The Role of Strategy in Mixed-Gender Group Interactions:

A Study of the Television Show "The Weakest Link"

Ana Valenzuela
Zicklin School of Business, Baruch College

and

Priya Raghubir
Haas School of Business, University of California at Berkeley

Authors' notes:

Authorship order is reverse alphabetical and reflects equal contribution. Please address all correspondence to Ana Valenzuela, Baruch College, 1 Bernard Baruch Way, New York, NY 10010, ana_valenzuela@baruch.cuny.edu

This research was supported in part by the Junior Faculty Research Grant awarded by the University of California at Berkeley to the second author.

We appreciate the assistance of NBC studios, which made available the taped episodes used to code the data in the observational study, the help of research assistants James Heyman and Claire Meissels, and the comments of Alice H. Eagly and Miguel Villas-Boas, as well as audience participants at the 2002 Society for Consumer Psychology meeting in Austin, Texas.

Abstract

This paper studies mixed-gender group interactions in a strategic game where group members are sequentially eliminated till a single winner takes all. Study 1 tests the hypothesis that female contestants are retained till final rounds where they are eliminated. Using observational data from the U.S. television show "*The Weakest Link*" (20 shows), results show that females are finalists but not winners. In a laboratory study (Study 2, 67 Berkeley undergraduates), we show that this effect is attenuated when winnings are shared among finalists (versus one winner takes all) due to the reduction in competitive pressures in the context.

Keywords: Game Strategy, Competitive Behavior, Group dynamics, and Gender Effects.

The Role of Strategy in Mixed-Gender Group Interactions:

A Study of the Television Show "The Weakest Link"

Prior literature on gender effects has delineated the many routes through which women may suffer a disadvantage in the workplace including differences in ability, life-style choices, motivation levels, biology and opportunity, to explain gender-related performance disparities. The current literature on gender effects takes a more nuanced view. Authors nearly unanimously debunks the "pipeline" theory, which argues that the supply of women is limited due to the life style choices they make, and the "deficit" theory, which argues that women do not have the characteristics required for upper management roles (for reviews see Eagly & Karau, 1991, 2002; Heilman, 2001). Current views examine the extent to which observers' cognitive and motivational characteristics can explain gender differences. Whereas cognitive antecedents relate primarily to the existence and use of gender stereotypes, motivational explanations relate to observers needs to maintain their gender stereotypes by processing information is a stereotype consistent manner.

This paper examines the role of strategy as an explanatory factor of gender-effects. We examine the context of a mixed-gender group interaction where group members are sequentially eliminated till a single winner takes all. We used the television show *The Weakest Link* as a metaphor for a group setting characterized simultaneously by both competition and cooperation, with a hierarchical structure over time where the winner "takes all." These kinds of group settings resemble those of an organization where entry ranks work together to achieve a common objective, but leadership has to be defended as managers climb the corporate ladder. *The Weakest Link* is a television game that consists

of an initial eight contestants who do not know each other who answer trivia questions cooperatively to achieve a maximum dollar prize amount with one winner who wins it all. The other seven win nothing, as they are consecutively voted out in rounds 1-6 of the 8-round game. The two finalists play the 7th and 8th rounds. The money earned in the 7th round is doubled and added to the accumulated winnings of rounds 1-6 to arrive at the overall pot of money. The player who answers the most questions correctly in the 8th round wins the game and the pot.

Contestants have to answer consecutive questions during a limited amount of time for each round. A string of eight correct answers allows for a maximum reward of $250,000 per round. If the string of correct answers is broken, the team returns to the beginning $1,000 level and loses the amount built up by the consecutive correct answers of previous players. At the end of rounds 1-6, each contestant individually and publicly votes for a player that s/he wishes to remove from the team. The player who receives the most votes is voted out with ties being resolved with the "statistically strongest player" from the previous round being given a tie-breaking vote. The statistically strongest (weakest) player is the player who answered the most (least) questions correctly. Accordingly, the number of players who play each round goes from 8, 7, 6, 5, 4, 3, to 2 in the seventh and eighth rounds, with a single winner after round 8.

Two studies examine the hypothesis that female contestants are retained till final rounds where they are eliminated; this is done by both male and female contestants, but for different reasons; and that this pattern is a result of strategic game play aimed at maximizing individual chances of winning. Study 1, using observational data from the U.S. television show "*The Weakest Link*" (n = 20 shows), show that females are finalists

but not winners. Study 2 uses a laboratory study (n = 67 Berkeley undergraduates) to show that this effect is reduced when winnings are shared among finalists (versus one winner takes all). The rest of the paper is organized as follows. We first summarize key findings related to gender effects, and then propose a strategic route for gender effects. The three studies are then described and results interpreted in the general discussion.

Gender Effects

Earlier approaches to gender effects focused on possible cognitive causes for differential gender-based evaluations. Such approaches held that gender cues are a simplifying heuristic, the use of which mitigates the need for an observer to undertake comprehensive processing (Fiske & Neuberg, 1990). Demonstrating cognitive biases in information processing, Martell (1991) showed that people were biased against women police officers when they were under time pressure or when they had other attentional demands.

