 episodes, and in each of the six rounds of the 20 episodes ($n = 120$). Additionally, we coded *in- vs. out-group vote* (the sex of the person a player voted out in each voting round was categorized as: same sex or opposite sex). As each of the 20 episodes has 33 votes ($n = 8, 7, 6, 5, 4,$ and 3 votes over rounds 1 through 6), there were 660 observations.

Results

We examined whether male and female players voted to eliminate members of their own sex or of the opposite sex as a function of the performance levels of males and females per round (% correct per round), the standard deviation in the performance levels of males and females per round, the performance of the player casting the vote (% correct by that player in the round), the sex composition of the round (% females in the round), and the extent of competition in the round (rounds 1 through 4 = lower, 5 and 6 = higher). The models for both male players as well as for female players were significant ($R^2 = .17$ for males, and .24 for females, see Table 2).

Female players were more likely to vote out males the higher the performance of females ($B = 3.96, Wald = 19.56, p < .05$), the lower the variance in female performance levels ($B = -1.82, Wald = 2.90, p < .05$), the higher the variance in male performance levels ($B = 3.80, Wald = 11.76, p < .05$), the lower the percentage of females in the group ($B = -4.40, Wald = 13.50, p < .05$), and the higher the level of competition ($B = 1.32, Wald = 10.44, p < .05$).

Analogously, male players were more likely to vote out females the higher the performance of males ($B = 2.40, Wald = 6.95, p < .05$), the lower the performance of females ($B = -1.76, Wald = 6.77, p < .05$), the lower the variance in male performance

levels ($B = -3.26$, $Wald = 8.47$, $p < .05$), the higher the percentage of females in the group ($B = 2.05$, $Wald = 3.33$, $p < .05$), and the worse the individual player's performance ($B = -1.46$, $Wald = 10.04$, $p < .05$).

Overall, both sexes are more likely to retain a player from their own sex and vote to eliminate a player from the opposite sex the higher the mean and the lower the variance in their relative performance, and the lower their own representation in the group.

We next examined the likelihood of in-group favoritism or derogation only as a function of the relative performance differentials and variance of the performance of one's own sex. H2 predicted that a high variance in performance levels of females provides the conditions conducive to females eliminating other females to be *queen bees* while a lower variance promotes in-group favoritism among females who behave like *power girls*. The variance in female's performance levels should not affect male voting behavior. These dynamics are presented in the bottom row of Figure 1 where males do not show an in-group effect, and females show either in-group derogation or favoritism.

Further, we predicted that a high variance in performance levels of males would increase the likelihood that males vote out other males (*alpha male*), versus favoring them (*old boy's club*). This variance should not affect females voting behavior. These are the dynamics in the last column of Figure 1 (where females show no in-group effect) and males show either in-group favoritism or derogation. To examine these hypotheses at the episode level, we cross-tabulated whether the votes were for a player of the same sex or the opposite sex, with whether the variance of performance of females (males) was lower or higher, separately for female and male players.

Female votes for other females (same sex) or males (opposite sex), were contingent on the variance in their own performance both when females performed better than males, and when males performed much better than females. When females performed better, the higher the variance in their own performance, the higher the likelihood that they voted off another female player (Percent votes for other females = 15.8%, 32.4%, and 67.7% for low, medium and high variance categories respectively; $\chi^2(2) = 15.12, p < .001$). The low incidence of voting out another female when variance is low is consistent with the power girls dynamic where strong performing female players retain other strong performing females in the game. The high incidence of voting out females when the variance is high is consistent with the *queen bee* dynamic, where stronger performing females eliminate weaker performing females to maintain or increase the relative performance differential between the sexes.

When males performed much better than females, the same pattern existed, but was less extreme (Percent votes for other females = 21.2%, 33.3%, and 53.1% for low, medium and high variance categories respectively, $\chi^2(2) = 7.32, p < .05$). The low incidence of voting out other females when the variance in the performance is low or medium, despite conditions where males perform much better than females, is suggestive of strategic performance improvement behavior, where females collude to keep each other in the game, potentially increasing their performance and chances of winning.

