Spring 2021

Effects of Social Context on Women's Political Engagement: Evidence from Focus Group Experiments in Tanzania

Elizabeth Kay McGuire
Yale University Graduate School of Arts and Sciences, lizkmcg@gmail.com

Abstract

Effects of Social Context on Women's Political Engagement: Evidence from Focus
Group Experiments in Tanzania

Elizabeth K. McGuire

2021

Both theory and policy experts have sought to understand how norms around
women's political participation may shift. This dissertation provides evidence that
the engagement of women can be affected by social contexts at the micro-level.
Through two randomized field experiments, I show that the social referents who
are present during a discussion about women and gender culture in Tanzania have
a significant impact, not only in how those topics are discussed, but also on the
behavior of the subjects with regards to women's political participation. Between
these two experiments, I conducted over 400 total focus groups in both rural and
urban Tanzania. The treatment in each varied the relationship of the people present
in order to measure the effects of social context.

In the first paper, I show evidence that norms are a connecting link between
attitudes and behaviors. Those norms may be demonstrated by novel discussion
of the topic, or a conversation with a previously silent majority. I look at the effect
of such discussions within primary social groups, the most basic social units. I
conducted a randomized field experiment in Dar es Salaam, Tanzania, to understand
how attitudes and behaviors might change with exposure to the opinions of friends

and family. I used 348 focus groups made up of either the subjects' friends, or their family members to test the heterogeneous effects of these groups. In this factorial design, half of the focus groups, in both family and peer conditions, discussed female representation and other gender norms in Tanzania. The other half discussed general questions of politics and culture in Tanzania. Via survey instrument and behavioral outcomes, I then measured their attitudes towards a number of gender norms using metrics common in the discipline. I found that while attitudes largely were progressive and unaffected by treatment, measures of behavior changed when respondents talked about these gender norms, particularly when they talked about those topics with friends. My research provides evidence that when a person observes that their existing attitudes match a newly revealed norm of their primary focus group, the cost of the behavior is reduced and the person is able to express their attitudes through behavior.

The second paper explores patterns within the data to understand what social forces may be driving these effects. I build on the analysis of the first chapter and incorporate a rich data set collected during that experiment to explore possible explanations for the heterogeneous effects of the previous paper. I will show that even when the norms among family members and friends are similar to each other, the ways in which those norms are communicated, and the dynamics within that communication can cause heterogeneous effects between the two groups.

The final paper in the dissertation considers how the presence of a white for-

eigner influences how women discuss gender issues in Tanzania. In this experiment, I varied the presence of a white, female researcher in 80 focus groups in rural Tanzania. This design allows me to estimate the average effects of the foreign researchers' presence by comparing relative levels of engagement with the focus group questions given across the focus groups. I find that focus groups with a white researcher present had longer discussions, particularly on questions about gender and culture. These questions are ones in which the gender and foreignness of the researcher would be more salient. In some cases, these results suggest that data collected in the presence of a conspicuously foreign researcher might be more thorough than in the absence of such. However, the information may also be pandering to the perceived values of the researcher, making some types of data less reliable. In either case the project underlines the importance of considering the measurer, not just the measurement, especially on topics with evolving or sensitive norms such as gender.

Effects of Social Context on Women's Political Engagement:
Evidence from Focus Group Experiments in Tanzania

A Dissertation
Presented to the Faculty of the Graduate School
Of
Yale University
In Candidacy for the Degree of
Doctor of Philosophy

by
Elizabeth K. McGuire

Dissertation Director: Dr. Katharine Baldwin

June 2021

ACKNOWLEDGEMENTS

I express my most sincere gratitude to my dissertation committee chair, Dr. Katharine Baldwin, for her considerable support and expertise. In addition to her inestimable academic guidance, she was also an encouraging and optimistic counter-balance to the ruthlessly helpful Dr. Alexander Coppock, whom I was usually profoundly grateful to have on my committee. I also have deepest gratitude for Dr. Frances Rosenbluth, whose passionate academic and moral support has been a personal and professional boon. Choosing this dissertation committee was the most important success of my graduate school experience.

I am indebted to Josephine Warioba, who managed the execution of my field experiment with remarkable aplomb. Her first child was born two weeks before the end of the project and Josephine insisted on completing her role without missing a beat. I do *not* thank baby Jemimah for the adorable inconvenience. Colossal credit also goes to the field team who carried out the smoothest implementation of a field experiment I have been a part of.

Many thanks to the Political Science Department at Yale University and all the associated faculty who taught me during the course of my study.I also thank the participants of B-WGAPE, the Yale APWG, and CAPERS for there periodic feedback over the life of these projects.

I am extremely thankful to my parents for their life-long love and support, and for proofreading aloud everything I wrote for the first 13 years of my education. I must also thank my sister, Amie, and brother, Spike, for their emotional and technical support, respectively.

I wish also to express my enduring gratitude to Daniel Nielson and Mike Findley for introducing me to field experiments and to the African continent. I am proud to consider them pivotal figures in my life, as well as friends.

Finally, I would like to thank all the people who have supported me to complete the project work, especially Cleo O'Brien-Urdy, Lilla Orr and many of my cohort for their generous technical assistance, both for this project and many a pset before.

I cannot begin to express the gratitude that I have for the dear friends, family, and teachers who made me, but I feel it deeply.

To Bethany, Soren, Connor, Alex, Jason, and Ruth: Thanks for coming.

Elizabeth K. McGuire
Yale University
Date: March 15, 2021

The glory of God is
INTELLIGENCE,
or in other words,
Light and Truth

Contents

List of Figures

List of Tables

Chapter 1

Norms as a Bridge between Attitudes and Behavior:

Evidence from a Field Experiment in Tanzania on Gender Norms

Abstract

The gender gap in political participation in politics exists even on the most basic levels of political participation. While support for women's participation in politics has been on the raise globally, there is a gap between attitudes and behavior. In this paper, I show evidence from the African context that norms are a connecting link between attitudes and behaviors. Individuals may not always have correct perceptions of norms, particularly in cases where the the topic is sensitive or when the norm is changing rapidly. In such instances, individuals may develop new understandings of norms as a results of novel discussion of the topic, or a conversation that reveals previously concealed norms. I looked at the effect of such discussions within primary social groups, the most basic social units. I conducted a randomized field experiment in Dar es Salaam, Tanzania, to understand how attitudes and behaviors might change with exposure to the opinions of friends and family. I used 348 focus groups made up of either the subjects' friends, or their family members to test the heterogeneous effects of these groups. In this factorial design, half of the focus groups, in both family and peer conditions, discussed female representation and other gender norms in Tanzania. The other half discussed general questions of politics and culture in Tanzania. Via survey instrument and behavioral outcomes, I then measured their attitudes towards a number of gender norms using metrics common in the discipline. I found that while attitudes largely were progressive and unaffected by focus group composition or topic, measures of behavior changed when respondents talked about these gender norms, particularly when they talked about those topics with friends. My research provides evidence that when a person observes that their existing attitudes match a newly revealed norm of their primary focus group, the cost of the behavior is reduced and the person is able to express their attitudes through behavior.

1.1 Introduction

One of the most basic questions in the study of gender politics and the political economy of gender is "How can we increase women's participation in politics?" Globally, only 25% of members of parliaments are female and only 10% of heads of state are female(Vogelstein and Bro, 2020). In more local politics, the gender gap is even more apparent in local legislatures, where only two countries have 50% or more women in those bodies (Parline, 2020).

Policy makers have attempted to increase local political participation by providing role models(Clayton, 2015), using elites to change attitudes(Morgan and Buice, 2013), and changing institutions to be more permissive(Rosen, 2013). Attitudes in many countries tend to be in favor of more women in politics(Afrobarometer, 2015), but the number of women who actually hold public office does not seem to reflect pro-women sentiment.

Psychologists and social scientists have long studied the connection between attitudes and behavior. If a person has a change in attitudes, one might suppose that a change in behavior will follow(Ajzen and Fishbein, 2000). While one may assume that attitudes and behavior would move together, research often observes a gap between the two(Lien, 1998; Ajzen and Fishbein, 2005; Sheeran and Webb, 2016). Attitudes may change without affecting behavior, and a change in behavior will not necessarily be the result of a change in attitude.

One reason for this gap may be that the actor observes that the norms around the topic are not in alignment with her own attitudes. The behavior that is consonant with these attitudes then becomes more costly, and she will be less willing to align her behavior with her attitudes (Cloward, 2014). If she were to observe that the norm is actually in alignment, she may shift her behavior to more accurately reflect her held attitudes.

Which are most relevant to influencing women's political participation? I argue that when it comes to the basic political participation, the most salient norms are not those within parliament or political elites, but those within the most basic social structures. I focus this research on primary social groups, namely friends and family, looking at these two distinct groups for differences in their effects on a change in behavior

My research focuses on the case of Tanzania, where while 38% of the Parliament is women, only 8% of candidates across all political races are women(IRI, 2015). On even more basic measures of political participation, we see a gender gap between men and women. AfroBarometer (2015) reports that 48 percent of women say that they never discuss politics, compared with just 26 percent of men (Afrobarometer, 2015). While the relatively high level of women in parliament is largely due to gender quotas, these more local levels of political participation seem some what more difficult to shift.

Women are less likely to run for office at both national and local levels. However,

the most basic levels of political participation show a gender gap. In behaviors as simple as discussing politics in a social setting, we see that women are less likely to participate (Afrobarometer, 2015). Women are also less likely to attend community meetings, less likely to join others to raise an issue, and less likely to vote than their male counterparts (ibid). These basic levels of political participation have very low barriers to entry and should be easy to engage in without any external resources or training. However, the gap persists.

The main goal of this article is to examine the role of social interaction and exposure to existing norms in changing behavioral outcomes. I argue that by exposure to the progressive beliefs of members of their primary social groups on the topic of women's participation in politics, young women will be more willing to act on their own progressive attitudes on those topics.

In this paper, I present evidence that the gap between attitudes and behavior can be bridged, at least in part, by a change in perceived norms. Attitudes and behaviors must be taken in consideration with the social context. Norms affect what is considered socially acceptable. Even if an attitude is widely held, if the perceived norm does not match the attitudes, the corresponding behaviors will not match the attitudes. However, if those privately and widely held attitudes are expressed in a social setting, then the perception of the norm shifts and the behavior will feel socially acceptable.

In a field experiment in Tanzania, I randomized the content and composition of 348 focus groups to expose subjects to the opinions of their friends or of their family members about gender norms and female participation in politics. Each focus group was organized and conducted in one of four treatment conditions, varying on two dimensions. I randomized the topic of the conversation based on the questions that were asked of the focus group: Half of the focus groups discussed politics and culture in general in Tanzania, while the other half discussed gender norms, particularly about women's participation in politics. On the other dimension, I randomized the social groups that discussed these topics by assigning half of the focus groups to be made up of only friends, while the other half consisted of only family members. Immediately following the focus group, participants responded to a survey measuring attitudes towards various aspect of gender culture in Tanzania, including participation in politics. The survey also included two behavioral measures that involved referring others to participate in activities that promote women's participation in politics. The same survey was administered across treatments. In this way I am able to observe the effects of revealing norms, and variation in the weight of norms across social groups.

These conversations about gender norms had a significant effect on behavioral measures, while the attitudes remained unchanged and generally progressive. When exposed to conversations about gender norms, particularly with their friends, women did not change their attitudes towards these norms, but did change their behaviors. Specifically, they increased the number of people whom they referred

to participate in pro-women events by 44 percent. The effect of friends on these outcome measures suggests that norms among friends have more importance than norms among family members in affecting behavior. Compared to other policies intended to shift women's participation in politics via institutional or attitudinal changes, exposing women to existing progressive norms is likely a far easier treatment.

The paper will proceed as follows: I will first compare my treatment to existing theory and policy around women's political engagement. I will also define norms and review research on norm learning, as well as the definition of primary social groups. I will then argue that a gap between attitudes and behavior must be considered in the context of the surrounding social norms. I explain my design to detect the effect of a change in perceived norms within the two primary social groups via a field experiment, and explain the surrounding Tanzanian context. I will show that many of the measures of gender norms had little variation and were near universally progressive. I will also show that metrics created from those that did vary also were largely unaffected by the treatment. I will then show that the behavioral outcomes were affected by the treatment particularly among friend groups. I will then discuss the implications of these results and conclude.

1.2   Theory: Revealed norms, attitudes and behaviors

My research explores how political participation, especially at the lowest levels, may be affected by revealing the norms of a person's primary social groups. I argue that while other ways to increase political participation among women attempt to change the collective norm, changing the perceived norm may have an important effect on these behaviors. Because attitudes are already generally in favor of women's participation in politics in many settings, revealing this norm will help bridge the gap between an individual's attitudes about participation and their behavior.

1.2.1   Political participation & primary social groups in Tanzania

As has been shown above, the gender gap in political participation exists at both national and local levels in Tanzania. One may think that women are not participating in politics because it is not socially acceptable, which is to say that the norm is not permissive of the behavior. This implies that if people generally believe that women should participate in politics, they will. However, this is often not the case. In Tanzania, using the example of the context in which my research takes place, 75% of people agree that women should be able to win an elected seat just as much as a man?. The super majority of opinion would suggest that changes in attitudes would have only marginal effects in this context, because the consensus is that women should be able to hold office. While the attitude (and even the institutions)

surrounding women engaging in politics is permissive, the behavior does not fully match.

I argue that attitudes interact with behavior in the context of norms. If the perceived norm does not allow for the behavior, then behavior will shift away from the held beliefs or attitudes. Furthermore, I suggest that while role models may have some effect, the most salient norms are not the norms signaled by elites and members of parliament, who may feel very far removed for an average woman, but rather those who are closest to a young woman. These primary relationships will be the most important in the decision of whether or not to engage in the most basic political actions.

Sociologists refer to these most basic social groups as *primary social groups*. These groups are generally in three categories: Family, friends, and neighbors (Litwak and Szelenyi, 1969). These relationships also have the most potential for norm learning and norm diffusion. The focus group that took place was conducted in the neighborhood that the subjects were recruited from. The additional participants were not likely to travel a long distance for a such an activity, so they would likely all be close enough to be considered neighbors. Therefore, for the purposes of this experiment, the treatment groups were only friends or family, assuming that many in both categories would likely also be neighbors.