Eagly and Karau's (2002) review of reasons for prejudice against female leaders suggests that prejudice manifests through a perception of incongruity between the descriptive stereotype of women's qualities (e.g., women are perceived to be more communal: concerned with the welfare of others) and the beliefs that a leader needs agentic qualities (e.g., assertive, controlling, and confident; see also Heilman 2001, 1983). It is this inconsistency between the perceived role of a leader and the perceived qualities of women that lead to women being perceived as less able and evaluated worse than men when those women are, in fact, leaders. Glick and Fiske (1996) reconciled the different views of women (nice people, but not necessarily competent in non-traditional

executive roles) in terms of their theory of "ambivalent sexism," which allows for both positive and negative views regarding women. Ambivalent sexists approve of women in traditional (e.g., homemakers) roles, but disapprove of them in non-traditional (e.g., managerial) roles.

Biernat and Fuegen (2001) proposed a different theory with a different prediction. This theory is based on the "shifting standards" model (Biernat, Manis, & Nelson, 1991). Biernat and Fuegen (2001) predict that women may be more likely than men of the same ability to be promoted at early and mid-level ranks, as the overall stereotype of women is that they are less able, leading to the conditional evaluation of a woman's ability, given her gender, being biased upwards: "good performance for a woman".

Motivational effects for gender differences have also been documented (e.g., Beyer, 1998; Buchner, Steffens, & Berry, 2000, Martell, 1991, 1996). For example, Martell (1996) demonstrated that stereotype consistent memory retrieval (i.e., recalling favorable and unfavorable information differentially for men and women) could not account for women's work being judged as less effective than men's when actual performance was the same and information about performance was unambiguous. However, she argued that this effect could be explained parsimoniously by a simple response bias that systematically appeared to favor men. In a similar vein, Buchner et al. (2000) found consistent evidence based on the paradigm of "false fame:" Study participants used a more liberal decision criterion to attribute fame to men as compared to women's names. This was less due to women being judged worse than their performance merited and more due to men being judged better than they performed. Gender discrimination implies that women will be less likely than men to reach the final round of a competition.

Strategic Antecedents of Gender Effects

Motivational and cognitive routes predict that performance and reward biases work in the same direction: if a specific gender is evaluated worse, it is also rewarded less. We now propose strategic game play as an alternate, though not substitute, route, through which the same empirical effects may occur. This route does not necessarily invoke conditional evaluations, and allows for differences in the direction of performance and reward. Strategic routes to gender-biases may be competitive and/ or cooperative.

Competitive dynamics. When there is competition between group members, then an individual's chances of being rewarded are based on her/ his own performance and the performance of others against whom she/ he is being evaluated. If individual group members use rewards strategically to improve their own performance or reduce the threat represented by the set of people against whom they are being evaluated, they will enhance their own chances of receiving future rewards.

Specifically, when competition is low, group members may cooperate with high performing others, but this may reverse as competition increases. This implies that during stages of low competition, evaluations are consistent with the reward structure, as those who are perceived to be good are retained in the group. However, when competition increases, then there is an increasing need for group members to retain those who are a lower threat (i.e., are worse than they are). It is at this stage that evaluations diverge from rewards. Group members who are evaluated worse, and, therefore, are less threatening to the rest of the group members, may be retained within a group simply because they are perceived to be easier to eliminate as competition increases. This implies that lower performers will be highly rewarded by others for strategic considerations.

In the presence of gender stereotypes where women are considered less able than men, this suggests that women are more likely to be retained so as to be easier to defeat when competition intensifies. This would lead to the hypothesized pattern of "Females as Finalists but not Winners."

Cooperative dynamics. Another route through which individuals can increase their relative performance level is by creating a context that allows them to perform better. Research on group performance based on group characteristics has shown that players own performance levels may be contingent on contextual variables including gender composition (Bray, Kerr &, Atkin, 1978). Even though the performance differential between all-male and all-female groups was non-significant (Bray et al., 1978), mixed gender groups seem to facilitate boys' achievements (Brophy, 1985) and limit girls' opportunities (Griffin, 1983). Lirgg (1994) found that girls perceive same-sex classes more favorably than mixed-sex classes, whereas the opposite is true for boys. Sleeper and Nigro (1987) found that, before beginning work on an experimental task, both male and female participants working with female partners expressed higher self-confidence than did those working with male partners. It is possible that people's self-confidence ratings reflected their beliefs about how competitive and aggressive a male or female partner would be (Deaux & Lewis, 1984). Similarly, Heilman and Kram (1978) reported that both men and women take more responsibility for success when they work with a woman than when they work with a man. This suggests that women (versus men) may be retained in a game so as to improve other players' own performance levels.

Gender of Participant as a moderator for strategic motivation. Overall, we predict that men and women will have different interaction styles within groups. Men tend to act

competitively within mixed-gender dyads in order to assert their dominance (Tannen, 1990) whereas women emphasize cooperative interactions, such as agreeing with or supporting the statement of others, that establish and maintain interpersonal interactions (Bales, 1970). From Bakan's (1996) perspective, this is a reflection of a feminine emphasis on communion, defined as a cooperative orientation that stresses interpersonal connections, contacts, and solidarity. As women's self-construals have been characterized as more inter-dependent, whereas men's self-construals have been characterized as more independent (Markus & Cross, 1990), and differential socialization directs boys toward competition and girls toward cooperation (Becker & Miles, 1978), we expect that the competitive dynamic would apply more to male players, whereas the cooperative dynamic would apply more to female players.