To summarize, the lower the variance in female's performance, the higher the likelihood that females would retain other females in the group. However, there was no difference in males voting patterns for other males contingent on the variance in male's

performance levels. Males were, overall, just more likely to vote out a person of the opposite sex ($n = 198/328$ or 60%, $p < .01$).

Finally, Table 3 presents the combined effect of male/ female performance differentials and variance on the likelihood that males and females vote out the opposite sex, thus, testing H3-H4. For ease of analysis, relative performance differential is coded at only two levels: combining the two groups where males performed better, and the two groups when they did not (no difference and females better); as is variance in the performance of both males and females.

This 2x2x2 matrix (males better/ not x male variance: higher/ lower x female variance: higher/ lower) shows support for H3-H4:

When males perform better and the variance in their own performance is low (first column of Table 3), then males show in-group favoritism (votes for females = 66% and 70% for lower and higher variance in female's performance respectively, p 's $< .05$). This is the pattern of the old boy's club (last column, middle row of Figure 1), where males show in-group favoritism and females do not display an in-group bias.

When males do not perform better, and there is a high variance in male's performance, females display in-group favoritism voting out males 77% ($p < .05$) of the time when the variance in their own performance is lower, and 63% of the time when it is higher ($p < .10$). Males do not display an in-group bias. This is the dynamic referred to as *power girls* in Figure 1 (middle column, last row).

When males do not perform better than females overall, and are similar to each other (lower variance), both males and females vote out females when female performance has high variance (Votes for the opposite sex = 79% and 26% by males and

females respectively, $ps < .05$ for both). This pattern is captured by the dynamic of *male chauvinism* where females show in-group derogation (due to high variance in females performance levels), and males show in-group favoritism (due to the low variance in males performance levels), leading to males being favored overall (Figure 1).

Overall, the voting pattern presented in Table 3 suggests that males are more likely to retain other males when the variance in male performance is low (Columns 1 and 3, Table 3, supports H3 i), but are relatively unbiased when the variance in their performance is high (Columns 2 and 4, Table 3, no support for H3 ii). Female participants are most likely to show in-group favoritism when their performance is relatively better than that of males and the variance in the performance of males is high (last column, Table 3, supports H4). Females also display in-group derogation when their relative performance is better and the variance in that performance is also higher (Column 3, lower row, supports H4).

General Discussion

In this paper, we proposed that strategic sex-based group dynamics develop due to variations in group composition: specifically, differences in sex composition, and the mean and variance of male and female players' abilities. We proposed that given competition within the group, the lower the representation of the sex group, the higher its relative mean performance, and the lower the relative variance in the performance of one sex versus the opposite sex, the higher the likelihood of members of that sex displaying in-group favoritism (versus derogation). The combination of males and females displaying favoritism, derogation, or no in-group bias decides the overall pattern of

mixed-sex group dynamics. We delineated nine patterns of game play (Figure 1). Two studies using different methodologies examined evidence for these patterns.

In study 1, we used game simulations modeled on the television show *The Weakest Link*. The four games illustrated four different dynamics as a function of the starting sex composition of the players and the relative mean and variance differential between male and female players in the group. These patterns were further supported in Study 2 where we analyzed 20 episodes of the television show *The Weakest Link*.

Literature on gender effects has found support for the observation that men and women may be differently evaluated and rewarded in a group context (for a review, see Eagly & Karau, 2002). In the context of performance evaluation, Murray (1996) examined how people assess performance by showing study participants a tape containing race, social class, and gender cues of four children. He found that black men received the lowest ratings and white men received the highest ratings, with female students receiving ratings in the middle. Woehner & Roch (1996) found that men were given higher evaluations than women even after controlling for actual performance. Furthermore, gender differences were still present when performance was self-assessed. For example, Lippa & Beauvais (1983) developed an achievement setting based on an experimental quiz show. Male and female subjects participated in a computerized quiz game in which they could choose topic areas and question difficulty. Females estimated their performance lower than did males, chose lower difficulty levels, and tended to choose more “feminine” questions than did males. In sum, the literature shows evidence of a tendency to discriminate against women (rather than men) in achievement-related

tasks (Martell, 1996). Martell (1991) showed that this is due less to women being judged worse than they perform and more due to men being judged better than they perform.