The differences between the two groups have some potential implications for understanding how gender norms may exist within them. Family groups tend to be more stable and long-term (Litwak and Szelenyi, 1969). These groups also have more control over necessary resources such as housing and income. Additionally, families, especially families in traditional societies like exist in Tanzania, often have hierarchical structure, wherein older generations are given implicit deference and authority. All of these have some bearing on how a woman may perceive norms and expectations of behavior.

Friends are more likely to have higher turnover in their membership. Because of this, friends may not participate in early development of perceived norms, as families do. These relationships also tend to be "opt-in," meaning that people have choice in whom they associate with, often based on some shared experience or set of values. Family relationships are "opt-out," meaning that the membership is the default and exiting the social structure takes (significant) effort. Again, the interpretation of expectations based on perceived norms may be affected by any these characteristics of the relationship. Because family and friend dynamics are so different, my design separates them to better understand the effects of each group on these basic levels of political participation. I hypothesize that the norms within these groups affect the attitudes and behaviors of the members differently.

1.2.2   Policies to increase participation

Because the gender gap in political participation is such a lasting issue, there has been significant work in both the academic and policy spheres to increase women's participation in politics. Typically, this focuses on increasing women's representation in elected office, particularly in national parliaments. The purpose of this goal is many-fold. An increase in women in legislatures improves descriptive representation. In a representative democracy, having the legislative body reflect the constituency is perhaps a worthy goal in itself. Descriptive representation also has a theoretical link to substantive representation, which is the difference in how women govern once they are in office (Wängnerud, 2009). Research based on randomized representation in India suggests that women do allocate public goods differently than their male counterparts (Chattopadhyay and Duflo, 2004). Similarly, female heads of state are less likely to engage in non-preemptive war (Dube and S.P., 2019). This does not necessarily advance women's participation in politics, especially at the very local level.

Gender quotas are some of the most popular options for increasing representation. Gender quotas take many forms. In Tanzania, the parliament has Special Seats for Women, which are allotted via proportional list system based on the share of the constituency seats that each party wins. No matter the mechanism, the goal is to increase statistical representation of women in the governing body. In theory, this has at least two effects on the country at large.

Gender quotas are intended to change norms around women's political engagement. If women and girls see women in power, the reasoning goes, they will have role models for their own political participation. This is an attempt to change attitudes and therefore increase the general norm of accepting women in positions of leadership broadly and in political life particularly. In fact, some research, again in the Indian context, shows that girls who are represented by women increase their aspirations for educational attainment (Beaman et al., 2012). Despite this and the fact that women have had special seats in Tanzanian parliament since 1985, the gap in political engagement persists. Even in parliament, though Special Seats for Women make up 30% of the parliament, only 25 seats, or 6% are held by women who won their seats in a constituency.

Other research has found that men in particular are more likely to support women participating in politics when cues from elites show support (Morgan and Buice, 2013). These elite cues are a signal that norms around gender have changed. This is a sort of pull, or increase in demand for women in top levels of political participation. Some more classical literature in sociology suggests that those changes in expectations around gender roles are important to changing norms around political participation (Volgy and Volgy, 1975). Quotas, elite cues, and changing gender roles all have some effect on gender norms. The question is how these norms can be translated to a change in behavior in the form of increased

women's political participation, particularly at local levels.

1.2.3   Norms: Perceived and collective

There are typically two definitions of norms used in general social science. In one definition, a norm is what a given person believes that other people, collectively, expect of them. In the other, a norm is the generally held beliefs about what is acceptable behavior in a given situation, regardless of how those expectations are perceived. This is important because the former definition allows for there to be a mismatch between how a norm is *perceived* and what the collective norm actually *is*. Lapinski and Rimal (2005) describe this distinction thus:

> At the collective level, norms serve as prevailing codes of conduct that either prescribe or proscribe behaviors that members of a group can enact. Individuals' interpretation of these norms, the construct of interest in this article, is referred to as perceived norms. Individuals may or may not construe the collective norm correctly; ... Because collective norms are seldom formally codified or explicitly stated (Cruz et al., 2000), there is likely to be divergence in how people interpret them. For this reason, an aggregation of perceived norms among members of a social system will likely not represent the prevailing collective norm.

This difference between a perceived norm and a collective norm is important for my design. I am going to be arguing that when people learn that the norms around women in politics are, in fact, permissive, then they will shift their behavior towards their own pro-engagement attitudes.

1.2.4   Attitude-Behavior gap

For my purposes, attitudes are defined as privately held opinions. In contrast, behaviors are manifested actions. In many cases, there has been an assumption that attitudes would be easier to change than behavior, as attitudes are privately held, whereas behaviors risk social sanctioning and or similar costs.

I propose that the gap between attitudes and behavior can be bridged by a change in the norm or a change in revealed norms. Zanna (1986) also theorized this connection the connection between norms and bridging the gap between attitudes and behavior, suggesting that a person will not have "attitudinally congruent behavior" in the presence of attitudinally in-congruent norms. My research finds experimental evidence for that connection.

The connection between perceived norms and attitudes appears to be endogenous. While this experiment shows that norms are constraining behaviors that would comes from attitudes, these constraints can also affect attitudes as people react to the situation created by those norms(Blaydes and Linzer, 2008).

1.2.5 Revealing norms

Attitudes lead to behavior only when existing norms allow for the behavior. Attitudes will not match behavior when norms are not in alignment with the behavior. If the *perception* of the norm changes to be in alignment with the attitudes, the behavior will also shift to align more with the attitudes. When a norm is learned to be in alignment with privately held attitudes (whether this learning takes place because of a change in the collective norm or a change in the perceived norm through observation of a previously concealed norm), the cost of the behavior decreases and the behavior will become more "attitudinally congruent" (Fazio, 1986). In such a case, the attitude does not need to change to see a change in behavior, only the understanding, perception, or observation of a norm.

Tankard and Paluck (2016) discuss this connection between the perception of social norms and social change. In their construction of the issue, observing the individual behavior is how norms are perceived. When they perceive a relevant norm, people will conform to that norm, even if it requires shifting away from preference maximization (Tutić, 2015). In my construction of the argument, attitudes precede the behavior, but need the perception of normative beliefs to create a change in the behavior. The change in perception of the norm could take place either by having a novel conversation on the topic that has not been previously breached (perhaps because the norm is assumed), or to a wider variety of people to detect the same silent majority.

Additionally, it may also be true that not all norms hold equal value. A person may weigh the norms of one social group more highly than another. While family groups may have a more direct effect on a person's day-to-day life, the opinions of friend groups and norms among them may have greater salience because the person has opted into this social network. How people react to a revealed norm may help us understand which of these groups carry more weight. Social interaction may take many forms, but discussion is a fundamental form of social-political engagement (Carpini et al., 2004).

This concept of the importance of the makeup of social groups builds on the idea that social interaction is the primary mechanism by which membership in a social group translates to political behavior(Mutz, 2002). Discussing topics with social groups informs members of the group about generally held opinions and may also shape their own opinion.

My study focuses on the effect of revealing the norms of these two different types of social groups on women's participation in politics. How does exposure to the norms within these different groups effect the subjects' attitudes and behavior around these gendered political practices? My hypothesis is as follows:

– The gap between attitudes and behavior can be bridged by a change in perception of norms.

– *When women observe that the norm is in agreement with their attitudes towards*

encouraging female participation in politics, they will shift their behavior towards those progressive attitudes.

– This effect may be heterogeneous between friend and family groups.

I test this hypothesis experimentally in Tanzania, using focus groups that were randomized by topic and makeup relative to the subject. The design and context of the experiment is outlined below.

1.2.6 The Tanzanian Context

I conducted this research in Dar es Salaam, Tanzania. Gender norms around political participation in Tanzania are complex and vary somewhat by level of government. Tanzania has had special seats for women in the parliament since 1985. The quota system allows for the indirect election of women to seats that are reserved for women (Yoon, 2008). The constituency, or regular seats in Tanzania are elected via single member districts in first-past-the-post-plurality. Along with submitting candidates for each (or most) constituencies, the parties also put forth a list of women for the special seats. The special seats for women are then allocated proportionally to the number of seats that each party won in the constituencies. The special seats make up 30% of the parliament. Women may also run for the constituency seats. Currently women make up about 38% of the Tanzanian parliament (IDEA, 2020). A one-third quota for women also applies in local politics, though the enforcement is less rigorous. The special seats may have changed some views on women in politics, but women made up only 8% of candidates in all elections in Tanzania (UNWomen, 2015).

Special seats hold all of the rights and privileges of constituency seats, though some Tanzanians wonder whether the occupants of special seats garner equal respect to their constituency peers.The disparity in the respect for special seats may point to why the policy has been ineffectual at changing the norm at lower levels of government. Though some evidence suggests that these special seats are used as a stepping stone or training ground for women to then move into constituency seats, Yoon also finds that the existence of these safer seats may discourage capable women from contesting in seats that are a greater risk to their position (Yoon, 2008).

Dar es Salaam is the largest city in Tanzania and is the economic capital, being a major port city for the East African region. The city used to be the political capital, but that distinction is (nearly) shifted to the interior city of Dodoma. Like many large cities, Dar es Salaam tends to be more ethnically diverse and more capitalist than other areas of the country, and also has a higher concentration of Muslims than most other regions. As in many countries, women in urban Tanzania are more likely to work outside of the home than in rural settings. Both male and female children are more likely to receive higher education. Citizens are more likely to have a more exposure to Western culture, as well as more access to outside groups, and media.

Because of historical trade routes in the global region, Dar es Salaam, which is a coastal city, has a higher Muslim population than the interior of the country, which is majority Christian. The dynamic between religious conservatism and liberal globalism makes Dar es Salaam an interesting case for the study of changing norms.

1.3   Focus group experiment design

The experiment was conducted in the summer of 2019 in Dar es Salaam, Tanzania. With the help of my field team, I conducted 348 focus groups, each consisting of one randomly selected primary subject and up to three other participants who were referred by the primary subject and who were known to her. The focus groups were conducted under one of four randomly assigned treatment conditions, varying on the primary social group used to create the focus group and on the topics of the focus group, priming discussion about either gendered or placebo topics. The four conditions were: (1) family of the subject, neutral political topics; (2) friends of the subject, neutral political topics; (3) family, gendered political topics; and, (4) friends, gendered political topics.

In the absence of a control group, the family groups serve as the statistical baseline in the comparison. The first arm then observes the effect of general conversation with family members on the outcome measures. This general political conversation serves as a placebo treatment. The second arm observes the effect of the same placebo conversation among friends, for comparison. The third arm enables me to observe the effect of a conversation about women in politics among family members, and the fourth, the same among friends.

Focus groups were conducted in a subset of neighborhoods in Dar es Salaam, Tanzania. Fifty-five neighborhoods (*mtaa*) from 10 wards were selected at random from a complete list of *mitaa* in Dar es Salaam. The neighborhoods selected covered all of the residential districts of the city. Business districts were dropped from the sample, as was a district that was prohibitively difficult to access. Six focus groups were conducted in each *mtaa*, for a total of 348 focus groups. Each focus group was randomly assigned one of four conditions, independent of municipal boundaries, for a total of 87 focus groups in each treatment cell. One primary respondent was randomly recruited for each focus group. The referred participants performed the treatment through their conversations, in which the primary subject was a participant. Table 1.1, below summarizes the treatment conditions.

The sample size is 348, as each focus group had only one primary subject, though identical outcome measures were collected for all focus group participants (n = 1278). This experiment is likely one of the largest of its kind to date. A recent, related study in Lebanon conducted an excellent experiment with 120 focus groups, resulting in 720 subjects (Paler et al., 2020). Other similar designs generally have less than 50 focus groups.

	Family	Friends
Non-gendered culture and politics (Placebo)	Discussing general politics with family members. **(87 focus groups)**	Discussing general politics with friends. **(87 focus groups)**
Gender norms and women in politics	Gender norms in politics with family members. **(87 focus groups)**	Gender norms in politics with friends. **(87 focus groups)**

Table 1.1: Summary of Treatment Conditions

To understand the effects of different social structures on the salience of these discussions, I also varied the makeup of the focus group. In half of the focus groups, the other focus group participants were all family members of the subject. In the other half, the participants were unrelated friends of the subject. The selection of the subjects and the other focus group participants is explained below. The design of the experiment is meant to simulate somewhat normal social interaction through guided focus groups. The focus groups make up the treatment of the experiment.

Primary subjects were chosen via random walk within the selected neighborhood using a randomized skip pattern to select houses. Once a house was selected, subjects were selected at random from a list of females in the household between the ages of 18 and 35. If she agreed to participate, the randomly selected woman is the primary subject. Her responses to the attitudinal and behavioral outcomes are the unit of observation.

Once the subject had agreed to participate, she was asked to refer either six friends or six family members according to her treatment assignment for participation in the focus group. By asking the primary subject to refer other focus group participants, the design increases the likelihood that other participants are those with whom the subject is most likely to talk on a daily basis. The design is somewhat similar in premise to some research done on norms regarding rape in Tanzania (Abeid et al., 2014).

The six referred focus group participants were contacted in a random order by a skip pattern and asked to participate in the focus group. The skip pattern continued until three other participants agreed to come to the focus group. The focus group took place one day after recruitment.

The focus groups took between 30 minutes and an hour to complete. To understand the effect of discussions of gender norms, the topics of the focus groups were either gendered questions or non-gendered questions. In the gendered condition, the questions were about gender norms and women's participation in politics in Tanzania. In the non-gendered condition, the questions were a placebo treatment, and discussed culture and politics in Tanzania in general. The variation between the two sets of questions was limited, as much as possible, only in their gendered

elements. For example, the non-gendered group asks "What role does education play in Tanzanian culture?" and the gendered group asks "Do you think that it is as important for a girl to get an education as it is for a boy to get an education? Why do you think so?" Each group was asked questions about culture, politics, education, and aspirations. The content of the conversations between the friends and family groups were quite similar, both often listing the same important cultural elements (food, dance, peace) and challenges for young girls (long walks to school, 'temptations', etc.) for example. A full list of focus group questions can be found in the appendices.

Focus groups were facilitated by a trained enumerator who facilitated discussion and also took notes on the content of the focus groups[1]. For the purposes of this paper, the notes are useful as a manipulation check to ensure that the focus groups did take place and that the intended topics were discussed as designated by random assignment. The facilitator also recorded at the end of each question whether or not the discussion was dominated by men (if men spoke more than 75% of the time) and also if the conversation was dominated by old people. All focus group participants were compensated for their time.

