



Undergraduate Honors Theses

2021-03-09

Group Gender Composition: A Field-Experiment Evaluating Individual Levels of Group Satisfaction and Perceived Group Effectiveness

Jacob Davis

Follow this and additional works at: https://scholarsarchive.byu.edu/studentpub_uht



Part of the [Economics Commons](#)

BYU ScholarsArchive Citation

Davis, Jacob, "Group Gender Composition: A Field-Experiment Evaluating Individual Levels of Group Satisfaction and Perceived Group Effectiveness" (2021). *Undergraduate Honors Theses*. 168.
https://scholarsarchive.byu.edu/studentpub_uht/168

This Honors Thesis is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Undergraduate Honors Theses by an authorized administrator of BYU ScholarsArchive. For more information, please contact ellen_amatangelo@byu.edu.

Honors Thesis

GROUP GENDER COMPOSITION: A FIELD-EXPERIMENT EVALUATING
INDIVIDUAL LEVELS OF GROUP SATISFACTION AND PERCEIVED GROUP
EFFECTIVENESS

by
Jacob Davis

Submitted to Brigham Young University in partial fulfillment
of graduation requirements for University Honors

Economics Department
Brigham Young University
April 2021

Advisor: Dr. Olga Stoddard

Honors Coordinator: Dr. John Stovall

ABSTRACT

GROUP GENDER COMPOSITION: A FIELD-EXPERIMENT EVALUATING INDIVIDUAL LEVELS OF GROUP SATISFACTION AND PERCEIVED GROUP EFFECTIVENESS

Jacob Davis

Economics Department

Bachelor of Science

Group collaboration is commonly used in educational and professional settings. This paper will identify whether female-majority groups are more likely to see positive group dynamics than female-minority groups. Group dynamics in this project will be measured as individual levels of group satisfaction and group effectiveness. I use data from a field experiment where university students who took a weekly study group class had their experience evaluated with monthly surveys. My results find no statistically significant relationship between female-majority group conditions and perceptions of group satisfaction or group effectiveness.

ACKNOWLEDGEMENTS

I express my appreciation to my thesis committee who gave valuable direction and feedback: Dr. Olga Stoddard, Dr. Eric Eide, and Dr. John Stovall. I also thank Dr. Chris Karpowitz, Dr. Jessica Preece, and Dr. Stephen O'Connell for providing an opportunity to assist with their research.

TABLE OF CONTENTS

Title.....	i
Abstract.....	iii
Acknowledgements.....	v
Table of Contents.....	vii
I. Introduction.....	1
II. Methodology.....	3
III. Data.....	6
IV. Results.....	9
V. Conclusion.....	15
VI. References.....	16

I INTRODUCTION

While group collaboration is commonly used in educational and professional settings, economic research has given little attention to systematically understanding group satisfaction and effectiveness across different gender compositions. Prior work has investigated gender composition effects in laboratory microfinance experiments with all female, all male, and mixed groups (Berge et al., 2016). Other work has evaluated three-person groups playing a large business game (Apesteguia et al., 2012) or a group dictator game (Defwenberg and Muren, 2006). Where these studies find that women-dominated groups have relatively worse outcomes, others find the number of women per group to positively relate to performance (Fenwick and Neal, 2001). Understanding gender composition effects in group work is increasingly important as other research has shown that teams produce more knowledge than individuals (Wuchty et al., 2007) and raise output in more complex production processes (Boning et al., 2007).

The effects of group dynamics persist into the workforce. A recent paper examining study groups found that females are more likely to drop out of male-dominated fields, exacerbating the minority status of women, potentially because they are less integrated to the social environment (Shan 2021). Other studies also show that women may shy away from competitive tasks or male-dominated fields due to lower self-confidence (Bengtsson et al., 2005; Kamas & Preston, 2012; Niederle & Vesterlund, 2007; Reuben et al., 2017). There is also a potential “stereotype threat” effect where individuals perform worse or expect less when their group faces negative stereotypes (Cadinu et al., 2003; Nguyen & Ryan, 2008; Spencer et al., 2016).