The explanations for gender-biases in evaluation and reward appraisal discussed in the literature are mainly cognitive and motivational (Biernat, Manin, & Nelson, 1991; Fiske & Neuberg, 1990). Our findings support the presence of strategic behavior as competitive pressures increase in the game. Our contribution is developing a strategic explanation of gender biases based on a variety of strategic game play due to sex. First, we replicate findings from prior research (Lirgg, 1994; Valenzuela & Raghurir, 2007) showing that female players perform better in a non-male majority setting. Second, we find that both men and women use sex in a strategic way. They use sex either as a source of collusion, as a self-enhancement tool, or as a way to pinpoint the player who would maximize an individual's chances of winning the final round when it is a winner-take-all scenario. Third, we find that differences in sex composition and performance differentials seem to explain the way in which sex is used strategically.

Study Limitations and Areas for Future Research

One limitation of the studies is that by using real game situations we measured rather than manipulated the theoretical variables of sex composition, and mean and variance of male and female performance. This was in the interest of realism and would require replication to other scenarios that may be similar, such as competition among students in a group-project, among team members in an organization, and among team members in other domains such as sports, debates, and business.

A second limitation to the external generalizability of the results is the use of game shows to test model predictions. While they have the advantage of observing actual

human behavior under conditions where economic rewards are substantial, it is possible that they invoke other dynamics that may be more muted in other small groups, such as organizational teams, sports teams, and group discussions. A third limitation of the data in the two studies was the absence of a situation where females substantially outperform males. It would be interesting to examine the robustness of the results to other domains where females have a task advantage over men, a factor that has been shown to affect member involvement in the group itself (Vancouver & Ilgen, 1989; Karakowsky & McBey; 2001). Finally, a pertinent question relates to whether the results noted here relate to a participant's gender or their sex? We identified sex, but did not measure gender, which could have affected the strength of our results. This is suggested as an area for future research.

References

- Bettencourt, B., Miller, N., & Hume, D. L. (1999). Effects of numerical representation within cooperative settings: Examining the role of salience in in-group favoritism. *British Journal of Social Psychology, 38*, 265-287.
- Biernat, M., Manin, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology, 60*, 485-499.
- Bray, R. M., Kerr, N. L., & Atkin, R. S. (1978). Effects of group size, problem difficulty, and sex on group performance and member reactions. *Journal of Personality and Social Psychology, 36*, 1224-1240.
- Brophy, J. (1985). Interactions of male and female students with male and female teachers. In L. Wilkinson, & C. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 115-142). Orlando, FL: Academic Press.
- Brewer, M. B. (1979). Intergroup bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86*, 307-324.
- Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology, 46*, 991-1004.
- Doise, W., & Sinclair, A. (1973). The categorization process in intergroup relations. *European Review of Social Psychology, 3*, 145-153.
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice towards female leaders. *Psychological Review, 109*, 573-598.
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum model of impression formation from category-based to individuating processes: Influence of information and

- motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74. San Diego, CA: Academic Press.
- Griffin, P. S. (1983). Gymnastics is a girl's thing: Student participation and interaction patterns in a middle school gymnastics unit. In T. S. Templin, & J. Olson (Eds.), *Teaching physical education* (pp. 71-85). Champaign, IL: Human Kinetics.
- Heilman, M. E., & Kram, K. E. (1978). Self-derogating behavior in women – fixed or flexible: The effects of co-worker's sex. *Organizational Behavior and Human Performance*, 22, 497-507.
- Jackson, J. W. (1997). Effects of socio-structural and psychological factors on five types of intergroup bias expressed by minority group members toward majority and minority out-group members. Doctoral Dissertation, Purdue University, West Lafayette, Indiana.
- Jackson, J. W. (1999). How variations in social structure affect different types of intergroup bias and different dimensions of social identity in a multi-intergroup setting. *Group Processes & Intergroup Relations*, 2, 145-173.
- John, O. P., & Robins, R. W. (1994). Accuracy and bias in self-perception: Individual differences in self-enhancement and the role of narcissism. *Journal of Personality & Social Psychology*, 66, 206-219.
- Karakowsky, L. & McBey, K. (2001). Do my contributions matter? The influence of imputed expertise on member involvement and self-evaluation in the work group. *Group Organization Management*, 26, 70-92.