Summary of Predictions. We expect that the two strategic routes described above, competition and cooperation, would lead to the following behavioral patterns:

1) Women may be rewarded more highly than they are evaluated leading to their being retained in the group beyond levels warranted by their performance.
2) Men may be rewarded less than they are evaluated leading to their being eliminated from the group earlier than is warranted by their performance.
3) Women would be less likely than men to win a final prize than is warranted by the frequency in which they are included in the final round of competition.
4) Women's performance is better the more women there are in the group.
5) Gender effects attenuate when the context moves from competition to cooperation.

Study 1: Observation of "The Weakest Link"

Method

The observational study was based on 120 voting rounds (20 episodes x 6 rounds) of *The Weakest Link* (Summer 2001). The data were coded at the individual level (eight players per game for a total of 160 players); and vote level (each game has a descending number of eight to three votes over the six rounds of the game for a total of 660 votes), and aggregated up to the round level (20 episodes x 6 voting rounds, for 120 rounds).

At the overall episode level (n = 20 episodes) the two authors coded:

(i) Gender of every player (men = 79; women = 81)

(ii) Round at which each player was eliminated (which could take values 1-6, and 8 if the player won the round).

These two variables were used to compute the gender composition of the winners, the final dyad, and the final voting round to test predictions 1-3.

At the individual round level (6 voting rounds x 20 episodes), the authors coded:

(iii) The gender of the person voted out as the weakest link.

(iv) The gender of the actual weakest link (lowest performer).

These two variables were used to examine gender effects when the player voted out in the round was, in fact, not the weakest performing player in that round (67 of 120 rounds) to test predictions 1-3.

At the individual vote level (660 votes), the authors coded:

(v) The number of correct answers given by each player in each round.

(vi) The total number of questions asked of each player in each round.

(vii) The vote cast by each player to eliminate a group mate.

These data were used to examine the robustness of the effects shown at more aggregate (episode and round) levels, and to provide preliminary evidence for predictions 4 and 5.

Five episodes were coded by both authors. Given that there was no discrepancy in the coding between the two on any of these variables, the remaining episodes were coded by one of the two authors.

The analyses are conducted at the episode level, round level, and vote level.

Results

Aggregate Episode Level

We predicted that women would be more likely to be retained in the group (prediction 1) while men would be more likely to be eliminated (prediction 2), but women would be no more likely to win the game (prediction 3). Together, this suggests a pattern where one should see females as finalists but not winners. At the overall episode level, Predictions 1-3 were tested by examining the gender-composition of the winners, the finalist dyad, and the penultimate round of the game.

Final Round Composition. The probability of a gender-balanced male-female final round was 0.75, significantly higher than chance would suggest ($t_{19} = 1.9$, $p < .05$). The incidence of an all male (.10), or all female (.15) final is not significantly different from chance levels. Females were however no more likely to win the final (females won 40% of the 15 male-female finals).

Penultimate Round Composition. The penultimate round with three players (one of which would be eliminated, with the remaining two reaching the finals) was

significantly more likely to have 2 females and 1 male (70% of the games that started with 4 males and 4 females) rather than 2 males and 1 female (23%, $p < .05$). This gender composition of the penultimate and final rounds does not reflect actual performance. Only 39% of the players categorized as "weakest links" are male, while 61% are females (binomial $p<0.01$). On the other hand, 63% of the "strongest links" are males, while 37% are females (binomial $p<0.01$).

Results-Individual Round Level

At the individual round level, we examined predictions 1-3 by examining situations where the player voted out in the round was, in fact, not the weakest performing player in that round (67 of 120 rounds). Players categorized as the "weakest link" were only voted out less than half the time (53 of 120 rounds). The following analyses examine whether there are gender effects in retaining weakest links at the round level.

Gender of actual "weakest link" who was retained. In the rounds when players other than the weakest link were voted out, female "weakest links" were retained almost twice as often (64.18%) as compared to male "weakest links" (35.82%, binomial $p < .01$).

Gender of player voted out when the "weakest link" was retained. In the rounds where "weakest links" were retained, the player voted out in their place was more likely to be a male player: males were voted out in 39 of the 67 cases (58.21%), while females were voted out in the remaining 28 cases (binomial $p < .05$).

Together, these results show that retaining a "weakest link" is not gender neutral. Specifically, while in 20 (of 67) rounds, the retention of the weakest link was gender neutral (*i.e.*, a female weakest link was replaced by another female; a male weakest link

by another male) and was at chance levels, in the remaining 47 rounds a pro-female replacement occurred twice as often as a pro-male one: Female ”weakest links” were retained at the cost of a male player significantly more often (n = 31) than male weakest links were retained at the cost of a female player (n=16, binomial $p < .05$).

Together, the behavior of retaining female players in a manner not warranted by their performance led to a pro-female bias for final rounds. This pattern suggests that females appear to be over-represented in the final rounds of the game given their lower performance levels, leading to “females as finalists but not winners.”

Results-Individual Vote Level

The following analyses test predictions 1-3 at the individual vote level to test the robustness of the aggregate result, and examine whether there is preliminary evidence for men and women’s different strategic motivations (prediction 4), as well as whether gender effects are greater when the game context is more competitive (*i.e.*, in later rounds of the game), versus cooperative (*i.e.*, in earlier rounds of the game; prediction 5).