Recent studies that examine group dynamics find specific gender disparities. Coffman (2014) finds that women in stereotypically male circles are less confident in their own ability and therefore less likely to contribute ideas to their group. Stoddard et al. (2020), examining majority and minority female study groups, show that women in female-minority groups are perceived as less influential and less authoritative by peers. That same study also finds that males do not encounter the same issue when outnumbered by females.

The asymmetrical behavior differences between male-minority and female-minority groups potentially suggests a Pareto-efficient increase in utility when groups are arranged to avoid placing females in male-dominated groups. If women enjoy more positive group dynamics from being in the majority, their level of satisfaction with their group may be higher than if they had been in the minority. A higher degree of satisfaction may then produce a more effective group. Exogenous variation in group gender conditions would allow for a causal estimate of group condition on group satisfaction and effectiveness.

This paper seeks to add to the literature the impact of group gender composition on individual's level of group satisfaction and group effectiveness. I identify whether female-majority groups are more likely to see positive group dynamics than female-minority groups. Group dynamics in this project will be measured as individual levels of group satisfaction and group effectiveness. I use the data generated through a field experiment which randomized students into study groups with various gender conditions. Participants in the field experiment took a weekly study group class and had their experience evaluated with monthly surveys. My results find no statistically significant

relationship between female-majority group conditions and perceptions of group satisfaction or group effectiveness. In the following sections, I explain the experiment design and methodology, describe the data, then explore regression results before concluding.

II METHODOLOGY

This section first describes the setup of the experiment then describes the variables of interest, individual self-reported measures of group satisfaction and group effectiveness.

Experiment Design

I assisted in conducting a randomized control trial (RCT) that examined student interaction in a study group class at a Western private university. Students self-selected into the study group class to supplement a general education (GE) course in which they were concurrently enrolled. The treatment condition placed more females than males in small study groups of six students. Control groups had a male majority among the six students. Gender conditions were assigned as 2F, 4F, or 6F, where the number corresponds to the number of female students placed in the group. Students were assigned to group conditions and groups randomly.

The female-majority treatment contrasts against observed real-world outcomes where, in certain professional spheres, women often find themselves outnumbered by men. My hypothesis is that groups with female majorities will report greater group satisfaction and effectiveness. The treatment of majority-female groups is in essence a

decrease in sexism relative to male-dominant groups. Female-minority groups are considered the control group. I will also designate the all-female group condition as a separate treatment condition from the female-majority treatment. Both treatments will be compared against the female-minority control condition.

Throughout the semester, students in the study group class completed assignments that supplemented a GE course which the students were also enrolled in. As a graded class, students had a natural incentive to participate in the group work. Along with the group work that pertained to the GE course, students completed several task-based labs. The students met in person weekly for 50 minutes. They also wore microphones so that the study groups could be recorded and evaluated.

Students completed monthly surveys regarding their experience in their groups. Among the questions were inquiries asking the student to evaluate the effectiveness and quality of their group. Students' perception of their group experience composes the primary data in this project. The monthly surveys allow for group dynamics to be analyzed over time. There were five monthly surveys. The first survey will not be used in this paper's analysis as students completed it prior to their first group meeting. The analysis will focus on the other four surveys which may give insight into how an individual's group satisfaction and effectiveness change over time and by group gender composition.

Group Satisfaction and Effectiveness Indexes

Students completed monthly surveys that asked specific questions about how satisfied the student was with their group. Students had the option to mark statements regarding group satisfaction as "Extremely satisfied," "Somewhat satisfied," "Neither

satisfied nor dissatisfied,” “Somewhat dissatisfied,” or “Extremely dissatisfied.” The statements asked the students their satisfaction about “how well the members of your study group get along with each other,” “what you are learning about [GE class] in your study group sessions,” “the study group assignments,” and “your study group overall.” I combined these satisfaction questions and combined them into an index from 0 to 1. A score of 0 would imply the lowest possible amount of satisfaction while 1 is the highest possible satisfaction. Each survey answer was coded as 1, 2, 3, 4, or 5, corresponding to higher amounts of affirmation with the survey question. I summed the highest possible positive responses, then subtracted off the lowest possible score (the most negative possible set of responses) to baseline off zero. Then, for each observation in the dataset, I added together the participant’s responses, subtracted the lowest possible score, and divided by the difference of the highest and lowest possible scores. Doing this created the index where 0 is the most negative possible score and 1 is the most positive possible score.