Immediately following the focus group, an enumerator team interviewed the respondents individually to create the outcome measures. The survey included both attitudinal and behavioral measures. The enumerator who conducted the survey was always a different person than the focus group facilitator, allowing the respondent to change expressed opinions from the focus group without social duplicity. The enumerator was never present during the conduct of the focus groups, and the interviews were conducted individually and in private. The survey was identical across all treatment groups. The survey enumerators were blind to the treatment condition as much as possible. Re-response rate was also near 100%, with only a handful of respondents needing to be surveyed on subsequent days. All respondents were compensated for their time for both the focus group and survey. The resulting data has 1248 observations, of which 348 are primary subjects.

I begin by describing the demographic characteristics of the focus groups by focus group. The age and gender makeup of the focus groups did not vary significantly across treatments as shown in Figure 1.3 below. Focus groups made up of the primary subjects family member were not likely to be older or more male than friend groups. Subjects generally referred young women to participate in the focus group with them, whether they were referring friends or family members.

The consistency of age and gender of focus groups across treatments is important because one might also expect that younger people are more likely to hold more progressive ideologies, especially in urban environments such as this. Some research has also shown that women are more affected by female role models as well (Wolbrecht and Campbell, 2017).

---

[1]These notes will be used more thoroughly in the second paper of this dissertation to understand the group dynamics of these focus groups.

Figure 1.1: Age and Gender by treatment
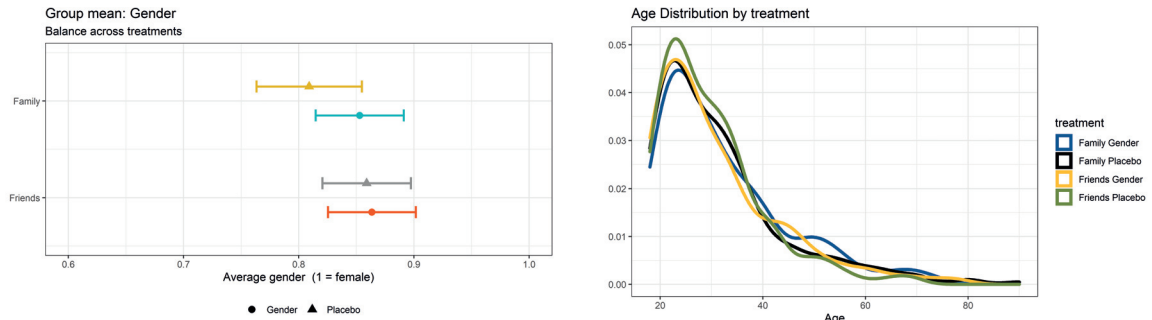
1.3.1 Attitude and behavior data

In this design, I collect three types of data. During the focus group, I collect information on the dynamics of the group, including if the conversation is dominated by one type of person. The survey, which was administered to all focus group participants, provides data on attitudes toward gender and women participating in politics. The survey was based on several existing studies that measure gender norms and role model effects, including a set of questions that have been used to measure gender norms in the United States for several decades (Charles et al., 2018 (Working paper)). It also includes short batteries of questions on household bargaining power (taken from the Demographic and Health Survey), political participation, aspirations for educational and career attainment for themselves and for young daughters (taken from Beaman and Duflo's work in India), and potential interest in holding political office. The data collected from the primary subject, who was selected via random walk, is the primary outcome of interest and is the dependent variable in all regressions around both opinion and behavior. The survey data from the other respondents is used in this paper to show that most of the participants have similar attitudes towards gender norms. These measures contribute to my understanding of the treatment (discussion among friends or family), not the outcomes.

Thirdly, subjects were also invited to two tasks with behavioral outcomes. Subjects were informed of a "Gender Festival" that would be taking place in Tanzania a few months after data collection. The gender festival includes information on women's initiatives, trainings, and various other meetings and events. To attend or want to attend would be to show a marked interest in the cause of elevating women's role in society. Respondents were asked if they would like to share their contact information in order to receive more information about the upcoming festival. They were also invited to share the contact information of others who might be interested in hearing about the event. Like the attitudinal outcomes, the behavioral measure was taken in private, away from the other focus group participants.

The behavioral measure was recorded as a binary of whether or not they referred

themselves for this information, as well as an integer count of the total number of people they referred for the same. They were also invited to refer themselves and others to be contacted by a local organization that helps women run for local political office, and the information was recorded as above. Such training to help women run for office is a regular activity for many NGOs in Africa seeking to advance women's participation in politics. UN Women, National Endowment for Democracy, and others have this as a regular program in various African countries. As there was a lack of variation in the binary variables (nearly everyone referred themselves and others), the salient measure of this outcome is the total number of people that they referred to receive information about the event. This outcome is both a measure of encouraging others to participate in these pro-female engagement activities, and a form of political engagement in itself, as encouraging political engagement is, at its core, also political engagement.

The effect of the treatments on each outcome was estimated with two regressions, testing the effect of the friends (as compared to the family) and the discussion of women in politics (as compared to the placebo) and then the same with an interaction term, as below.

$$Y_i = \beta_1 X_{friends} + \beta_2 X_{gender} + \beta_0 + \epsilon$$

$$Y_i = \beta_1 X_{friends} + \beta_2 X_{gender} + \beta_3 (X_{friends} * X_{gender}) + \beta_0 + \epsilon$$

The data collection finished in August of 2019. The design and analysis plan for the project on the attitudinal and behavioral data hypotheses was pre-registered with EGAP, March 10, 2020, before the treatment indicator had been added to the data. Though the data had already been collected, the treatment indicator was not merged with the outcome variables until after the pre-registration was submitted. The analysis code was written using the real outcome data, merged with a shuffled treatment variable. Once the pre-registered analysis was submitted, the shuffled treatment variable file was swapped for the real treatment variable file and the same analysis code was run to return the results. The change in treatment files was the only change to that analysis. Initial analysis was done in early April, 2020.

In the theory explained above, I hypothesize that when existing progressive attitudes do not match the current behavior, then exposure to previously unperceived progressive norms will result in a shift in behavior. Below I will show that many of the attitudes towards gender norms and female participation in politics expressed by survey respondents were nearly nearly universally in favor of women's participation and women's rights. These progressive opinions from the participants, even in the placebo treatments, demonstrate that norms that the subjects were learning about (if they were learning) were a progressive norms. Those that did have variation were largely unchanged by the focus group discussion. Though the measures of attitudes were not affected by the treatment, I show that the behavioral outcomes did shift in a way that matched the existing attitudes. I will also show

13

that these conversations about gender norms were most effective when had with friends, rather than with family members.

The unit of observation is the primary subject, who was a female, age 18-35, selected via random walk. The other participants in the focus group were referred by her, so they are not randomly assigned to treatment as she was. While I have survey data for all of the focus group participants, unless otherwise indicated, this analysis only concerns those women who were randomly assigned to treatment in the experiment. The much of the remaining data will be used the second chapter to understand the mechanisms observed in this paper.

In the next section, I will use these data to show that while attitudes did not change with these discussions, behaviors were affected when the primary subjects were exposed to the opinions of others and therefore learned the existing norm. Furthermore, these results are found only in the friend-based focus groups, not in the family groups.

1.4 Results: Attitudes and behavior

1.4.1 Null effects on attitudes

The measures of attitudes in the survey included measures of household bargaining, educational investment in the girl child, views on the morality of women working outside of the home, etc. Instead of taking each of the questions in turn, I grouped the questions by topic in order to create an index for each measure. For example, in household bargaining, I asked the following: Who makes decisions about large purchases? Who makes decisions about small purchases? Who makes decisions about visits to relatives? These are all ways of measuring different aspects of household bargaining, so they could be combined into an index.

Not all measures collected on a particular topic were used, however. As shown below, many of those measures did not have significant variation. In many cases, even with questions that work well in other research contexts, nearly all subjects responded in the same way. In all cases of outcomes measures which had no variation, the universal response was affirming of women´s role in politics and generally expressed progressive views of gender norms. I found this lack of variation in these opinions across treatment groups. Figure 1.2 below shows the distribution of a few of these attitudinal measures, for all focus group participants. Those measures that did not have significant variation were not used in the indices, as prescribed in the pre-analysis plan registered before the treatment variables were merged with the outcome data.

The lack of variation means that these attitudinal outcomes are not particularly helpful in measuring a change in attitudes, but they do show that the existing norms were progressive. In all cases in which there was significant variation on an outcome measure, the consistent opinion was on the progressive end of the scale.

Figure 1.2: Sample of outcome variables with no significant variation
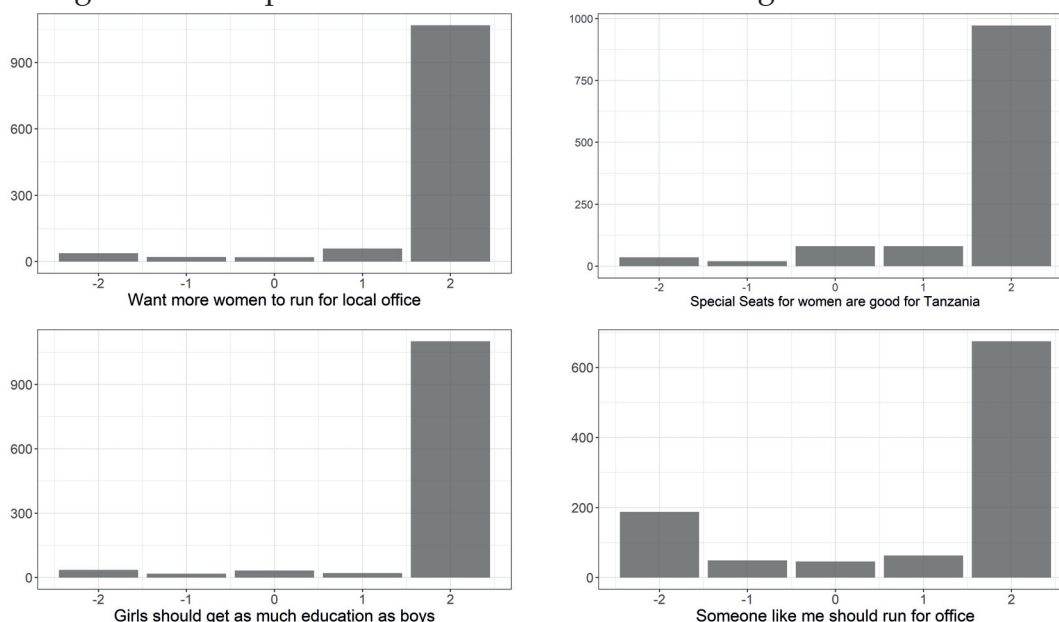
A review of the content of the focus groups confirmed that there were a scant few comments that expressed an objection to female political participation. The norms were progressive, so updating the perceived norms via the focus groups would lead the listener to believe that progressive attitudes were the norm.

With the remaining measures (those which did have enough variation to show potential effects), I used factor analysis in the pre-analysis plan to identify which variable could be combined into indices. I identified sets of variables that correlated both statistically and thematically. This meant that in some cases, the *a priori* categories based on thematically similar questions, did not in fact, correlate well. In some cases, thematic categories were split to better represent correlations within the data.

The measures that were thematically similar and highly correlated, I combined to create a total of nine indices. Returning to the previous example of measures of household bargaining power, all of the questions were highly correlated. Because these all measured the same topic (household bargaining power) and were statistically correlated, I combine them into a single index of household bargaining. [2] In this way, the indices should reflect both *ex anti* theory, and *ex post* patterns based on the data. The indices are:

1. Educational investment in girls
2. Household bargaining power

---

[2] In the case of measures of norms around women participating in politics, the measures were split into two groups of two questions each that were correlated within but not between each other.

15

3. Ideal age for marriage

4. Man as the only breadwinner

5. Mother/child relationship

6. Working mother effect on children

7. Morality of female political participation

8. Norms around female political participation

9. Own political participation (willingness of the respondent to participate in politics themselves)

Because the pre-registered analysis used the actual data (without the attached treatment information), these indices were created before the analysis was run and the indices pre-registered along with the rest of the analysis. For a full treatment of which variables were combined to make indices, see the Appendix.

Using these indices, I am able to show that treatment did not have a significant effect on these indices of attitudes. Speaking with friends or family members, about either general politics or women in politics left the attitudes of the subject unaffected. Figure 1.3, below, shows the group means (with 95% confidence intervals) for an index that measures support for investing in the education of a girl child. The values of this index were between -1 and 1, with 1 being more support for investment in the girl child. Most importantly, all four group means are statistically indistinguishable from each other, across the treatments. This pattern is similar across the nine indices. The full set of results for all pre-registered outcomes can be found in the Appendix.
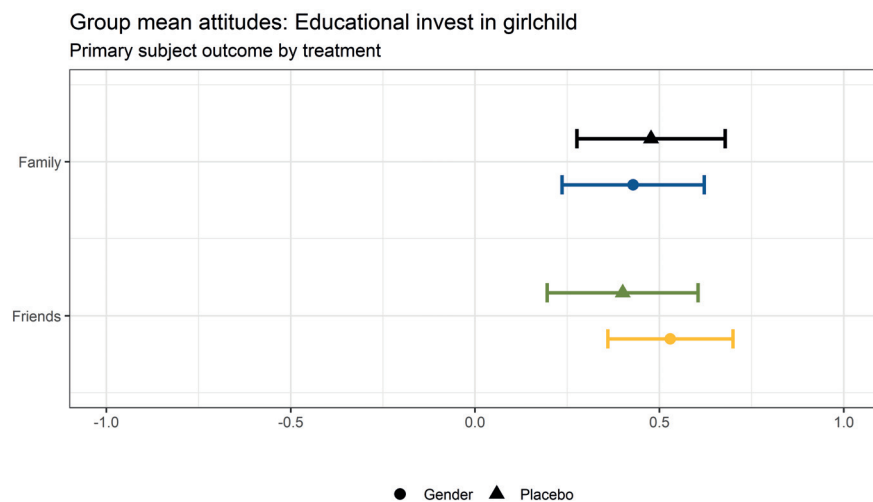

Figure 1.3: Constant attitudes across treatment

Consistent with my hypothesis, attitudes were not affected by exposure to the treatment. Learning the opinions of friends or family members about women's po-

litical participation did not change the individuals' opinions. Without a behavioral outcome, this may look as though these conversations did not have an effect at all and that such conversations would not be useful in increasing such participation. However, according to my hypothesis, exposure to these norms will lower the social cost of this participation and increase the behavior as it shifts towards the attitudes.