- Kerr, R., & Beh, H. C. (1995). Attributions of causality in different grades of Australian football league players. *Australian Psychologist, 30*, 102-108.
- Kruglanski, A. W., & Freund, T. (1983). The freezing and unfreezing of lay-inferences: Effects on impression primacy ethnic stereotyping, and numerical anchoring. *Journal of Experimental Social Psychology, 19*, 448-468.
- Levitt, S. D. (2004). Testing Theories of Discrimination: Evidence from the Weakest Link. *The Journal of Law and Economics, 47*(2), 431-453
- Lippa, R. A., & Beauvais, C. (1983). Gender jeopardy: The effects of gender, assessed femininity, and masculinity, and false success/failure feedback on performance in an experimental quiz game. *Journal of Personality and Social Psychology, 44*, 344-353.
- Lirgg, C. D. (1994). Environmental perceptions of students in same-sex and coeducational physical education classes. *Journal of Educational Psychology, 86*, 183-192.
- Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology, 23*, 77-87.
- Martell, R. F. (1991). Sex bias at work: The effects of attentional and memory demands on performance ratings of men and women. *Journal of Applied Social Psychology, 21*, 1939-1960.
- Martell, R. F. (1996). What mediates gender bias in work behavior ratings? *Sex Roles: A Journal of Research, 35*, 153-170.
- Murray, C. B. (1996). Estimating achievement performance: A confirmation bias. *Journal of Black Psychology, 22*(1), 67-85.

- Otten, S., Mummendy, A., & Blanz, M. (1996). Inter-group discrimination in positive and negative outcome allocations: Impact of stimulus valence, relative group status, and relative group size. *Personality and Social Psychology Bulletin*, *22*, 568-581.
- Pyszczynski, T., Greenberg, J., & LaPrelle, J. (1985). Social comparison after success and failure: Biased search for information consistent with a self-serving conclusion. *Journal of Experimental Social Psychology*, *21*, 195-211.
- Raghubir, P., & Valenzuela, A. (2006). Center of Inattention: Position Biases in Decision Making. *Organizational Behavior and Human Decision Processes*, *99*(1), 66-80.
- Seta, C. E., & Seta, J. J. (1992). Observers and participants in an inter-group setting. *Journal of Personality & Social Psychology*, *63*, 629-643.
- Shaw, M. E. (1976). *Group dynamics: The psychology of small group behavior*. New York: McGraw Hill.
- Sleeper, L. A., & Nigro, G. N. (1987). It's not who you are but who you are with: Self-confidence in achievement settings. *Sex Roles: A Journal of Research*, *16*, 57-69.
- Steiner, I. D. (1972). *Group process and productivity*. New York: Academic Press.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Asustin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-48). Monterey, CA: Brooks/Cole.
- Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*, 778-793.

Vancouver, J.B. & Ilgen, D.R (1989). Effects of Interpersonal Orientation and the Sex-Type of the Task on Choosing to Work Alone or in Groups. *Journal of Applied Psychology, 74*, 927-934.

Valenzuela, A. & Raghurir, P. (2007). The Role of Strategy in Mixed-Gender Group Interactions: A Study of the Television Show *The Weakest Link*. *Sex Roles: A Journal of Research, 57* (3&4), 293-303.

Wills, T. A. (1981). Downward comparison principles in social psychology. *Psychological Bulletin, 90*, 245-271.

Woehner, D. J., & Roch, S. G. (1996). Context effects in performance evaluations: The impact of rated sex and performance level on performance ratings and behavioral recall. *Organizational Behavior and Human Decision Processes, 66*, 31-41.