The first analysis examines whether there is a tendency for members of one gender to be voted out (or retained) more so than members of another gender, whether this effect is itself contingent on the gender of the player who votes, and the round of the game. The second analysis examines the same prediction by analyzing the likelihood of being voted out, controlling for the actual performance level of the player in that round, is affected by player’s gender and the round of the game. The final analysis examines whether performance level of male and female players is contingent on the gender composition of the group.

Effect of Round and Gender of person voting. We recoded each vote as for a person of the same gender (*e.g.*, a female player voting for a female player, or male player voting for a male player) or of the other gender (*e.g.*, a male player voting for a female player, or a female player voting for a male player), and examined whether this pattern varied as a function of the round of the game and the gender of the person voting. Across all rounds of the game, players of the other gender were voted out significantly more often (59%) than players of the same gender (41%), χ^2 (1) = 19.38, $p < .001$. This pattern could have been consistent with in-group favoritism if it had not been contingent on voting round and gender. Voting round (1-6) and gender of the player voting interacted with the gender of the player voted out (3-way interaction), χ^2 (5) = 13.31, $p < .05$. The pattern of voting is depicted in Figure 1.

-- Insert Figure 1 around here. --

An analysis of this interaction showed that, in the first and last voting rounds, there was no difference between men's and women's voting patterns for a person of the same or other gender, χ^2 (1) = .60 and .30 for round 1 and round 6 respectively, p's > .40. However, in rounds 2-4, men were more likely to vote out women (67%, 69%, and 66%, respectively; all binomial tests of difference in proportion, $p < .05$), whereas women split their votes equally between genders (binomial p's n.s.). Round 5 showed a marked departure from this pattern. Overall, men were voted out significantly more often than women in this round, χ^2 (1) = 5.75, $p < .05$. The two genders showed slightly different voting patterns: Female players were significantly more likely to vote out men than women (73%, binomial $p < .01$), whereas male players just significantly increased the number of votes directed toward other male players (47%) as compared to their voting

behavior in rounds 2-4 (33%, 31%, and 34%, respectively, binomials $p < .01$). To summarize, men voted out women at more than chance likelihood in rounds 2-4, but both men and women voted out men in round 5.

Effect of round and gender of voter controlling for Actual performance level of player. We now incorporate individual performance level as an explanatory variable to see if these results hold despite performance differences. To examine this issue, we conducted a logistic regression of whether a player was voted out or not; including as explanatory variables: the player's gender, performance level, and round of the game. This model had a higher explanatory value than base-line probability, $\chi^2 (6) = 106.23$, $p < .001$, and revealed a significant three-way interaction, $Wald = 6.59$, $p < .01$. To examine the pattern of this interaction, we conducted follow up analyses separately for male and female players.

Both models had a significantly better explanatory value than pure base-line probability, $\chi^2 (3) = 56.37$ and 52.51 for females and males, respectively, both p's $< .01$. Higher performance reduces the odds of being voted out for both men, $Wald = 24.04$, $p < .001$, and women, $Wald = 14.99$, $p < .001$, implying that these results have face validity. However, in the model run for male players, round of the game interacts with performance, $Wald = 9.67$, $p < .005$, suggesting that as the game progresses men are less likely to be retained despite their performance. There was no such effect for women.

To summarize, both men and women voted out players according to their performance during the early stages of the game, but this trend changed at the end of the game. In the penultimate voting round (round 5), men and women started voting out men more aggressively, even when their performance was relatively superior. The spike in

votes in round 5 against male players who performed at high levels suggests that there may be strategic game play involved when competitive pressures increase supporting prediction 5 which argued that gender effects would be exacerbated in competitive (versus cooperative) contexts. It is possible that stronger male players are voted out so that they do not become a threat at the end of the game.

Effect of gender-composition on actual performance. Prediction 4 proposed that one route through which strategic gender effects could manifest is if women perform better the greater the proportion of women in the group. To examine whether gender composition (ratio of male/ female players) affects women's and men's performance, we categorized each round of each game as "male-dominant" (n = 186) or "female-dominant" (n = 556). The number of questions answered correctly by each player in the round was subject to a 2 x 2 (gender of player by gender-mix) ANOVA. This analysis revealed a main effect of gender, F (1, 738) = 11.04, $p < .01$, Means = 2.03 vs. 1.78 for men versus women, respectively).

Follow-up analysis showed that for female players the effect of gender-composition was significant: their performance was worse when the group was male-dominant (M = 1.53 vs. 1.84 for male and female- dominant), F (1, 373) = 4.52, $p < .05$. The same ANOVA for male players did not show any significant effects ($F < 1$). Therefore, women, but not men, perform better in low-male-ratio contexts. Our findings support the cooperative strategic dynamic for gender-based voting behavior: Women may retain other female players to improve their own performance.

Discussion

To summarize, the results of the observational study of *The Weakest Link* suggest that gender effects occur for strategic reasons. We documented the effect "Female players as finalist but not winners" and showed evidence consistent with the strategic antecedents of gender effects: competitive and cooperative dynamics. These gender-based group dynamics predict that male group members retain lower performing female players until the end to enhance their individual chances of winning the final (the competitive dynamic) and that female group members retain female players in their group as their presence enhances their own performance (the cooperative dynamic). As a result, women are more likely to be retained in a group than is warranted by their performance, but they are no more likely to win the final game.