1.4.2 Positive effect on behavior

Consistent with my hypothesis, attitudes were not affected by exposure to the treatment. Learning the opinions of friends or family members about women's political participation did not change the individuals' opinions. Without a behavioral outcome, this may look as though these conversations did not have an effect at all and that such conversations would not be useful in increasing such participation. However, according to my hypothesis, exposure to these norms will lower the social cost of this participation and increase the behavior as it shifts towards the attitude.

1.4.3 Positive effect on behavior

Consistent with my hypothesis, despite the lack of effect on the attitudes of the subjects, I do find an effect on the behavioral outcome measure. Again, the behavioral outcome is the total number of people whom a subject referred to a pro-women's participation activity. Among primary respondents, the maximum number of people referred was four (other respondents referred as many as seven.) The data did not include significant outliers to drive results. These behavioral outcomes are a more costly measure as they have to provide the name and contact number of someone who may be interested. As they are inviting people to these events, providing the name is, in itself a low-level political act. This is also similar to a discussion of mechanisms by Goyal, which shows that women are mobilized to political participation by other women. Networks of women become politically active together, especially when encouraged by another woman (Goyal, 2020).

Between the treatment conditions, I do see a change in the behavior outcome, but only in the focus groups made up of friends of the subject. Figure 1.4 below shows the group means across treatments for the behavioral outcome. This clearly shows that while conversations in the family groups did not have an effect on the number of referrals, in conversations with friends, learning about the opinions of others increased the number of referrals given.

These results are quite large. The average number of people referred in the placebo condition among friends was 1.5. When subjects spent time listening to the opinions of their friends about gender culture and politics in Tanzania, they referred on average 2.5 individuals to these activities, exactly one individual more than in the placebo group. The focus groups in which the subject participated with
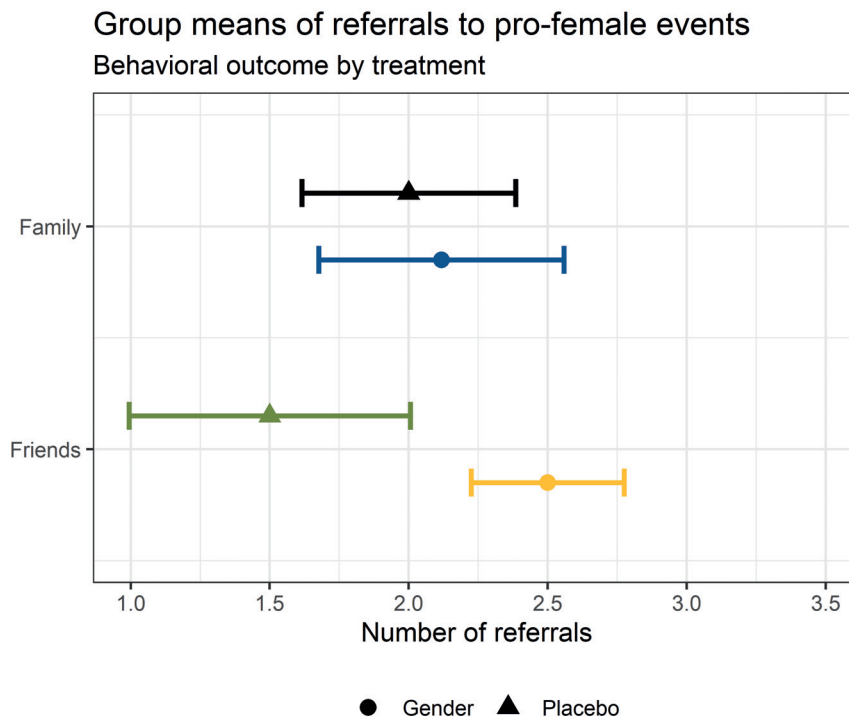
Figure 1.4: Group means of behavioral outcomes by treatment.

members of their family did not have the same effect on the subjects. The full regression results are in the Appendix.

Exposure to norms via conversation had no significant effect on the attitudes of the subjects. Those attitudes were largely progressive, in both placebo and treatment conditions. However, with exposure to opinions of their friends, subjects did shift their behavior to increase low level participation. These outcomes are consistent with my hypothesis.

One possible explanation for this difference between friend and family groups is that the content of focus groups with friends was significantly different from the content of focus groups with family members. While a full review of the transcripts is forthcoming in another paper, I did a random audit of family and friend focus groups in the gender discussion treatment to look for any significant differences.

I randomly selected 20 focus groups transcripts, 10 from friends and 10 from family focus groups in the gender discussion treatment. After reading through all 20 transcripts, I found that conversations were not discernibly different between friends and family members when talking about women in politics and culture. Across all questions, there was near universal support for women's education and participation in society [3].

---

[3]These transcripts will be discussed in chapter 2 of this dissertation. While the sentiment was

Because both family and friend groups near universally expressed pro-women attitudes during the focus group, the treatment seems to consistently be exposure to progressive norms. The difference between family and friend groups, is not entirely clear from this data, but focus groups, as a treatment, were expressing pro-female norms in a way that motivated the subjects to behave differently.

1.5 Conclusion

Behaviors do not always cleanly correlate with attitudes. People may often behave in a way that is not quite consistent with their beliefs. In this paper, I provide evidence for the idea that norms can help bridge that gap. Norms often act as a constraint on our own preferences. If a subject learns that their privately held attitudes match the previously concealed norms, the cost of acting in accordance with the norm decreases. Therefore, the external behavior moves closer to the internal attitude.

Through a randomized experiment I show that when subjects are exposed to the opinions of others regarding women's cultural and political roles, they are more likely to refer others and themselves to participate in training to run for local office and a gender festival. The referrals are a small act of political participation that show a shift in behavior, even when attitudes remain the same. However, I only find this result when the person is speaking with their friends, not with their family members.

While many policies have been used to try to increase women's political engagement, these policies have had varying degrees of success, especially at the most local, fundamental levels of political engagement. Demonstrating that the norm is permissive with increased political engagement may be one of the easiest, lowest-cost ways to shift that behavior. Similar findings show that people are more likely to vote if they are told that their neighbors all vote as well(Gerber and Rogers, 2009).

While policies like gender quotas and elite signals may take significant effort, including changing democratic structure, encouraging conversations with friends about these topics is likely significantly lower-cost. Furthermore, increasing the number of women in parliament is a top-down approach that is quite socially distant from the most local levels of political participation. My findings provide evidence that primary social groups may be an effective avenue to increase women's political participation on the most basic level.

The idea that people need to know that their opinion is socially acceptable in order to put it into action is, in many ways, already in deliberate practice. Mass protests are intended to demonstrate to others that the presence of a large group

nearly universally progressive, the text does reveal some interesting patterns between the two groups.

of people with the same opinion. Petitions try to convince politicians that their constituents have a norm in favor of a particular policy. A large number of blank black Instagram posts send a low-cost signal to others to demonstrate the prevalence of a particular attitude. The context that exists around an attitude matters. My research shows that even if many people hold the same attitude, if the majority remain silent, the perceived norm may not match the collective norm, which can have a chilling effect on the public's actions. When the actual norm is revealed, the cost of action is reduced, even when attitudes remain the same. As a result, the gap between attitudes and behaviors can close, at the very least on the margin.

These results are consistent with recent findings in Saudi Arabia, which show that men were more likely to encourage their wives to work outside on the home, once they learned that other men also approved of their own wives working (Bursztyn et al., 2020). The similar effects in two context suggest that the effect has external validity beyond the Tanzanian context.

The mechanism by which the these outcomes occur cannot be inferred by the results of this experiment. Further mediation analysis is needed to specify beyond the causal inference of this design (Bullock et al., 2010). For example, it is unclear what aspect of friend relationships causes these heterogeneous treatment effects of conversation.

One possible mechanism that might explain this difference may be the hierarchical structure of family relationships. Parents and older adults may have implied authority over other members of the family. A vertical structure like this is not necessarily found in friend relationships, which do not have an implicit hierarchy. The authority may also carry into a hierarchy of opinions. If a key figure is not present when revealing norms, the perceived norm might not shift as readily, since the opinion of the key figure is not revealed. In a conversation with family members, a family leader, such as a father figure may be missing from the conversation. Even if the rest of those present reveal that they are in favor of women's political participation, if the opinion of the father figure is not revealed, the perceived norm of the unit may not shift, as the opinion of the missing figure is a sizable gap in the information.

In this case, most of the focus group participants were female. In a group of friends, there may be no conspicuous absence. But in a family group, the missing voice of the father figure may be a drag on the shift in perceived norms. One way to test this may be to run a similar experiment that randomizes the presence of the subject's father or other family authority. Such an experiment could test the effect of a thought leader in the norm-revealing conversation.

The observed results were not caused by differences in the age or gender makeup of the focus groups. As was mentioned previously, there was not a significant difference in the age or number of men in the focus groups across treatment. While there may be hierarchical structure based on age and gender, especially in the Tanzanian context, that structure alone cannot account for the results, independent

of family relationships.

Further similar focus group experiments could vary how much each person was required to speak, in order to understand if the egalitarian nature of such friendships is the cause. Or, an experiment could vary the presence of authority figures, such as fathers or breadwinners, to understand if the hierarchical nature of families may be dampening the effect. Further research is needed to understand what about friends as a primary social group causes the observed effect.

Dar es Salaam is fairly heterogeneous, especially when compared to more rural areas of Tanzania. In the villages, most of the population come from the same ethnic group, religious background, and specific cultural context. If this experiment were repeated in those rural contexts there may be different results based on the fact that the differences previously pointed out between family and friend groups would not be as pronounced. Life-long friends would have early access to norm creation, and leaving a friend relationship would likely mean leaving the village entirely. Repeating this experiment in such a context would be important to know if this result can carry to more homogeneous communities.

These results show great promise for a low-cost path to increasing women's political participation on local levels. While further investigation is needed to understand the precise mechanism, the effect of simple conversations among friends may be all the encouragement that some need to become more engaged.

Chapter 2

Explanations for observed effect of friends on political participation

Evidence from additional data

Abstract

My previous research has shown that by discussing female political participation with friends, women will change their behavior regarding political participation, even without changing their attitudes. That experiment also shows that similar conversations with family members do not have the same results. In this paper, I build on the analysis of that experiment, using additional data that was collected during the experiment to better understand the heterogeneous effects: why was the effect seen in friend groups but not family groups? I find evidence that the change in behavior was not due to priming, but rather, gives suggestive evidence that this heterogeneity is due to differences in the amount of novel information that is given in friend groups relative to family groups. I leave open the possibility that the bulk of the difference in effects are due to differences in salience in the information received.

2.1 Introduction

In my previous paper, I demonstrated the potential for increasing women's political participation through a relatively light-touch intervention which had women discuss gender norms within their primary social groups. I also found important heterogeneity depending on whether women discussed female political participation with family members or friends. Only in the latter group did I find that women would change their behavior after conversations about women's participation in politics. I theorized that this was because the women themselves already held those beliefs, but that only when they saw that the norm matched their beliefs were they willing to take the more socially costly measure of acting on them. Conversations with family members did not have the same outcome. In those groups, the subjects changed neither their opinions nor their behavior because of a conversation about these topics. I argued that the transmission of norms was different between friends and family in a way that allowed the subject to shift her views of the perceived norm.

In this paper, I will explore possible explanations for these heterogeneous effects. I will provide some evidence that even when the norms among family members and friends are similar to each other, the ways in which those norms are communicated and the dynamics within that communication can cause heterogeneous effects between the two groups. I will show suggestive evidence collected from the focus group participants and the focus group discussions themselves that supports the theory that it is not the content of the focus groups, but rather, the amount of exposure to the opinion that changes the subject's behavior as she updates her views of what the norms are.

Based on existing theory, I propose three mechanisms by which these focus groups could change the behavior of the subject. Conversations with friends members might transmit norms as a result of one of these mechanisms. The first mechanism suggests that information from friends is more *salient* to the subject than information from family members; therefore, it is more likely to change behavioral outcomes. For the second, I argue that focus groups with friends may differ in how much *new information* the subject gets during the focus group conversation. If this were the case, it may be because the subject already knew the opinions of family members, friends were more likely to speak equally, or that the content of the focus groups was such that friends were expressing progressive norms more often than family groups.

Finally, it may be that the subjects were more likely to speak during a focus group with friends, and by doing so are, in fact, not so much learning a norm, but priming themselves in a way that prompts them to act in accordance with those beliefs for the sake of self-coherence. A *desire for self-coherence* after expressing their views on the topics would then lead them to effects that we observe.

Another way to consider the two mechanisms is to think of them as *sending* and

receiving signals about norms. In the information mechanism, I discuss ways that signals can be more effectively sent: via new information, greater participation, and clear communication. The sending of these signals maps clearly onto my theory that the differences in behavior are a result of women updating their perception of the norm with new information.

To test these mechanisms, I will use secondary data collected during the focus groups and from the focus group participants experiment. The data includes notes taken during the focus group that show what each participant said during the focus group, a tally of how often each subject spoke during each question, and some estimate of how often the conversation was dominated by an elder or a male. These tests, while not experimental, may provide some insights that suggests the mechanisms at play (Bullock et al., 2010). Such results provide interesting avenues for research in the future to better understand these outcomes. I will also provide some summary statistics that better outline what the focus group, as a treatment, was actually like.

I will begin by outlining the existing research on this topic, particularly the in-group dynamics that speak to the outlined mechanisms. I will then summarize the experiment and the findings, as well as outline the data I collected, which will be used in this paper, as well as the tests that I will use to examine each of the possible mechanisms.I will then describe in some detail the conversations that took place during the focus group. I show that the treatments were carried out satisfactorily as intended. I will also show that conversations about gender politics and culture in Tanzania were almost universally progressive quite progressive, in both family and friend focus groups. Finally, I will report on the findings from the proscribed tests, report on the findings, and conclude.

2.2   Theory

In previous research, I theorized that when a young female observes that the norm is in agreement with her own beliefs, she will shift her behavior to be more consistent with those beliefs. In a large-scale focus group experiment conducted in Tanzania, I demonstrated that when exposed to the (largely progressive) opinions of their primary social groups, women were more likely to take small actions towards increased political participation of women, with effects concentrated among women who discussed female political participation with friends. While this pattern fits with the theory that was proposed, multiple mechanisms could underpin this effect. Additionally, the heterogeneous effects between family and friend groups provides both an additional puzzle and a way that we may be able to better understand these mechanisms: What is happening in friend groups that is not happening in family groups that causes these norms to more readily manifest in behavior?