Table 1

Overall Game Characteristics of the Four Simulations

	Game 1	Game 2	Game 3	Game 4
Male: Female Starting Ratio	3:5	3:5	6:2	5:3
Finalist Dyad	Male-Male	Female-Female	Male-Female	Male-Male
Average number of questions answered correctly per round (<i>SD</i>)				
Male	2.83 (1.47)	2.50 (1.05)	5.38 (2.72)	5.33 (1.86)
Female	2.17 (1.47)	6.00 (1.67)	1.38 (1.19)	2.83 (1.17)
Paired t_7	.67	10.25*	3.19*	3.10*
Mean and variance of performance of Males and Females				
Male-Female Mean Difference	Low difference	Females > Males	Males > Females	Males > Females
Male-Female Variance difference	No difference	Females > Males	Males > Females	Males > Females
Sex-based group dynamic				
Sex In-group Dynamic	Males: In-group Favoritism	Females: In- group Derogation	Males: In-group Derogation	Females: In- group Favoritism
Game Dynamic	The Old Boy's Club	The Queen Bee	Females as Finalists	Bounded Rationality

* $p < .05$

Table 2

Results of Logistic Regressions for Study 2

Measure ¹	<u>Males</u>				<u>Females</u>			
	B	S.E.	Wald ³	Exp(B)	B	S.E.	Wald	Exp(B)
SD of female performance	1.98	1.07	3.42*	7.23	- 1.82	1.07	2.90*	.16
SD of male performance	- 3.26	1.12	8.47**	.04	3.80	1.11	11.76**	44.67
Female: % correct responses	- 1.76	.68	6.77**	.17	3.96	.90	19.56**	52.60
Male: % correct responses	2.40	.91	6.95**	11.01	-.46	.73	.41	.63
% females	2.05	1.12	3.33*	7.75	- 4.40	1.20	13.50**	.01
Level of competition ²	-.56	.37	2.30	.57	1.32	.41	10.44**	3.73
% of correct answers by the player voting	- 1.46	.46	10.04*	.23	- 1.81	.46	2.32	.50
Constant	.40	.91	.19	1.49	.27	.92	.09	1.31
-2 log likelihood	392.36				388.69			
Nagelkerke R ²	.17				.24			

1. Voting for the opposite sex (= 1). Level of analysis = round (6 per episode x 20 episodes)

2. Rounds 1 to 4 = 0, 5 and 6 = 1

3. ** $p < .05$, * $p < .10$

Table 3: Study 2 results

Female Performance: Lower Variance				
	Overall Males Better		Females Better or No Difference	
	Lower Variance in Males Column 1	Higher Variance in Males Column 2	Lower Variance in Males Column 3	Higher Variance in Males Column 4
Males	.66** (n = 56)	.59 (n = 37)	.53 (n = 49)	.54 (n = 28)
Females	.56 (n = 61)	.63 (n = 32)	.57 (n = 44)	.77** (n = 26)
Prediction	In-group Favoritism by males.	In-group Derogation by males.	In-group Favoritism by females.	In-group Favoritism by females.
	Supported			Supported
Female Performance: Higher Variance				
	Males Better		Females Better or No Difference	
	Lower Variance in Males Column 1	Higher Variance in Males Column 2	Lower Variance in Males Column 3	Higher Variance in Males Column 4
Males	.70** (n = 33)	.55 (n = 53)	.79** (n = 19)	.58 (n = 53)
Females	.51 (n = 39)	.58 (n = 52)	.26** (n = 27)	.63* (n=51)
Prediction	In-group Favoritism by males.	In-group Favoritism by males.	In-group Derogation by females.	In-group Favoritism by females.
	Supported		Supported	Supported

Figure 1

Mixed-Sex Group Dynamics		Females		
		In-group derogation (Females vote off females more than males)	In-group favoritism (Females vote off Males more than females)	No in-group bias
Males	In-group derogation (Males vote off males more than females)	Eat Your Own (Both sexes vote off members of their own sex)	Females as Finalists (Both sexes favor females)	Alpha Male (Males favor females)
	In-group favoritism (Males vote off females more than males)	Male Chauvinism (Both sexes favor males)	Battle of the Sexes (Both sexes vote off members of the opposite sex)	Old Boys Club (Males favor males)
	No in-group bias	Queen Bee (Females favor Males)	Power Girls (Females favor females)	Bounded Rationality

Key to Antecedent of the Dynamic

MOTIVATIONAL DYNAMICS

STRATEGIC DYNAMICS

COGNITIVE DYNAMIC

List of Figure Captions:

Figure 1: Conceptual Model of the patterns of Cognitive, Motivational, and Strategic Group Dynamics in Mixed-Sex Groups