Study 2: Competitive versus Cooperative Contexts

A primary goal of Study 2 was to examine prediction 4 that argued that gender effects would reduce in cooperative scenarios. While Study 1 examined the effect of level of competition in the context by using round of the game as a proxy (with later rounds being more competitive), in Study 2 we change the utility of competing versus cooperating by manipulating the rules for winning at the end. We proposed that when a game has a cooperative end rule (finalists share), as compared to the competitive "winner keep all" rule, then it is less relevant to keep a player on till the final so as to be able to defeat him/her. If female players are retained as finalists primarily for the purpose of being easy to defeat, and if the goal of the game changes so that a finalist does not have to be defeated for another finalist to win, then women should be less likely to be selected to play the final round.

To examine this prediction, we asked male and female observers to simulate voting behavior for an episode of *The Weakest Link* and examined whether manipulating the end rules would affect which gender they would like to play the final round with, and why. The approach of asking people with whom they would like to play has been used in the past to study gender effects (Rudman, 1998). Rudman (1998) showed that, when people were asked to select a partner for a *Jeopardy!* game, women were more likely to choose a male partner, whereas men appeared to be equally likely to choose a male or female partner.

A secondary goal of this study was to demonstrate that women players are perceived as being lower performers than male players, an assumption that is central to the idea that women would be retained not because they are considered good, but because they are considered non-threatening. Given the main effect of gender in the performance measures in Study 1, it could be argued that considering female players as lower performing may have been based on objective reality rather than a subjective gender stereotype bias. Therefore, it is important to document this belief controlling for actual performance.

Method

Experimental participants were 67 undergraduate students in an introduction to marketing class at a west-coast university who undertook the study for partial course credit (Men = 33, Women = 34). The design was a 2 (end-game rule: cooperative/ competitive) between-subjects design. Participants observed selected segments of an episode of *The Weakest Link* during the experimental session. The specific game they

saw had an equal number of male and female players at the start, an all women final, and the "strongest link" was voted out in round 6 of the game.

After they watched the introductions of the players, participants made predictions about who would be the first two players voted out and which two players would be finalists. All participants then saw rounds 1 and 2 of the game; the videotape was stopped just before the votes were revealed in round 2. All respondents predicted who would be voted out, and then estimated the number of questions answered correctly and in total per respondent in that round. This was used to examine whether perceived performance of female players was lower than that of male players, controlling for actual performance differences. They were also asked to rate how much each player contributed to the group earnings on a 7-point scale anchored at "(-3) = Hindrance and "(3) = Help." This scale was used to examine whether there were gender effects in the use of a subjective scale (as suggested by Beirnat et al., 1991) and whether these effects were, in fact, different than gender effects observed using an objective scale (estimated performance).

After they made their estimates, participants were again asked for their predictions about who would be finalists in the game and who would be the winner. They were then shown round 5 of the game (four players remained: two female and two male players). They were then asked to estimate the actual performance of these four players and again to estimate who would be a finalist and who would win the game. Again, these measures were used to examine differences between actual and perceived performance as a function of players' gender.

At the end of the experimental session, we introduced the end-game rule manipulation that made the final round of the game either competitive (as in the original) or cooperative. In the competitive condition, they read:

> REMEMBER how the rules of the game WORK: Winner takes all and the loser gets nothing. Only one person can win. However, the group must work as a team to maximize the earnings from the game right up to round 6 of the game. In the 7th round of the game only 2 finalists play and overall earning are doubled. However, in round 8, only one of the two finalists gets it all. The loser goes home with nothing. Remember these rules, and imagine that YOU are playing the game and have perfect information about the other players' performance.

In the cooperative condition, participants read:

> IMAGINE the rules of the game are DIFFERENT: The two finalists are both winners. Each gets 50% of all earnings. The group must work as a team to maximize the earnings from the game right up to round 6 of the game. In the 7th round of the game the 2 winners play and overall earning are doubled. The same happens in round 8. After that round, total earnings are split equally between the 2 winners. Remember these rules, and imagine that YOU are playing the game and have perfect information about the other players' performance.

Participants designed their game strategy next. They chose which player they would vote out in each of the rounds, and described why. This was the primary dependent variable used to examine whether changing the competitive context would reduce gender effects in the choice of a finalist. Finally, all participants recorded their gender, whether they had ever watched the show, and were then debriefed. The procedure took 1 hour.

Results

We first examine the effect of competitive context on the viewers choice of who they would like as a finalist when they played the game to examine prediction 5, then analyze the reasons provided for why player were voted out for their strategic content as a function of the end-rule of the game, the round in which the voting was done, and the gender of the player voted out. Finally, we examine the perceived performance of different players as a function of their gender to assess whether female players are perceived to be lower performing than male players controlling for their actual performance levels.