Discussion and discourse is becoming a common method to attempt to change

norms within a community (Wegs et al., 2016). My experiment shows that the behavioral outcomes may be affected by discussion, even when attitudes are not. I have outlined three general categories for potential mechanisms of these effects: self-coherence, exposure to new information, and salience of information [1]. Below I will outline the theory motivating the design of the focus group experiment in Tanzania, as well as additional literature from the social sciences concerning these three mechanisms.

2.2.1   Norms: Perceived and collective

There are typically two definitions of norms used in general social science. In one definition, which I refer to as a *perceived norm*, a norm is what a given person believes that other people, collectively, expect of them. In the other, referred to as a *collective norm*, a norm is the generally held beliefs about what is acceptable behavior in a given situation, regardless of how those expectations are perceived. Importantly, the former definition allows for there to be a mismatch between how a norm is *perceived* and what the collective norm actually *is*. Lapinski and Rimal (Lapinski and Rimal, 2005) describe this distinction as follows:

> At the collective level, norms serve as prevailing codes of conduct that either prescribe or proscribe behaviors that members of a group can enact. Individuals' interpretation of these norms, the construct of interest in this article, is referred to as perceived norms. Individuals may or may not construe the collective norm correctly; . . . Because collective norms are seldom formally codified or explicitly stated (Cruz et al., 2000), there is likely to be divergence in how people interpret them. For this reason, an aggregation of perceived norms among members of a social system will likely not represent the prevailing collective norm.

This explanation on the difference between a perceived norm and a collective norm is important for my design and findings, as I argue that when people learn that the norms around women in politics are, in fact, permissive, they then shift their behavior towards their own pro-engagement attitudes.

2.2.2   Attitude-Behavior gap

For my purposes, attitudes are defined as privately held opinions(Allport, 1935). In contrast, behaviors are manifested actions(Fazio, 1986). In many cases, there has been an assumption that attitudes would be easier to change than behavior, as attitudes are privately held, whereas behaviors risk social sanctioning and/or similar costs. My experiment found that while attitudes do not change upon exposure, the behavior does.

---

[1]I thank Sarah Khan for the suggestion of this framing

3

2.2.3 Revealing norms

While norms and attitudes can both affect behavior (McDonald and Crandall, 2015), I argue that attitudes will lead to behavioral changes only when existing norms allow for the behavior. Attitudes will not match behavior when norms are not in alignment with the behavior[2]. If the *perception* of the norm changes to be in alignment with the attitudes, the behavior will also shift to align more with the attitudes. When a norm is learned to be in alignment with privately held attitudes (whether this learning takes place because of a change in the collective norm or a change in the perceived norm through observation of a previously concealed norm), the cost of the behavior decreases and the behavior will become more "attitudinally congruent" (Fazio, 1986). In such a case, it is not necessary to have a change in attitude to observe a change in behavior. Rather only the understanding, perception, or observation of a norm needs to change.

The focus group experiment that I designed in Tanzania shows that when participants are exposed to the norms of their primary social groups (friends and family), they shift their behavior to match their attitudes, even when their attitudes do not shift. However, not all primary social groups have the same effect. Friend groups are more effective at this change. By comparing the differences between the treatment groups on data collected during the experiment I will show evidence for the mechanism of these heterogeneous effects.

2.2.4 Information

The mechanism most directly related to the theory is that, during discussion, the focus groups are revealing information to the subject that she did not already know. The new information then informs subsequent behaviors. If this is the case, then I would expect to see that friend groups are communicating more new information to the subject than family groups.

In order to communicate a newly revealed norm, the information needs to be a) novel, b) communicated and c) understood to be widely held. Friend groups may be different from family groups in any combination of those aspects.

Through the course of domestic life, family members communicate more consistently with each other, even if only due to proximity. Such communication with family members could provide a way for family members to convey norms more effectively than friend groups, closing the gap between collective and perceived norms. If during a discussion the norms that are being communicated are already known, they will not likely have a large effect on the listener, as the discussion would not lead to a change in the revealed norm.

---

[2]While norms are one constraint on behavior, they are certainly are not the only one. Norms that are permissive of the attitude and behavior are necessary but not sufficient condition for the attitude to manifest in behavior.

In this case, the revealed norm would be the progressive norm. Women's political participation is an evolving norm, especially in Tanzania where the office of Vice President has recently been occupied by a woman. The default is the conservative opinion, which would be the perceived norm until the progressive norm is revealed. The novel information learned by the subject would then be the progressive, pro-women's participation attitudes of their friends or family members. For our purposes, the communication of previously unheard progressive opinion on the topics would lead updating the subjects' views on the percieved norm.

The norm must also be communicated. In order for a norm to be revealed, the norm must somehow be demonstrated. In my design, outlined in the next section, norms are revealed via conversations with primary social groups. If the conversation does not actually discuss the norm or does not clearly express the norm, it will not matter if the norm is novel or widely held.

In the case of women's political participation in Tanzania, groups of friends might actually have more progressive views than groups of family members. Even if the friend groups do not hold more progressive views than their family counterparts, they might not express those views in a way that clearly communicates the norm.

Finally, the information about norms must be communicated in a way that demonstrates to the listener that the views are generally held. If a view is only expressed by one person, then the listener may feel that it is a personal opinion. If instead those sentiments are echoed by the group, then it can be understood to be a norm.

Contrarily, it may be that the opinions of some members of the groups carry more weight than others. If those opinions are expressed, so they may be they only views that need to demonstrate a norm (Ellickson, 2001). However, if individuals who are important to setting and keeping norms are not vocal, the absence of their comments will inhibit the transmission of the norm. For example, an older woman might not be likely to dominate the conversation, but they may be the gatekeepers to changing norms[3].

Taken together, we might expect that in a family structure– even when the participants hold progressive views, and if someone expresses those views– if the views are not *generally* expressed by the group, the listener may not perceive them to be a norm. They may assume them to be held only by a small subset of the group, and therefore are not commonly held enough to change the relevant norm.

2.2.5   Priming and self-coherence

While the design of this research is meant to allow subjects to learn the beliefs of their friends and family members, the treatment is interactive. The subjects not only hear what their social units have to say, but also express their own views to

---

[3]Some research suggests that women are likely to perpetuate norms of FGM and child marriage in order to conform to the existing norm (Cloward, 2014)

the social units. The mechanism, then, might not be found in the hearing (**?**), but in the speaking.

The literature supports the idea that people are prone to act according to views that they have previously expressed (Swann jr and Buhrmester, 2011), not just to avoid the audience costs of being viewed as inconsistent or hypocritical by the listener (Fearon, 1994; Epstein, 1973), but also to avoid cognitive dissonance in themselves (Sela and Shiv, 2009-03). The inclination towards self-coherence moves people to ensure that their beliefs and actions are in alignment. Wheeler et. al. (2008) find that people are self-primed by expressing their own opinions. Interaction with these social groups is the primary mechanism by which membership in the group translates into political behavior (Mutz, 2002). In as much as voicing one's opinion on a politically relevant topic is a political behavior, by participation in the focus group, the subject has completed a basic political behavior, and is primed to complete another, in the form of the behavioral outcome that is measured (Smeesters et al., 2009a).

In this case, a similar mechanism may be at play. By communicating her views to others, the subject may be reinforcing and expressing her views to herself, perhaps for the first time. If a subject is more likely to express her views in a friend group than in a family group, it may explain the heterogeneous effects that we observe.

This examples refers to the literature mentioned above that suggests friend groups tend to have more egalitarian communication than family groups. In such a group, the subject, along with others in the group, is more likely to express her views, and by doing so prime herself for the congruent action (Molden, 2014).

The self-coherence mechanism considers the effects of a woman sending signals to her social groups and how this expression will lead her to act accordingly. Even though the outcome measures were taken in private, away from those in the group who may have heard her express those views, acting in a way that is self-coherent can reduce cognitive dissonance (Festinger, 1962; Harmon-Jones and Mills, 2019; Festinger, 1957). The theory put forth in the experimental results of this experiment is not consistent with this self-coherence mechanism.

2.2.6 Salience

A third mechanism deals not with how well the norms of friends and family members are communicated, but whether the subject actually cares about those norms when they are signaled. It is possible that women simply care more about the norms of their friend groups than those of their family.

Some research suggests that for adolescents, the norms transmitted by the parents matter more than norms signaled by peers (Smetana, 1999). Generally, people are found to conform more to the norms of a group that they identify closely with, so the norms of a family group may be more salient (Terry et al., 1999).

Other observational research has suggested that speaking with peers makes

adolescents more tolerant of others, although the opinions of parents is also a predictor of those opinions (Ekström and Östman, 2013). While a significant amount of research has studied the communication of political ideology from parents to children(**??**), more recent work has suggested that political discourse is more collaborative even in families (McDevitt and Chaffee, 2002; York, 2019). Furthermore, research on political participation in particular has suggested that peer relationships have the largest influence (Quintelier, 2015).

2.3   Design

2.3.1   Experimental procedure

The data for this paper was collected during a focus group experiment in Dar es Salaam, Tanzanian. A more thorough explanation of the experiment can be found in that paper. The experiment consisted of 348 focus groups, each of which included one subject, who was randomly selected, and up to three other participants who were referred to the experiment by the subject. These other I conducted participants were either friends or family of the subject, based on her treatment assignment. In addition to randomizing the subject's relationship to the other participants, I also randomized the questions that were asked the focus group as discussion topics.The four conditions were: (1) family of the subject, neutral political topics, (2) friends of the subject, neutral political topics, (3) family, gendered political topics, (4) friends, gendered political topics. Without a true control group, the family groups serve as the statistical baseline in the comparison. The neutral political topics serve a placebo.

The first arm then the effect of general conversation with family members on the outcome measures. The second arm observes the effect of the same placebo conversation among friends, for comparison. The third arm enables me to observe the effect of a conversation about women in politics among family members, and the fourth, the same among friends.

Focus groups were conducted in neighborhoods in Dar es Salaam, Tanzania. The 55 neighborhoods (mitaa) were selected at random from a complete list of mitaa in Dar es Salaam and six focus groups were conducted in each mtaa. Each of the 348 total focus groups was assigned one of the four treatment conditions, with 87 focus groups in each cell. The treatment was the conversation held with the participants in the focus group, in which the randomly selected subject participated. Table 2.1, below summarizes the treatment conditions.

The total sample size is of the experiment is 348, as each focus group had only one primary subject, chosen via random walk. Identical outcome measures were collected for all focus group participants (n = 1278). The design of the experiment is meant to simulate somewhat normal social interaction through guided focus groups. The subject could refer whomever she wanted for the focus group, within

	Family	Friends
Non-gendered culture and politics (Placebo)	Discussing general politics with family members. **(87 focus groups)**	Discussing general politics with friends. **(87 focus groups)**
Gender norms and women in politics	Gender norms in politics with family members. **(87 focus groups)**	Gender norms in politics with friends. **(87 focus groups)**

Table 2.1: Summary of Treatment Conditions

the limits of the treatment condition, in order to increase the likelihood that the other participants would be natural interlocutors with the subject. The design is somewhat similar in premise to some research done on norms regarding rape in Tanzania (Abeid et al., 2014).

Once the participants were recruited, the focus was conducted the following day. A trained focus group facilitator conducted the focus groups, which took between 30 minutes and an hour to complete. In the gendered condition, the questions were about gender norms and women's participation in politics in Tanzania. In the non-gendered condition, the questions were a placebo treatment, and discussed culture and politics in Tanzania in general. The variation between the two sets of questions was limited, as much as possible, only in their gendered elements. For example, the non-gendered group asks "What role does education play in Tanzanian culture?" and the gendered group asks "Do you think that it is as important for a girl to get an education as it is for a boy to get an education? Why do you think so?" Each group was asked questions about culture, politics, education, and aspirations.

The outcomes were measured via survey, immediately following the focus group. An enumerator team interviewed the respondents individually, measuring both attitudinal and behavioral outcomes. The enumerator was never present at the focus groups, and the interviews were conducted individually and in private. Re-response rate was also near 100%, with only a handful of respondents needing to be surveyed on subsequent days. The resulting data has 1248 observations, of which 348 are primary subjects.

The attitudinal outcomes included several measures of various gender norms. Several groups were taken from existing literature long-standing surveys in order to make them more comparable. For the behavioral outcome, we told the participants about two pro-female empowerment activities: a gender festival that would take place a few months later, and an NGO that trained women to run for local public office. We asked the women to if they would be interested in receiving more information about each of these and then asked if she knew any other people who may also be interested. The outcome measure then, is the total number of women that the respondent referred to these two events.

2.3.2 Data Collection on Focus Groups and Focus Group Participants

The theory of this experiment, as outlined above, focuses on the transmission of norms. The outcomes in the experimental analysis analyzed the effects of a communication of norms. To understand that treatment, we need more information, not only about what norms were explicitly expressed, but also any information about how they were transmitted, including group dynamics.

I collect a rich set of information on the content and dynamics of the focus groups, in addition to information on focus group participants and their outcome variables. First, the outcome measures of attitudes and behaviors were measured not only for the primary subjects who were randomly selected, but also the other focus group participants who were referred by the subject.

The second source of data is the notes taken by the facilitators during the focus groups. Facilitators took notes during the focus group to record what was said and by which participants. While these are not word-for-word transcripts, they do record what thoughts were expressed and which topics came up in response to the questions. They also can reveal qualitative information about group dynamics and sentiment of the focus groups. The information in the transcripts is essential for understanding the nature of the experiment. While random assignment determined who was recruited for the focus groups and what the question prompts were, the actual treatment was determined by the focus group participants. Their responses to the questions, as well as the dynamics within the group make up the treatment. The things that were said are a key (but not sole) element of this treatment. Which participants spoke and in what frequency and order are also import data to understanding the visible dynamics of the group.

For the purposes of this paper, I selected a random sample of these transcripts for qualitative analysis. The sample consisted of about one-third of the transcripts, weighted toward the conversations that had the gendered questions. A holistic reading of these transcripts, as well as coding patterns within the conversations provide illuminating insights into the dynamics of the treatments.

Finally, the facilitators recorded information on two specific social dynamics during the focus groups. At the end of each question the facilitator recorded whether either of two types of people dominated the conversation (spoke more than 75% of the time): men and older people. These are two groups that we may expect to get some deference in a conversation, and whose opinion may carry greater weight in norm-setting The facilitator also recorded that information about the focus group discussion as a whole, creating a rough measure of some social dynamics that we might expect to see in this context.