Student responses to the satisfaction statements were relatively correlated with an average correlation coefficient of 0.473 across each survey. The weakest correlation, 0.26, was between student satisfaction regarding what students were learning and the study group assignments. The highest correlation, 0.66 across each survey, was between student satisfaction with their study group overall and what the students were learning. Examining the correlations between these questions shows that satisfaction may have been primarily driven by students’ perception of how much they were learning during the study groups. Whether group dynamics impacted how much students were learning will have large implications for student satisfaction.

The monthly surveys also asked the students about their perception of group effectiveness. Because each question measures a different aspect of effectiveness, creating an index with these questions creates a generalized estimate. Like the questions about group satisfaction, responses to the group effectiveness statements had potential responses of “Strongly agree,” “Agree,” “Neither agree nor disagree,” “Disagree,” or “Strongly disagree.” The students were asked to respond to statements about whether group discussion helped the students better understand the issues, group work made everything slower and harder to accomplish, a few people tended to dominate discussions, members treated each other with respect, members were too quick to agree with each other, study sessions were less helpful than expected, study sessions brought new perspectives, and too many opinions hampered productive study sessions. Following a similar method for the satisfaction index, I condensed responses for effectiveness questions into an index between 0 and 1. A score of 0 would imply the most negative possible responses while a score of 1 is the most positive possible.

The responses to survey questions are not a perfect capture of group satisfaction and group effectiveness. Even still, the indexes should be correlated with the students’ perception of their groups. The sign, positive or negative, on a statistically significant treatment coefficient would provide an acceptable answer to the hypothesis that individuals in female-majority and all-female group conditions enjoy a higher level of group satisfaction than those randomly assigned to be in the female-minority groups.

III DATA

The project data includes student demographic information and student survey responses. The partnering university provided student demographic data. The surveys were conducted monthly through Qualtrics, and the demographic data and survey data were merged in Stata. I created indexes based off a collection of questions about group satisfaction and group effectiveness, paying careful attention to whether affirmative responses were positive or negative. The data is reliable and relevant for this project because the data originates from a real-world experience with motivated students who self-selected into a study group class. However, the participants in the study are primarily young adult age, white, and attending college. The external validity of this project's data may be limited to similar demographics in a similar situation. Despite limitations, the data provides a snapshot of group dynamics from legitimate group work.

Table 1: Randomization Balance Table By Gender and Condition

	Min F, Female	Min F, Male	Maj F, Female	Maj F, Male	All F, Female
Age	19.1	20.2	19.6	20.3	18.9
White	92.3	94.8	94	92.6	94.4
Married	1.5	5.2	2	2.9	2.1
ACT score	28.1	28.7	27.3	28.4	28
Class year status	1.3	1.2	1.3	1.4	1.2
Leadership experience	84.6	89.6	79.3	89.7	85.5
International student	4.6	3.7	6	7.4	7.2
Parental income category	4.1*	4.2	4	4.2	4.4
Political affiliation scale	2.9	2.6	3	2.3	3.1
Ambivalent sexism index	41.8	56.7	41.3	56.7	38.9
Egalitarian index	65.1	56	62.3	54.9	62.5
Individualism index	61.1	68.8	63	69.8	64.4
Pro-sociality index	72.8	72.1	73.6	69.5	73.6
Conflict avoidance index	46.2	42.1**	45.6	47.1**	44.8
Individual efficacy index	53.3	59.9	53	61.5	56.1
Group efficacy index	58.9	63.4	59.2	61.6	58.6
N	65	135	150	68	195