Moderating effect of competitive context on choice of finalist. To examine prediction 5 that when the context moves from competition to cooperation strategic effects attenuate players chosen to be finalists are less likely to be lower performing female players, the choice of finalist was categorized according to their gender and cross-tabulated with the experimental manipulation of end game rule (cooperative/ competitive). As predicted, when the end-game rule was a competitive "winner-take-all" (as in the original game), then 20/34 (58.8%) of participants chose to play with a female finalist. On the other hand, when the end-game rule was cooperative, only 7/33 (or 21.2%) chose to play with a female finalist, overall χ^2 (1) = 9.85, $p < .002$. A similar pattern of results was obtained for both female and male experimental participants. Thus, as predicted, the preference for a female finalist was higher, despite overall lower performance, when the end-goal was a competitive "winner take all."

As all participants made their decisions given the actual performance of the players at various rounds of the game, this preference cannot be explained by recall or

evaluation biases in a stereotypic-consistent direction. It appears to be a strategic decision based on the individual's goal of winning or sharing the final pot of money. The next analysis examines this hypothesis directly using the reasons provided by viewers for why different players were voted out in the different rounds of the game.

Moderating effect of end-game rule on reasons to vote in the simulated choice. When participants were asked to be the 9th player and decide their voting strategy, as the rounds of the game proceed, people were increasingly likely to vote out an individual because they were a threat rather than because they were poor performers. In round 6, the penultimate voting round, 76% (26/ 34) of participants mentioned that they were voting out a player because s/he is a potential threat when the rules were competitive, compared to 30% (10/33) when the rules were cooperative, $\chi^2 = 15.29$, $p < .01$. The pattern continued in the final simulated voting round, round 7, $\chi^2 = 11.70$, $p < .01$; 94% versus 45% voted out a person because s/he was a threat in the competitive versus cooperative versions of the game. Thus, overall, players are more likely to vote out high performers toward the end of the game, particularly in the competitive version of the game.

Moderating effect of player's gender on the reason they were voted out. Male players were voted out 61.82% of the time in the penultimate voting round of the simulated version of the game. To examine whether male and female players were voted out for different reasons, we examined the reasons provided to vote out a player by the player's gender in the penultimate round of the simulated game (round 6). In the competitive version of the game, male players were predominantly voted out because they were a threat (95.2%), whereas female players were voted out because they were considered a threat only 66.7% of the time, overall $\chi^2 = 3.71$, $p < .05$. In the cooperative

version of the game, both genders were voted out significantly less often because of being a threat, overall $\chi^2 = 8.41$, $p < .01$ for men and $\chi^2 = 5.65$, $p < .05$ for women. Threat-based reasons for voting were reduced to 53.8% for male players and to as little as 25% for female players, $\chi^2 = 3.86$, $p < .05$. That is, men were voted out more often than women because they were a threat, especially in the competitive version of the game.

Moderating effect of round on reasons ascribed to a player being voted out. Reasons provided for players being voted out were coded into poor performance-related, threat-related, and other by an assistant who was blind to the hypotheses. The coding was checked by both authors, and no errors were found. In the first two rounds, respondents near unanimously (98.51%) named poor performance as the reason why the 1st (95.5%) or the 2nd player (77.6%) were voted out; only 13.40% mentioned that the player was perceived to be a threat, z = 9.92, $p < .05$. The percentage who mentioned that the player was a threat increased from 13.40% to 86.6% from the first two rounds to the last two rounds of the game, $z = 8.47$, $p < .05$, which is consistent with the idea that, as the game progresses, competitive pressures increase at the expense of cooperative pressures.

Across these analyses, results support prediction 5 that argued that gender effects exist for strategic reasons of competition between players with females being retained as they are perceived to be easier to defeat in the final round of the game, and these gender effects reduce when competitive pressures reduce. The next two analyses examine the underlying assumption on which this logic is based -- that women are perceived to be lower performers than men, controlling for actual performance levels.

Perceived performance of players in round 2. To examine whether gender affects perceived performance, we conducted a regression on the estimated accuracy of the

players in round 2 of the game as well as round 4 of the game. Predictor variables were the actual percentage of correct answers and respondent and player gender (male = 0, female = 1). The regression was significant, $R_a^2 = .160$, $F(4, 456) = 22.96$, $p < .001$. The coefficient associated with actual performance was positive and significant, $\beta = .352$, $t = 8.03$, $p < .001$, which suggest that the greater the actual accuracy of a player, the greater his/her estimated accuracy, which implies the face validity of the test. However, after we controlled for actual performance level, the performance of female players was estimated to be lower than that of male players as evinced by a negative beta coefficient for the player gender variable, $\beta = -.145$, $t = -3.30$, $p < .01$. Respondent's gender did not exert a significant effect, $\beta = .005$, $t = 0.11$, $p > .90$, which suggests that male and female observers were equally likely to estimate female players' performance as lower than male players' performance, even when actual performance was controlled.

Perceived performance of players in round 5. The pattern is similar for estimates of the four players' performance in the 5th round of the game, $R_a^2 = .463$, $F(3, 264) = 77.78$, $p < .001$, with a positive coefficient for actual performance, $\beta = .637$, $t = 13.52$, $p < .001$, a negative one for player's gender, $\beta = -.12$, $t = -2.55$, $p < .05$, and no effect of respondent's gender, $\beta = -.046$, $t = -1.02$, $p > .30$. This pattern of results confirms that people believe that females perform worse than men on this task.