2.3.3 Observational Implications of the Mechanisms

I will first explore what was actually discussed in the focus groups. The content not only helps me understand whether and in what context positive opinions of women's participation in politics were expressed, but also to ensure that the treatment was carried out as expected. While the placebo questions did not bring up gender issues, there may still be an excludability violation if those topics ever or often came up organically in the placebo condition. The concern would be especially valid if this happened in one type of social group more than another. Reading a large sample of the transcripts helps to confirm that the treatment that was assigned was carried out as expected.

Using this data, I will explore some of the possible mechanisms for the changes in the behavioral measures that I observed in the experimental analysis. As stated above, the three mechanisms that I will explore are: exposure (receiving new information), self-coherence (the subject strengthens her opinion via communication), and salience (the opinions of friends matter more.) The last of these is exceptionally difficult to explore with the current data. Absent explicit expressions within the transcripts, this theory will perhaps only receive support by elimination of the others.

Friend groups could expose the subject to more novel information than the family groups in at least three ways. First, it may be that because more time is spent with family members than with friends, the opinions of family members on these topics are already known. If this were the case, we would expect that those who are speaking with their family members would have the same, progressive behavior towards women in politics, whether they are speaking about placebo topics or about gendered topics. Since the observed norm and the actual norm would already be in alignment, I would expect that the outcome for family focus groups, no matter the topic, would match the friend groups in the gendered topics condition. The behavioral outcomes from the primary participants has confidence intervals that are too large to make this judgement. By including the data from the other focus participants, I hope to shrink those confidence intervals to reveal this pattern in the behavioral outcomes.

The second way that information about norms may be different between family and friend groups is that the dynamics within the friend focus groups may allow for better communication about the previously concealed norm. If a single participant dominate the conversation in the focus group, then the subject may not believe that the expressed opinions are, in fact, a norm. Rather, it may sound like just the opinion of that person. The more people that a subject hears express these opinions, the more likely they are to believe that this opinion is the norm. Therefore, more egalitarian participation among friend groups may lead to a better transmission of the norm that women can and should participate in politics.

To measure this, I used the notes taken during the focus groups to measure how

often each person spoke and whether or not they spoke on each question. Using this data, I can look at how many questions included comments from every participant, as well as the total number of comments by each participant. I will use the former to detect a difference in full participation between the treatment groups. Using the latter, I will look at the variance in the total number of comments in each focus group as a measure of how equally distributed the participation in the was in the focus group. The difference in mean variance between the focus groups will be a measure of egalitarian participation.

Finally, the third way in which friend groups may be communicating more new information than family groups about norms is the actual content of the focus groups. The expression of previously concealed norms is perhaps the most basic view of the communication of gender norms: did friend groups communicate more novel information about (progressive) norms of women in politics? To understand this, I will use qualitative data from reading a random sample of about one third of the focus group transcripts. I will examine how often views that were against women's participation in politics were expressed and in which groups. If such views were more often in family groups, the difference in information may be a mechanism for the effect that we observed. Also, if the pro-participation views were expressed more forcefully in friend groups, this may also account for the effect. I will also use some visualization of the content of the notes to better understand the respective discussions.

I will then explore the possibility that women in friend groups changed their behavior not because of the content of the focus group, but because they themselves are communicating their own beliefs. It is not in learning previously concealed norms, but in the priming of speaking her own attitudes and having others hear them. The change in behavior that is observed is not due to learning a concealed norm, but confirming to herself her own attitude. Because we see different effects between family and friend groups, we would expect that, were this the mechanism, women would be expressing their views more often in friend groups than in family groups.

To test this, I will look at the average number of comments per question given by the primary subject. I will also look at the average number of questions she participates in. If the observed effect of behavior is due to the need for self-coherence of the subject sharing her opinion on the gendered topics, I would expect to see higher participation of the subject in the friend groups.

In summary, I have proposed three mechanisms that may explain why friend groups elicit a change in behavior while family groups do not. Table 2.2 outlines the observable implications of each of the mechanism proposed above, for clarity.

Mechanism	Empirical Test	(Results)
Information Received		
Opinions previously known	Behavioral outcomes for all participants	
Egalitarian participation	- Variance of comments - Measures of full participation on questions	
Content of discussion	- Word clouds	
Self-Coherence	- Comments per question from subject - Binary participation in question from subject	
Salience of Information	*Needs additional data*	

Table 2.2: Table of possible mechanisms and tests

2.4 Description of Focus Groups

The experiment varied the condition of the focus groups on two factors: the relationship of the people in the focus group and the content of the focus group. It is essential to confirm that the treatment was carried out as designed. Focus groups, especially those made up of people who are already familiar with each other, have complex dynamics that are difficult to fully understand. The purpose of this research was not to attempt to find the depth of these dynamics, but rather to point to some patterns within the focus groups which may be useful in understanding the observed effect.

As has been described previously, the people in the focus groups were recruited from referrals from the primary subject, who was selected via random walk. She was asked to refer people to the focus group according to random assignment, either only friends or only family members. These people were contacted by a mobiliser who organized the focus groups. The following day, a focus group facilitator met with those who had been mobilized. The participant confirmed the name and relationship to the subject of each of the focus group participants. The facilitator was instructed to not allow people who were not contacted previously to participate.

Were a primary subject would have had to refer someone to the focus group that was from a different social group than the one she was prompted to refer, the treatment would not be as assigned. There may be some motivation to do this, as there may be an assumption of compensation for their time and the subject may want to divert such compensation to a preferred individual or group. However, because focus group participants were contacted by the mobilizer and then con-

firmed by the facilitator, the primary subject would have had to coordinate fairly quickly and covertly to make such a deception work. Because I cannot confirm the family relationships via objective source, I cannot fully confirm that there were no ITT issues, but I can say with reasonable confidence that it would be unlikely.

The second factor was the content of the focus group. The content was varied via random assignment of the questions that were asked of the focus group. Questions about gender politics were intended to elicit discussion of women in politics, investment in the girl child and other gender culture topics. The other set of questions did not ask about these topics, but it is still possible that they would come up organically. For example, in a discussion about the parliament, the topic of special seats for women might arise.

The content could also vary based on the different opinions of the people in the focus groups. I will discuss the transcripts of the focus groups below. The attitudinal outcomes of the focus group participants shows that this is not the case. The post-treatment attitudes of the focus group participants did not vary by treatment condition. Figures 2.1 and 2.2 show two examples of outcome variables of the attitudes of all of the focus group participants by treatment group. The first is a measure of the participant's responses to questions about their own inclinations towards political participation. The second is responses about bargaining power in their own home from the same group. Both of these outcome measures are a index of similar and correlated measures on the topic. The group means of these outcomes are not significantly different. The remainder of the attitudinal outcomes for all participants can be found in the Appendix.
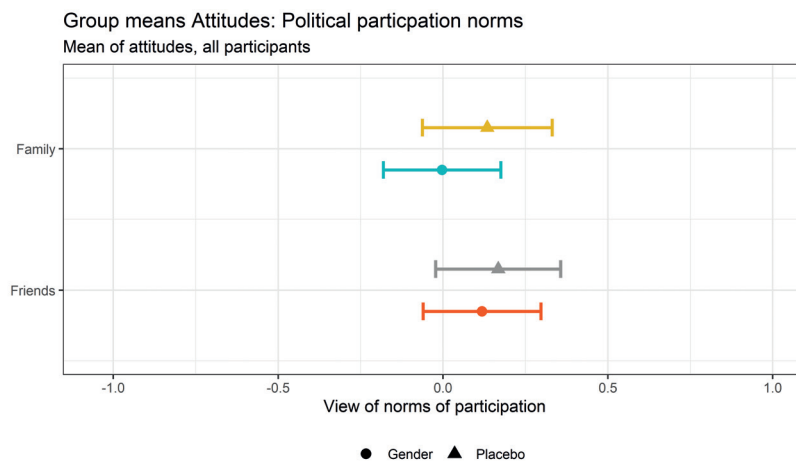

Figure 2.1: All participants attitudes on political participation by treatment group

The best way to confirm that the placebo discussion topics were not gendered is to refer to the transcripts that were written by the focus group facilitators during the focus group. Because these notes were taken while the facilitator was conducting
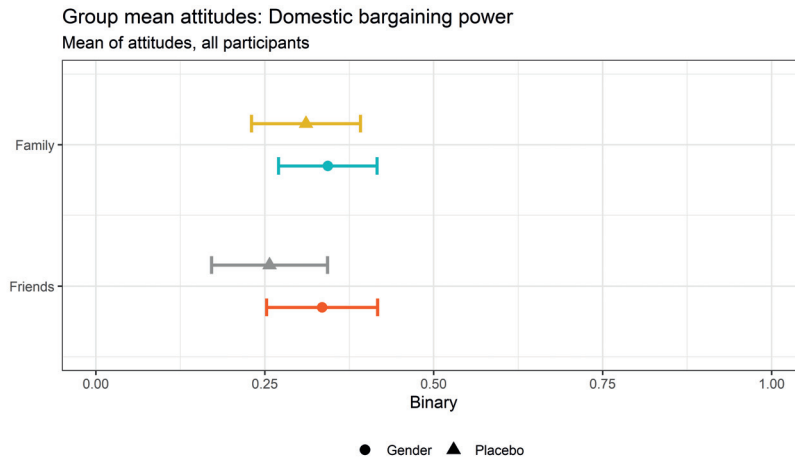
Figure 2.2: All participants views on domestic bargaining power by treatment group

the focus group, the transcripts are not word-for-word transcriptions, but rather summaries of the content of each comment. The transcripts can still provide insight into the topics of the discussion, as well the speaker, relationship to the subject, and order of the comments.

2.4.1 Discussion Content in Placebo Condition

In response to the control condition questions, which deliberately avoided gendered topics, women and gender issues did arise in about one third of the focus groups that were intended to be placebo conditions. Topics of gender and tradition are fairly salient in Tanzania, particularly in large cities such as Dar. In about half of those cases, the only mention of gendered topics was to list FGM or circumcision, either in answering the question about culture in Tanzania, or so say that one benefit of education was to reduce these practices. These were not about women and politics *per se*. Below are some examples of the comments that were made on the topic:

> *"Education is important because it prevents FGM because it would have continued to cause many deaths"* –Female respondent, age 23.

> *"Education is very important because it has reduced bad traditions like circumcision of women (FGM). It also reduces ignorance among women."* –Female, Primary subject, age 21.

Additionally, some participants also volunteered comments that reinforced gender roles and norms, such as commenting that if life went well for a person, the person would "Get a good husband if she is a girl" without mentioning a corollary

14

for a male. Another commenter said that education is good "for women and men to know their responsibilities." In only one case did the conversation actually touch on women in politics, in which one person mentioned that in the modern parliament "women are seen and remembered" as they had not been before. While in some cases a comment on FGM or another gender element might spur comments in agreement, in this and most other cases, the comment was stand alone and the transcript does not another participant echoing or contradicting the comment.

The spill over actually would cause an underestimate of the results, since some gender topics were being discussed in intended placebo condition. However, if these comments were to be framed as "doses" of pro-female conversation, the fact that all but a few comments in the treatment condition were discussing the intended topic makes the hand full of comments in the placebo conditions feel qualitatively minimal. Importantly, I do not see a significant difference in the number of these comments across family or friend conditions. Though in the sample, comments against FGM were more prevalent in the friend group than in the family group, in the handful of cases in which it was brought up.

2.4.2 Discussion Content in Gender Condition

Through reading a large, random sample of transcripts I found that in the treatment condition, which focus groups were asked gendered questions, the answers did, in fact, stay on the relevant topic. The treatment was successfully carried out and the focus group participants did share their thoughts on the topic. I found no clear example of the group talking about a topic other the gendered question when asked. In the treatment, the conversations were appropriately gendered. The only possible exception to this was in the questions which asked about what a good life would look like for a girl in Tanzania when she grew up. The responses to this question were remarkably similar whether the question asked about a young girl, or a young (ungendered) child. In this case the fact that respondents were commenting that a girl could be a doctor, teacher, or President of Tanzania [4], was treatment, even though the gender is not mentioned by the respondents.

The treatment condition also was almost universally pro-women's empower-ment and women's participation in politics. In the second question, participants were asked to reflect on ways men and women are expected to act differently in Tanzanian culture, if at all. In nearly all cases, the respondents did enumerate some traditional gender roles, such as that women work in the home or on the farm, while men are required to provide financially and do more physically demanding labor like construction. There were a few comments that mentioned that such roles

---

[4]At least as often, it suggested that a girl child may grow up to be *Vice* President of Tanzania. While this may seem odd, it is important to remember that since 2015, the Vice President of Tanzania has been a women, Samia Suluhu. The existence of a role model seems to have had an effect on the views of possibilities in girl's achievements.

were not as relevant in modern society, noting that both men and women work equally now. Both men and women were equally likely to enumerate these gender norms.

However, on further examination, these were merely descriptions of the norms, not necessarily a signal that the participants held traditional views on these roles. In fact, there were only a handful of comments that could be interpreted as being against women participating in politics, or prescriptive of women staying at home instead of working. In one case, the lone male in the group mentioned that in an ideal life, a woman would be married to a reputable man, wearing modest clothes, being a good mother. It is possible that these comments would also have been said about a male child, were the question posed, but in this context, it seems to be an endorsement of traditional gender norms.

In another case, one of the two males in the focus group mentioned that Special Seats for women in parliament were a waste of money and was generally discouraging of the practice. However, later, when asked about a good life for a girl child, this same person mentioned that she could be leader. In these few comments that were stating regressive gender norms, the comments were as likely to be given in friend groups as family groups, and were given by women as often as by men.

When asked about the Special Seats for women in parliament, there were a few participants who were not in favor of the seats. The objection was, however, not about women in politics, but rather, that these women were not good legislators because they did not have a constituency to be accountable to. Interestingly, these comments tended to either be held by the entire focus group or not mentioned at all.

In general, there was broad consensus that it was good to have women's representation in parliament, that girls should be educated as much as boys, and that girls could grow up to be leaders and economically successful. Below are some example quotes from these focus groups.

> *"We are expecting [women in parliament] to be our leaders in society. . . A girl should also receive the same educated as men to help liberate the nation from poverty "* –Female, age 22.