Note: I found statistically significant differences at the 5% level for two variables. The within-gender difference in parental income among females in Minority-Female groups goes away when not using dummy variables for the income categories. The within-gender difference for the conflict avoidance index between males across groups is significant at the 1% level. * $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Table 1 shows descriptive statistics for the relevant data, organized by group condition and gender. There are more females in general in the study group class. Because of the female-to-male ratio, the all-female group condition was created which is considered a treatment separate from the majority-female treatment. The variety of variables and indexes show that randomization was successful. Across a broad range of

categories, students of the same gender look the same between the different group conditions. The difference in parental income for females in minority-female groups goes away when using dummies for the income categories. The only non-gender difference across groups is the conflict avoidance index scores among males. This result could be due to chance from testing so many variables. If males in minority-female groups do have a lower conflict avoidance score, it should pronounce the detriment women see when placed in the minority as the males in those groups are less likely to avoid conflict.

Table 1 also shows balance test between gender. Among the variation between males and females in the study, age is a statistically significant difference. The average female in the study is a little over a year younger than the average male. Regressing age on gender results in a coefficient of -1.03. This result is nuanced by the measure of age in whole years. While there is an overall age difference between males and females, female ages are not statistically different across group conditions.

A key aspect of the treatment is the difference between males and females, especially as measured by an ambivalent sexism index. Another key distinction is highlighted through student responses to a battery of questions that provide a measure of hostile and benevolent sexism from the baseline survey. The measure comes from an index of the sexism related questions where 0 shows the least sexism while 1 corresponds to the highest sexism score. Table 1 shows a higher average sexism score among males than females across each group condition. Regressing the sexism score on a female indicator produces a statistically significant negative coefficient. A difference between all female, majority female, and minority female groups is a lower, on average, sexism score

as more females and fewer males are present. This difference could account for higher or lower amounts of an individual's level of group satisfaction and effectiveness.

IV RESULTS

Across Time

Do individuals in a female-majority group report higher satisfaction? The following tables explore the regressions from the project data. The first analysis regressed the satisfaction index on majority-female treatment status for each of the four monthly surveys. The same regression was repeated with the all-female treatment. The coefficients are summed in the table below, organized by survey. The falling number of respondents is due to a small amount of attrition when students withdrew from the class. The specific attrition rates for minority, majority, and all female groups were 3.9%, 2.7%, and 2.6%, respectively. The treatment statuses do not have a statistically significant difference in attrition rates. When also factoring in demographics such as gender, the groups maintain equal attrition rates; women in minority-female groups are not more likely to leave the study than other students.

Table 2: Satisfaction Over Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Survey 1		Survey 2		Survey 3		Survey 4	
Maj F	-0.004 (-0.22)		-0.007 (-0.33)		-0.024 (-1.27)		-0.018 (-0.83)	
All F		0.009 (0.47)		0.008 (0.38)		-0.025 (-1.25)		-0.031 (-1.43)
Const.	0.69*** (56.62)	0.69*** (53.56)	0.65*** (44.66)	0.65*** (44.73)	0.65*** (45.43)	0.65*** (45.39)	0.63*** (40.33)	0.63*** (40.28)
N	410	389	406	383	399	378	394	373

Note: This table presents results from regressions of effectiveness over time. I find no significant relationship between the treatment status and effectiveness across any of the surveys.

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

A basic regression of the satisfaction index on majority-female or all-female treatments produces no statistically significant results. Because they are insignificant, coefficients for the majority-female treatment mean that students in majority-female groups were, on average, just as satisfied or unsatisfied.

I repeated the same process for evaluating effectiveness over time. Again the results are organized with the coefficients for majority-female treatment and all-female treatment appearing under each corresponding survey. None of the treatment coefficients were statistically significant. Majority-female groups, except in the second survey, had a higher average perception of group effectiveness. All-female groups had higher perceptions of group effectiveness over the first month relative to minority-female groups, but then faded to levels less than the control.