Subjective performance evaluation. Results were the same using the subjective measure. A regression that incorporated actual percentage, respondent's gender, and player's gender was significant, $R_a^2 = .199$, $F(3, 451) = 38.60$, $p < .001$, with a positive coefficient for actual performance, $\beta = .353$, $t = 8.20$, $p < .001$, a negative coefficient

associated with player's gender, $\beta = -.215$, $t = -4.99$, $p < .01$, and no effect of respondent's gender, $\beta = .001$, $t = .01$, $p > .90$.

To summarize, when we controlled for actual performance levels, we found that both male and female observers estimated that women performed worse than men at the end of the second round of the game and at the end of the 5th round of the game.

Discussion

Study 2 results show that women are retained in the game because they are perceived to be less threatening than men. This strategic reason for retaining a woman so that she is easier to defeat in the final round can explain the pattern of why female players are finalist but not winners. When end-game rules are competitive, male players are considered a threat in the penultimate round of the game. In this setting, individual chances of winning become increasingly important as the game progresses (and competitive pressures overtake cooperative pressures), which leads players to want to play in the final round with a weaker player. Therefore, female players are chosen significantly more often than male players as a partner to play against in the final. However, when the end-game rules are cooperative, men are seen as less threatening, and so are less likely to be voted out before the final round.

General Discussion

We conducted two studies to examine the presence and reason for gender biases in mixed-gender group interactions. In an observational study we used the *Weakest Link* as a context that allows for a real-life investigation of the biases in people's voting

behavior in a group setting where there is tension between retaining strong players due to their ability to answer questions and earn money for the group and eliminating them due to the competition they pose at the individual level. The initial analysis of the game's final-dyad composition showed that more female players make it to the final rounds than should have if only their performance were taken into account. A test of observational data from the actual game show supports the pattern of "Female players as finalists but not winners," where male group members retain non-threatening female players until the end to enhance individual chances of winning and female group members retain female players in their group as those women's presence enhances their own performance or, at least, creates a "nice" environment in which their performance improves. As a result, female players are more likely to be retained in a group than is warranted by their performance, but they are no more likely to win the final game. We propose that this occurs because of *strategic considerations*, whereby the use of gender in decision-making can be explained in terms of rational utility-maximization.

We proposed that gender effects may occur for strategic reasons over and above the motivational or cognitive ones as studied in the literature. In Study 2 we manipulated the competitive versus cooperative context of the game and showed that the tendency to retain lower performing female players until the end when they were easy to defeat was reduced when the end-game was not a "winner takes all" scenario.

Theoretical Contributions

This article makes a contribution to the study of gender effects. Literature on gender effects has shown that men and women may be differently assessed. In the

context of performance evaluation, Murray (1996) examined how people assess performance by showing a tape that contained race, social class, and gender cues of four children (two Black and two White). He found that Black boys received the lowest ratings, White boys received the highest ratings, and Black and White girl students received ratings in the middle. Woehr and Roch (1996) also found that men were given higher evaluations than women even after controlling for differences in actual performance. Gender effects have also been demonstrated in quiz shows. Lippa and Beauvais (1983) developed an achievement setting based on a quiz show. Male and female participants played in a computerized quiz game in which they could choose topic areas and question difficulty. Results showed that women estimated their performance lower than did men, chose lower difficulty levels, and tended to choose more feminine questions than did men. Such effects are explainable by differences in self-confidence (Lenney, 1977). The finding that men are more confident in their abilities has been a fairly robust and documented in stereotypically masculine domains such as math ability (Mura, 1987), although there do not appear to be gender differences in confidence in stereotypically feminine areas such as interpersonal perceptiveness (Lenney, 1981). In addition, individuals who approach tasks with a low expectancy of success perform worse (Feather, 1966), even when expectancies are experimentally induced (Diggory, 1966). Such a theoretical account is consistent with our results: If performance is contingent on gender-mix (Brophy, 1985), and it affects evaluations of the experience (Lirgg, 1994), it is a viable strategic reason to vote out a player on the basis of gender.

Although earlier researchers have examined the cognitive and motivational antecedents for gender-biases in performance appraisal, we propose a hitherto new route

through which gender influences behavior: *strategy*. We argue that the need to maximize individual chances of winning is a rational strategic reason that can lead to gender effects: the desire to compete against a player perceived to be easier to defeat. Players appear to retain (what they perceive as) lower performing female players not to achieve the cooperative goal of increasing the size of the pot, but because, at the end of the game, it increases their individual likelihood of winning the pot as winners are decided by better relative performance. This increases the competitive motivation to eliminate strong links early in the game and to retain a weaker player to maximize one's own chances of winning the game. Playing with a lower performing player at the end maximizes their own chances of winning the pot of money at the time the dynamics of the game become increasingly competitive.

We contrast our findings to those of Biernat and her colleagues' "shifting standards" model, yet acknowledge that the strategic route we propose is not a substitute but may complement the fact that women are rated against a lower standard (Biernat & Feugen, 2001; Biernat & Kobrynowicz, 1997; Biernat et al., 1991). However, our findings move away from Biernat's model in two different ways. First, we identified biases in assessments of women's performance using both objective and subjective measurements. Second, we found that women are perceived to perform worse than they actually do, whereas Biernat's model makes the opposite prediction based on stereotypes.