> *"[Education] is important because it will help a girl not be dependent on a man."* –Female, age 20

> *"[Special Seats] help to motivate their fellow women to join social issues in society* –Female age 54.

> *"In the past, women were not listened to but through these special seats, now women are listened to."* –Female age 54

"Home labor is also a challenge [to girls education] because most of her time is being wasted doing work in the home." –Male

2.4.3   Patterns within the Treatment Condition

A qualitative reading of the focus group also revealed some patterns within the treatment. The vast majority of focus group participants were female. In fact, only 15.5% percent of the participants were male, and in 54.5% focus groups there were no males in the focus group at all. This did not vary across treatment condition, even between friend and family groups

However, when men were present, males did did have an effect on the focus groups. When the focus groups happened to include men, the facilitators recorded whether or not men dominated the conversation (spoke roughly more than 75% of the time.) Across the entire focus group, men were not more likely to dominate the conversations in gendered conversations. However, men were more likely to dominate a placebo conversation in family groups than in friend groups, which hints at an underlying dynamic. Figure 2.3 shows these group means across the treatments for all questions.
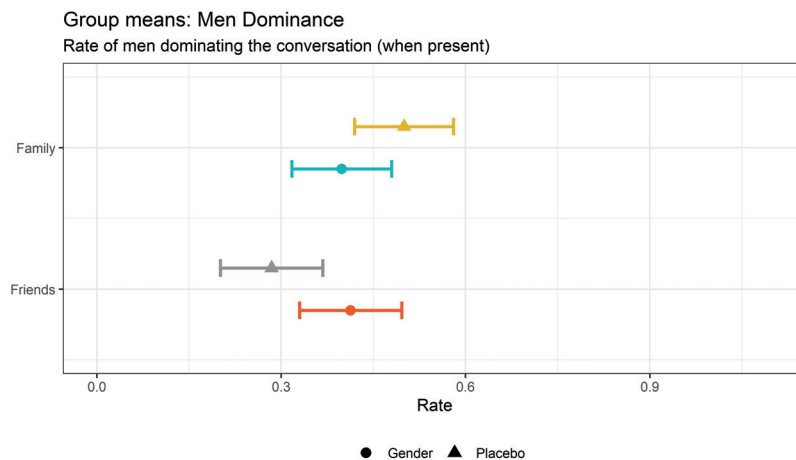
Figure 2.3: Male dominance of conversation by treatment

However, in the individual questions, we see that men, when present, were more likely to dominate the conversation during focus groups that discussed gender topics in both family and friend groups.Other research suggests that men are generally more likely to dominate the conversations(Karpowitz and Mendelberg, 2014). Figure 2.4 shows this pattern in Question 2. The group means plots for all questions can be found in the Appendix, as well as the full regression table.

Qualitatively, I found some evidence that men, when present, were critical thought leaders in the focus groups that they were participating in. Men, especially
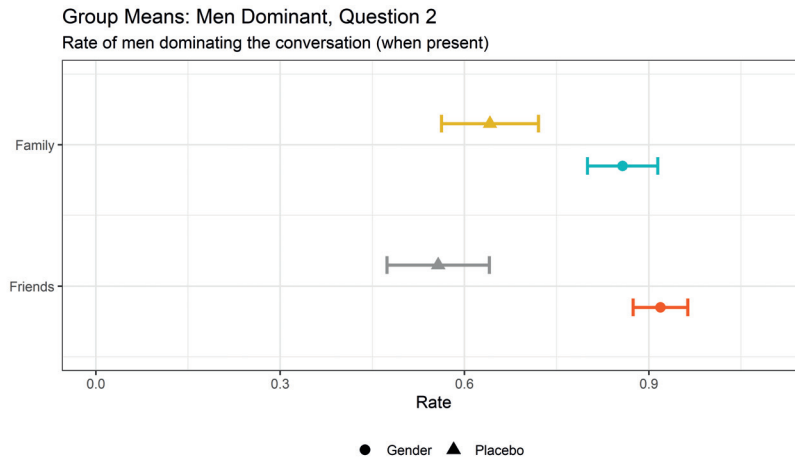
Figure 2.4: Male dominance of conversation by treatment: Question 2

older men, were often the first to speak after the question was posed to the group. In one particular case, in which the only man in the group was also somewhat older, he was the first to speak on each of the questions. While not tested here, one may think that the first comment, especially when coming from an elder, may set the tone for which norms will be expressed in the focus group. This, combined with the observed patterns that men in friend groups were more likely to dominate the conversation on gendered topics, speaks to some interesting dynamics that may have larger effects outside of the confines of this experiment.

As mentioned above, gender norms were often mentioned in the focus groups, even when they were not explicitly endorsed. While men were more likely to speak first, they were not the only ones bringing up these gender norms. Women were just as likely to mention these expectations, as well as mention the issues around FGM in the placebo condition. Even when men were the ones to mention these existing norms, they, as well as their female interlocutors, were likely to mention pro-participation stances later in the conversation.

In one case, however, a male participant, age 32, mentioned more than once that the Special Seats for women were, in his mind, not a good thing. Not only did he think the institution was a waste of money, but he also explicitly said that it would be better if all of the legislators were men. Despite all of this, he still mentions that if life went well for a young girl, when she grew up, she would be a leader. All of this is to say that the opinions held by the participants are not easily categorized as "pro-women's participation" or "anti-participation."

2.4.4 Explicit expressions of surprise

One hypothesis for why I find the stated effects is that the subjects are learning more about what their friends think about the topics. If that were the case, we might suspect that some participants may verbally express surprise about what they were hearing from their friends. These expressions of surprise or other emotional reactions to comments from other focus group participants could also be an indicator of the salience of the information. If we see more comments that are reacting to the opinions of others, this would me that not only are the subjects learning, but that they find the information interesting and are paying attention to the comments of others. In a low-salience situation, I would not expect to see much reaction to the opinions of other participants at all.

However, I found no such overt expressions of surprise in my reading of the focus group transcripts in either friend or family groups. The absence of these expressions in the notes may not mean that the subjects did not learn anything about the opinions of those in this primary social groups, nor that the information was not salient. As has been mentioned before, these are rough transcripts taken by the focus group facilitator as they were conducting the focus group. The facilitator could not write down every word, and was instead focusing on recording substantive comments on the topic itself. Even if a subject did express verbal surprise at an opinion, the facilitator would not have been taking recorded of that type of off-handed comment. The only ancillary comments that were recorded by the facilitator, where comments such as "hyvio ndyvio," (*even so*), which is a way to agree with what has been said while abdicating the obligation to add anything further.

2.5 Observable Implications of the Mechanisms

In the theory section above, I proposed several possible explanations for the observed effect on the behavior of the subjects. I now seek to find evidence for these theories from the additional data collected in the experiment. Because the data used here is not necessarily experimental, I offer these as suggestive results, which are potential avenues for further research.

2.5.1 Exposure to new information

The content mechanism hypothesizes that it is a difference in content of the conversation that changes the transmission of norms. It suggests that it is the norms expressed that change the perceived norms. Understanding this mechanism is a matter of knowing the content of the focus groups, including what norms are signalled in the conversation.

As has been described above, the treatment and control conditions were quite distinct. Conversations in the gendered treatment did in fact talk about gender norms, while those in the control condition largely avoided topics about gender politics and culture. Both groups almost universally expressed positive opinions about women's expanded participation in civil society. While there were a small handful of comments that were explicitly anti-women's civic participation, these comments were in both friend and family groups, and the number of observations were few enough that it was not clear that either type of group was more likely to include such comments.

From the random sample of 104 focus groups transcripts, I was unable to qualitatively detect any clear patterns about the content of the groups that clearly distinguished the content of the friend groups from those of the family group.

As comments that expressed opinions that are against women's participation in politics were rare, the differences between the norms that were expressed would have to be more subtle. There are many ways that more subtle signals can be communicated within social groups, especially when the group members are intimately familiar with each other, as is the case in these focus groups. The nuances are difficult to quantify, especially for a researcher who is unfamiliar with the context. However, the notes that were taken during the focus groups can perhaps provide some insights into patterns within the groups.

Using these transcripts, I created word clouds that help visualize the words that were used in each treatment group. These visuals present the words used in the transcripts, with the size of the words being relative to how often they were mentioned. The total unique words in each type of focus group was relatively small. The conversations in each treatment type produced less than 200 unique words in English (with the removal of conjunctions, articles, and prepositions.) My own review of the translations find them reliable, and, because the focus groups are relatively short, is it not surprising that unique words were limited. The notes were translated in a proscribed, random order so that any changes in translation skills across time would be evenly distributed throughout the treatment conditions.

Whereas the focus of this section is to understand why family groups do not produce the same effects as friend groups when discussing gendered politics, I present here, Figures 2.5 and 2.6, only the two gendered conversations. The word cloud from the respective placebo conversations can be found in the appendix.

Figure 2.5: Wordcloud: Family gendered topics group

Figure 2.6: Wordcloud: Friends gendered topic group

In both family and friend groups, we see that the words "women" and "men" were mentioned quite often. Given the discussion, this makes sense. In the family group, we see the word "temptation" more often than in the friend group. "Temptation" is mentioned in the context of asking about challenges which keep a young girl from finishing school. It is more or less a euphemism for sex, and refers to the fear that girls will become victim to or distracted by young men, leading to situations that force her to leave school (these are not enumerated, only alluded to.) Temptation is mentioned in the friend group, but much less often.

It is also clear that words like "help" and "good" are used often in both groups, suggesting that women have the potential to be a benefit to society. The friend group also mentions the word "proud" often; however the word is not used at all in the family group. "Proud" may indicate that the norm of an internal satisfaction is more clearly communicated in friend groups than in family groups. The friend groups also mention the word "will" often, while the family group use the word "may." This suggests an emphasis on self-efficacy (Bandura, 2010) in the friend group. If the positive outcome is signaled to be more likely to yield results, then women may be more likely to perform the behavior. While each of these words have at least one homophone, these visuals suggest that the friend groups are expressing more certainty about the norms that they are expressing, and the family groups are softening the views slightly.

These few qualitative observations are certainly not conclusive, but suggest that while both groups are expressing that women can or should participate in politics, friend groups may be doing so with more assurance, and more focus on intrinsic, not just extrinsic, benefits to such changes.

2.5.2 Exposure

One possible explanation for the change in behavior is that the subjects were exposed to new information and that this exposure was heterogeneous between family and friend groups. In other words, something about the dynamics between friend or family groups cause more novel information to be transmitted in the friend group than in the family group. There are three ways that this may take place.

Previously known vs novel information

It may be that we do not see a change in behavior in family groups because the opinion of the family was already known. A subject may already know their family members about these topics because they are more likely to have already discussed them previously. Home life may give subjects more opportunities to discuss a variety of topics, including gender norms and women's political participation. As a result, during the focus group, the subject does not learn any new information, so

her behavior will already be coherent with the information she would have learned.

If this is the case, I would expect that both the placebo and treatment conditions in the family to not only be approximately equal to each other but also approximately equal to the treatment condition of the friend group. The results that we see using the randomly selected primary subjects are suggestive of this, but the confidence intervals are too large to be confirm.

Because the other subjects in the experiment were not randomly selected, I did not include them in the original analysis. The lack of random selection meant that there was no evidence those subjects would be uncorrelated to the outcome measures. Relaxing this scrutiny slightly, I also have no theoretical reason to believe that the selection of the other focus group participants would be correlated with the outcomes variables. Therefore, for the purpose of seeking to understand this pattern, I will include them in the analysis.

Figure 2.7 below shows the group mean of the primary subjects, as was shown in the previous chapter. Figure2.8 shows the same group means, but including the full set of focus group participants, not just the subjects who were randomly selected for participation.
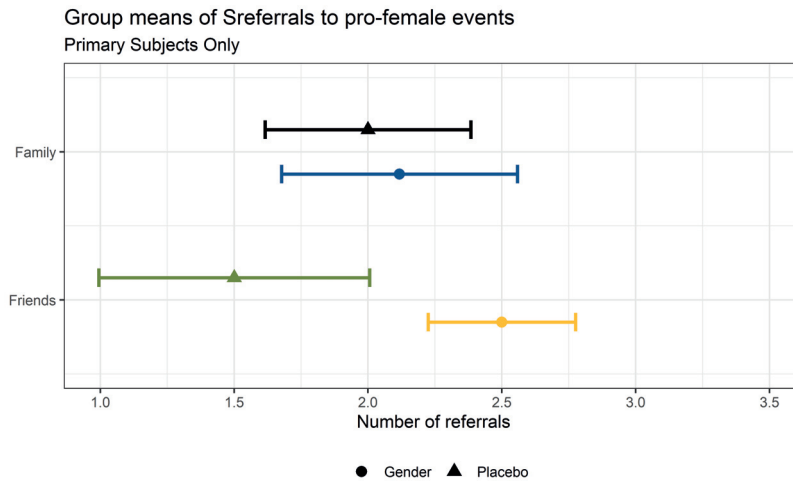
Figure 2.7: Primary subjects behavioral outcome by treatment group
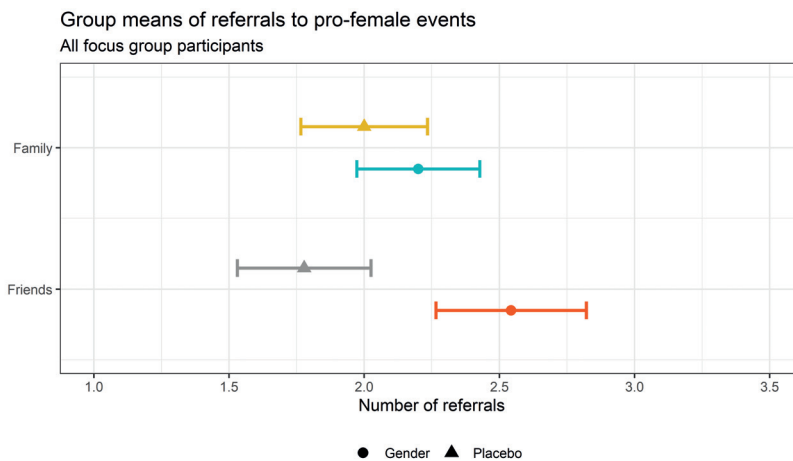
Figure 2.8: All participants behavioral outcome by treatment group

The initial result–that conversations with friends about gender topics yield more referrals to the pro-participation events–is strengthened by the inclusion of the rest of the focus group participants. However, if the opinion of family members was already known, we would expect both family conversation groups to have outcomes approximately equal to the outcomes of the gendered conversations among friends. While the outcome for the gendered conversation among family is similar to the outcome for gendered conversation among friends, the placebo conversation among family members does not have similar outcomes to the gendered/friend conversation. The finding is counter to the hypothesis that subjects were not updating their mental model of the norms, the outcomes for both conversation types in family groups would match the gendered + friend conversation groups.