Table 3: Effectiveness Over Time

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Survey 1		Survey 2		Survey 3		Survey 4	
Maj F	0.002 (0.21)		-0.012 (-0.91)		0.001 (0.08)		0.007 (0.55)	
All F		0.002 (0.18)		-0.011 (-0.84)		-0.001 (-0.05)		-0.007 (-0.57)
Const.	0.64*** (75.36)	0.64*** (75.98)	0.65*** (70.34)	0.65*** (71.84)	0.64*** (69.38)	0.64*** (68.10)	0.63*** (67.72)	0.63*** (70.9)
N	412	389	407	383	403	380	396	374

Note: This table presents results from regressions of effectiveness over time. I find no significant relationship between the treatment status and effectiveness across any of the surveys.

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

End Survey

The results in this section use only data from the final monthly survey.

Theoretically, the levels of satisfaction and effectiveness in this survey capture the overall sentiment of the student toward their group. Table 3 presents regression results from end-survey satisfaction on the majority-female treatment. Table 4 does the same with end-survey effectiveness rather than satisfaction. The first and second column look at a simple regression with and without demographic controls. The third and fourth columns include an interaction between treatment status and the female indicator. The interaction is useful to see if the treatment status affects females differently than males in the treatment groups.

Table 4: End Survey Satisfaction

	No Controls	With Controls	With Interaction and No Controls	With Interaction and Controls
Maj Treat	-0.0178 (-0.83)	0.0013 (0.06)		
ACT Score		0.00555 (1.78)		0.00487 (1.53)
Age		0.0135 (1.77)		0.00921 (1.06)
Year Status		-0.0525*** (-3.48)		-0.0514*** (-3.40)
White		0.0468 (0.99)		0.0442 (0.93)
International		0.0302 (0.54)		0.0276 (0.49)
Min F, Female			-0.0424 (-1.27)	-0.0309 (-0.86)
Maj F, Male			-0.00446 (-0.14)	0.00601 (0.18)
Maj F, Female			-0.0432 (-1.67)	-0.0189 (-0.65)
Constant	0.629*** (40.33)	0.196 (1.02)	0.642*** (33.96)	0.312 (1.41)
N	394	385	394	385

Note: This table presents results from regressions of end survey satisfaction. I find no significant relationship between the treatment status and satisfaction.

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

With and without the treatment-female interaction, the treatment status does not have a statistically significant effect on individual level of group satisfaction. The result is consistent with and without controls of ACT score, age, year status, an indicator of white ethnicity, and an indicator of international student status. While the treatment may not have created an increase in satisfaction, females in minority-female groups also did not report, on average, lower levels of satisfaction with their group. Year status did have a positive effect that was significant at the 0.001 level. Freshmen students primarily composed the study group class, but it appears the few seasoned students involved felt lower levels of satisfaction with their group.

Table 5 looks at the same analyses except with group effectiveness as the dependent variable. Similarly, there are no statistically significant coefficients on either the treatment variable nor the interaction between treatment and female indicator. I expected an increase in satisfaction to lead to higher effectiveness. Because there was not an observed increase in satisfaction among individuals in treatment groups, the lack of change in perceived effectiveness is not surprising. The controls in these regressions do not have any significant coefficients. There potentially could be a positive effect from student's ACT scores, but it is only significant at the 0.1 level. The ACT coefficient result could also be due to random chance from multiple hypothesis testing.

Table 5: End Survey Effectiveness

	No Controls	With Controls	With Interaction and No Controls	With Interaction and Controls
Maj Treat	0.00697 (0.55)	0.0166 (1.26)		
ACT Score		0.00327 (1.76)		0.00354 (1.86)
Age		0.00722 (1.59)		0.00831 (1.62)
Year Status		-0.0158 (-1.76)		-0.016 (-1.77)
White		0.0195 (0.69)		0.018 (0.64)
International		0.00389 (0.12)		0.00547 (0.17)
Min F, Female			-0.0256 (-1.29)	-0.0134 (-0.63)
Maj F, Male			-0.00864 (-0.44)	-0.00366 (-0.19)
Maj F, Female			0.00201 (0.13)	0.0208 (1.2)
_cons	0.628*** (67.72)	0.384*** (3.36)	0.637*** (56.53)	0.360** (2.75)
N	396	387	396	387