On the other hand, we did not find clear evidence for in-group biases (a higher than chance likelihood of voting off players of the other gender). The literature on inter-group relations has indicated 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 players on *The Weakest Link* show in-group biases, the gender of the person voted out should be contingent on the gender of the person casting a vote: women should be more likely to retain women, and men more likely to retain men. However, in our studies, we found that male and female observers and players behaved similarly. This could reflect the changing face of society (gender identity becoming less salient), a situation (strong rules and aims to reach) that does not afford in-group bias or mere cross-categorization.

Areas of Future Research

We speculate that these different gender dynamics may develop due to differences in three basic factors. First, they may depend on the initial starting gender composition. When there is gender balance in the starting ratio of male and female players, then the need to maintain equity leads to a pattern of gender-balancing whereby people of different genders are sequentially and alternatively asked to leave a group. That is, if a man leaves in the first round of retrenchment, a woman leaves in the second round, and vice versa. On the other hand, when there is an unequal starting composition, then members of the minority group collude to increase their strength against the majority group. For example, if the ratio of men to women is skewed against women, then the women will band together to try and ensure a female representation in the group.

Second, the different gender dynamics may depend on the mean and variance of players' abilities by gender. When the two genders perform at different levels, the higher performing group can use someone from the lower performing group as a token member, that is, a person who can be retained when necessary only to be discarded later. When

there is performance variance within a group (especially a lower performing gender group), then the higher performing members within the group could retaliate against the lower performing members to preserve the gender-group's superiority. In addition, people's own performance affects the extent to which they use data-based cues to assess another person's performance. For example, Pyszczynski, Greenberg, and LaPrelle (1985) explored biases in information search that lead to self-serving attributions for success and failure and help maintain positive self-evaluations. They showed that information search about performance outcomes is biased so as to provide evidence consistent with a favorable self-evaluation. Participants who had not performed well asked for more information 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 good performance of others.

Third, the development of gender dynamics may vary with the extent of competitiveness versus cooperation in the group. When there is cooperation we expect to see more collusive patterns within a gender group, whereas when there is competition, we expect to see members of one group keeping weak members irrespective of their gender, as they are perceived as less threatening later. To the extent that one particular group is weaker than the other (in perception or in reality) the members of that group may be retained as token participants. Future researchers should follow up to identify evidence for the remaining dynamics and explore the role of the three antecedent factors proposed in determining strategic group gender dynamics.

References

Bakan, D. (1966). *The duality of human existence: An essay on psychology and religion.* Chicago: Ran McNally.

Bales, R. F. (1970). *Personality and interpersonal behavior*, New York: Holt, Rinehart and Winston.

Becker, W. M. & Miles, C. (1978). Interpersonal competition and cooperation as a function of sex of subject and sex of counterpart, *Journal of Social Psychology*, *104*, 303-304.

Beyer, S. (1998). Gender differences in self-perception and negative recall biases. *Sex Roles: A Journal of Research*, *38*, 103-134.

Biernat, M. & Feugen, K. (2001). Shifting standards and the evaluation of competence: Complexity in gender-based judgment and decision making. *Journal of Social Issues*, *57*, 707-724.

Biernat, M. & Kobrynowicz, D. (1997). Gender and race-based standards of competence: Lower minimum standards but higher ability standards for devalued groups. *Journal of Personality and Social Psychology*, *72*, 544-557.

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.

Buchner, A., Steffens, M. C., & Berry, D. C. (2000). Gender stereotyping and decision processes: Extending and reversing the gender bias in Fame Judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1215-1227.

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.

Diggory. J. (1966), *Self-evaluation: Concepts and studies*, New York: Wiley.

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. (1991). Gender and emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, *60*, 685-710.

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

Feather, N. T. (1966). Effects of prior success and failure on expectation of success and subsequent performance. *Journal of Personality and Social Psychology, 3*, 287-298.

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, 23*, 1-74, San Diego, CA: Academic Press.

Glick, P. & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, *70*, 491-512.

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. (1983). Sex bias in work settings: The lack of fit model. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior, 5*, Greenwich, CT: JAI.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues, 57*, 657-674.

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.

Lenney, E. (1981). What's fine for the gander isn't always good for the goose: Sex differences in self-confidence as a function of ability area and comparison with others. *Sex Roles, 7*, 905-923.

Lenney, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin, 84*, 1-13.

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.

Markus, H., & Cross, S. (1990). The interpersonal self. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp.576-608). Guilford Press: New York, NY.

Martell, R. F. (1996). What mediates gender bias in work behavior ratings?" *Sex Roles: A Journal of Research, 35*, 153-170.

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.

Mura, R. (1987). Sex-related differences in expectation of success in undergraduate mathematics. *Journal of Research in Mathematics Education, 18*, 15-24.

Murray, C. B. (1996). Estimating achievement performance: A confirmation bias. *Journal of Black Psychology, 22*, 67-85.

Pyszczynski, T., Greenberg, J. & LaPrelle, J. (1985). Social comparison after success and failure: Biased search for information consistent with a self-servicing conclusion. *Journal of Experimental Social Psychology, 21*, 195-211.

Rudman L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counter-stereotypical impression management. *Journal of Personality and Social Psychology, 74*, 629-645.

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, 16*, 57-69.

Tannen, D. (1990). *You just don't understand: Women and men in conversation.* New York: Ballantine Books.

Figure 1. Likelihood of Voting for the Opposite Gender