Note: This table presents results from regressions of end survey effectiveness. I find no significant relationship between the treatment status and effectiveness.

t statistics in parentheses

* p<0.05 ** p<0.01 *** p<0.001

V CONCLUSION

Several other researchers and I conducted a field experiment where we randomized students into 6-person study groups with minority-, majority-, and all-female conditions. Students completed assignments and discussions with their groups throughout the semester long course. The students answered monthly surveys that allowed measurement of their satisfaction with their group and their perception of group effectiveness. This paper looked at the relationship between female-majority groups and individual levels of satisfaction and group effectiveness. I was unable to reject the null hypothesis that students in female-majority groups had no difference in satisfaction than female-minority control groups. This result persisted when including an interaction between treatment status and a female indicator and with demographic controls.

While I did not find a statistically significant result, I do not conclude that women in female-majority groups do not enjoy higher satisfaction. The measure of student satisfaction with their group was dependent on a collection of answers from monthly surveys. Students may have endured survey fatigue and might not have carefully considered every question. In other words, the surveys may not have captured students' opinion with enough precision to find a causal relationship with treatment status. Further study should use other estimates of student satisfaction to elucidate the potential relationship. A measure of students' satisfaction with their group using behavioral passive data would be ideal. More research in the area of group dynamics among various gender compositions will be useful for providing better, more productive outcomes at the individual, group, and organization level.

VI REFERENCES

- Apestequia, J., Azmat, G., & Iriberry, N. (2012). The Impact of Gender Composition on Team Performance and Decision Making: Evidence from the Field. *Management Science*, *58*(1), 78-93.
- Bengtsson, C., Persson, M., & Willenhag, P. (2005). Gender and overconfidence. *Economic Letters*, *86*(2), 199-203.
- Berge, L. I. O., Juniwaty, K. S., & Sekei, L. H. (2016). Gender Composition and Group Dynamics: Evidence from a Laboratory Experiment with Microfinance Clients. *Journal of Economic Behavior and Organization*, *131*, 1-20.
- Boning, B., Ichniowski, C., & Shaw, K. (2007). Opportunity Counts: Teams and the Effectiveness of Production Incentives. *Journal of Labor Economics*, *25*(4), 613-650.
- Cadinu, M., Maass, A., Frigerio, S., Impagliazzo, L., & Latinotti, S. (2003). Stereotype threat: The effect of expectancy on performance. *European Journal of Social Psychology*, *33*(2), 267-285.
- Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, *129*(4), 1625-1660.
- Dufwenberg, M., & Muren, A. (2006). Gender Composition in Teams. *Journal of Economic Behavior & Organization*, *61*, 50-54. doi: 10.1016/j.jebo.2005.01.002
- Fenwick, G. D., & Neal, D. J. (2001). Effect of Gender Composition on Group Performance. *Gender, Work, & Organization*, *8*:2, 205-225.

- Kamas, L. & Preston, A. (2012). The importance of being confident; gender, career choice, and willingness to compete. *Journal of Economic Behavior & Organization*, 83(1), 82-97.
- Nguyen, H.-H. D. & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, 93(6), 1314.
- Niederle, M. & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics*, 122(3), 1067-1101.
- Reuben, E., Wiswall, M., & Zafar, B. (2017). Preferences and biases in educational choices and labour market expectations: Shrinking the black box of gender. *The Economics Journal*, 127(604), 2153-2186.
- Shan, Xiaoyue. (2021). Does Minority Status Drive Women Out of Male-Dominated Fields? *Job market paper*.
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67, 415-437.
- Stoddard, O., Jessica, P., & Karpowitz, C. (2020). Strength in numbers: A field experiment in gender, influence, and group dynamics. *Unpublished manuscript*.
- Wuchty, S., Jones, B., & Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science* (New York, N.Y.). 316. 1036-9. doi: 10.1126/science.1136099