Egalitarian participation

It may also be that conversations among friends have different dynamics than conversations among family members. My theory suggests that when women lean that others agree with the idea that women should participate in politics and contribute to civil society they will be more willing to participate themselves. If, however, during a focus group, one person tends to dominate the conversation, then the subject may not actually learn that the rest of the group has these beliefs. Therefore, if focus groups among friends are more egalitarian, and have more full participation, then we may see a larger shift in the perceived norm. If conversations among family members are more likely to be dominated by one person, then the perception of the norm may not be sufficiently changed to cause a change in the behavior.

To test this, I create two measures of equal participation among groups using information from the transcript data. First, from data collected by the focus group facilitators, I have a measure of how many times each participant spoke during the focus group. If the variance of this measure within each focus group is small, then I can say that the focus group was more egalitarian.

Figure 2.9 shows the variance of the number of comments for each participant within the focus group as a whole (not taking each question as an observation). Here, it is clear that the variance was much higher in focus groups that were discussing gender topics than those with placebo questions. It is also interesting that in the placebo treatment, friends actually had more variance between the comments than family groups. The larger variance suggests that in the placebo treatment, family groups were actually *more* egalitarian than friend groups. The more egalitarian discussion in family groups mirrors the finding above that men dominated the conversation more in friend groups as well. However, in relation to our question, I detect no discernible difference in the variance of the number of comments between friend and family focus groups in the gendered-topic conversations. The measure of equal participation does not provide evidence that friend groups were more egalitarian, allowing for a better transmission of gender norms.

A second measure of participation is how often all members of the focus group commented on a particular topic. Again, using data from the focus group transcripts, I can record whether or not every member of the focus group participated in the discussion, both for the focus group in its entirety and for each question. When all members of the focus group participate in the discussion, I can say that the conversation is more egalitarian than cases in which some participants did not comment. When the focus groups consisted of no more than four people, this is especially true.

I created a binary indicator of whether or not all participants commented on each question. I then changed that into a measure of the share of questions in each
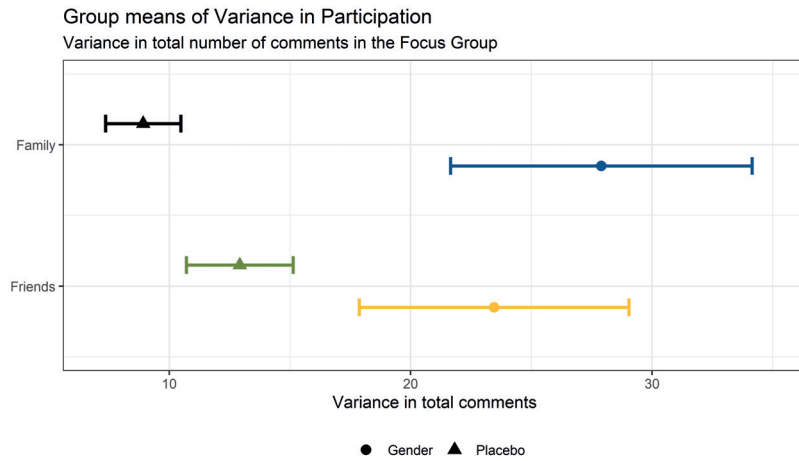
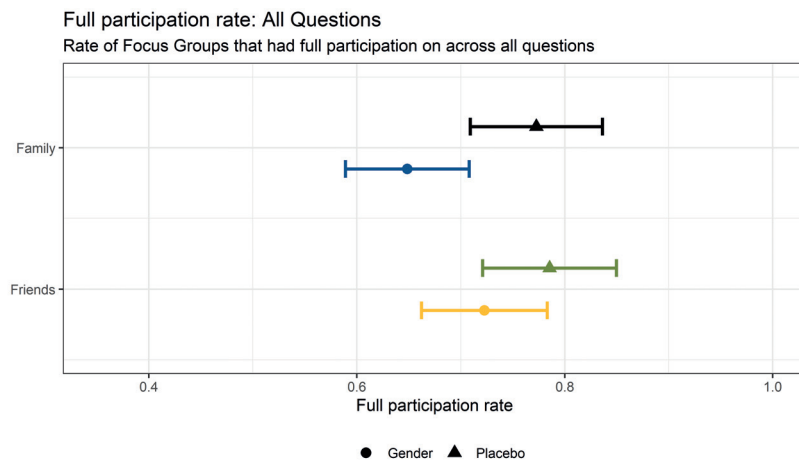Figure 2.9: Variance of participation for all participants

Figure 2.10: Rate of full participation in each question

focus group that had full participation[5].

Figure 2.10 (above) shows a lack of significant difference in the participation rate between family and friend groups. One minor difference is that in the family groups, the questions in the focus groups were more likely to have full participation in the placebo condition than in the gendered topic conversation, suggesting that in the gendered topic, some people are less likely to speak. Again, the family members were more likely to have egalitarian participation in the placebo treatment than in the gendered conversation. However, this is difficult to interpret compared to the friend groups, though the fact that egalitarian participation goes down when talking about gendered topics in family groups may suggest that the norms would not be as easily transmitted in these situations.

2.5.3 Self-Coherence

The final theory is that the mechanism at work it not about receiving now norms, but about the subjects expressing herself. It might be that the change in behavior comes from the priming effect of having spoken in favor of those norms. If we observed that the subjects in friend groups– perhaps because on internal groups dynamics– have a higher likelihood of speaking more, those subject may be priming themselves more than those in family groups.

I used two similar measures to test this hypothesis. First, using the transcript data, I measured the number of comments per question that were given by the primary subject. Second, I show that the share of the questions that the primary subject commented on (Number of questions that the subject commented on, divided by the number of questions in the focus group.).

In both cases, the group means show no significant difference across any of the four treatment conditions. Figures 2.11 and 2.12 below show the group means for these metrics. These measures do not provide evidence that the reason for the change in behavior was due to priming effects instead of a communication of previously concealed norms.

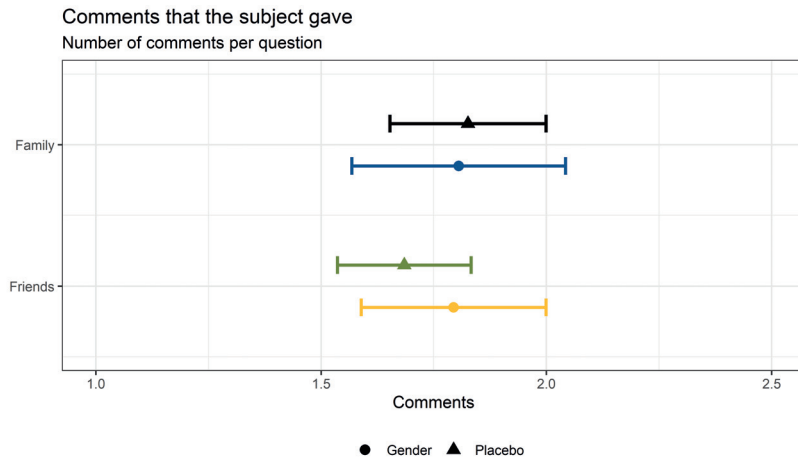---

[5]The gendered topic condition had seven.

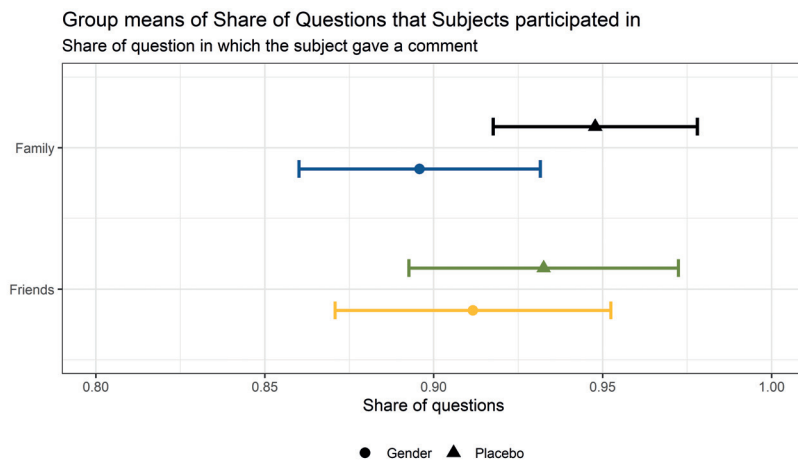Figure 2.11: Comments per question: Primary subjects

Figure 2.12: Rate of participates for primary subject

2.6 Conclusion

My focus group found heterogeneous effects of explicitly discussing gender in politics across primary social groups. Talking about issues of gender and politics lead women to refer more people to participants in pro-women's empowerment events; discussions of those topics with family did not have the same effect. In this paper, I sought to better understand the reason for these heterogeneous effects. I posed three different mechanisms by which we may see these effects: Novel information, self-coherence or priming, and information salience. After reviewing the content and treatments involved in the focus group, I reviewed evidence from additional data collected in the experiment to try to support or refute those possible mechanisms. Table 2.3 summarizes the findings of each of these hypotheses and their tests.

Hypothesis	Empirical Test	Results
Information Received:	–	–
Opinions previously known	Behavioral outcomes for all participants	*Unsupportive*
Egalitarian participation	- Variance of comments	*Unsupportive*
	- Measures of full participation on questions	*Unsupported*
Content of discussion	- Word clouds	*Supportive*
	- Key word analysis	??
Self-Coherence	- Comments per question from subject	*Unsupportive*
	- Binary participation in question from subject	*Unsupportive*
Salience of Information	*Needs additional data*	–

Table 2.3: Table of possible mechanisms and tests

The metrics used to test the priming mechanism showed no difference in the amount of interaction between the four treatment conditions. Because interaction with the focus group by the subject did not vary across the treatment groups, I find no evidence that this mechanism could be the reason that we see effects in the friend group but not in the family group.

The novel information mechanism has three possible ways that that friend groups may differ from family groups. Family opinions may already be known, leading to the no update in gender norms; friend groups may be more pro-women's participation, leading the subject to learn about those opinions; or friend group discussion may be more egalitarian, leading to learning more information from the

29

group.

Based on the analysis I had of these hypotheses, there seems to be weak, but not exclusively null results. While the opinions expressed were almost universally pro-participation in both family and friend groups, a visual analysis of the text reveals that there may be some patterns in the way that the topics were discussed that could change the reaction from the subject. Likewise, while there was no difference between friend and family groups in measures of egalitarian participation, family groups were *more* egalitarian than friend groups in placebo conversation than in gendered conversation. Even taken as a whole, this evidence does not provide compelling support that friend groups provided more novel information than family groups about otherwise concealed norms about gender in Tanzania.

The last hypothesis was that the information about the beliefs of friends was more salient to the subject than the same information about their family members. It may be that subjects simply care more about what their friends think about these topics. The data that was collected in this experiment does not offer any reliable window into whether or not the hypothesis is true. Further research is needed to discern whether this mechanism has any weight.

This research could easily take the form of a survey experiment, in which participants are surveyed about their opinion and then the results of those surveys are presented (on average, without identifying information) to a subject. Similar outcomes of political participation could then be measured and then compared across conditions of whether the subject viewed the opinions of friends or family members.

Further versions of the focus group experiment could also vary the way in which the focus group is carried out, requiring all participants to speak on each topic or not, in order to have a more reliable measure or egalitarian participation, isolating from those effects from other variables.

While the rich data available provides valuable insights into the dynamics within these focus groups, further research is necessary to make conclusive casual inference.

Chapter 3

Do Western Researchers Affect How We Observe Gender Dynamics in the Field?

An Experimental Study

Abstract

Do research subjects in the field respond differently to questions in the presence of a foreign researcher? In particular, do white researchers affect the dynamics of focus groups that are conducted in rural Africa? In a field experiment in Tanzania, I randomized the presence of a foreign female as part of a field team conducting focus groups in 80 study villages (as opposed to an all-local team with the same gender composition). The design allowed me to estimate the average effects of the foreign researchers' presence by comparing relative levels of engagement with the focus group questions given across the focus groups. I found that focus groups with a white researcher present had longer discussions, particularly on questions about gender and culture. These questions are ones in which the gender and foreignness of the researcher would be more salient. In some cases, these results suggest that data collected in the presence of a conspicuously foreign researcher might be more thorough than in their absence. However, the information may also be pandering to the perceived values of the researcher, making some types of data less reliable [1].

3.1 Introduction

Researchers conducting field work in international contexts often worry about the effect that their presence might have on their subjects while working in the field. The concern is especially prominent when white researchers are working in developing countries in which white people are a minority and therefore a more conspicuous presence. Other authors have discussed the implications of a conspicuous foreign presence on field research design (Kapiszewski et al., 2015; Paluck, 2010; Udry, 2003).

In this paper, I focus specifically on the effects of the presence of a foreigner on the research process. Researchers who are foreign to the research site are often conflicted about how much time to spend in the field, due to two competing concerns. On the one hand, they would like to provide high levels of oversight of their projects; on the other hand, they do not want to distort responses through their presence. Recent research by Cilliers, Dube and Siddiqi (2015) finds that the presence of a white foreigner does indeed change how respondents play laboratory games. My research contributes to understanding how foreigner presence in a field team changes the responses provided in both qualitative and quantitative research.

At the moment, much of the advice offered in the existing literature is impressionistic rather than empirically supported. While some literature exploreturs the effects of interviewer identity, gender, and proximity on survey responses (Adida et al., 2016; Blaydes and Gillum, 2013; Campbell, 1981), many aspects of field research design remain under-studied. Most relevant for my project, Weinreb (2006) finds that locally known interviewers achieve higher response rates and collect more reliable data than unknown (stranger) interviewers.

Many concerns about the effect of foreign researchers on the responses of research subjects stem from the concept of social desirability bias. While focus groups participants may feel at ease to communicate among themselves when no foreigners are present, a white, female researcher may project expectations that the respondents feel they need to meet. The most visible parts of her identity (being white and being female) will trigger the greatest sensitivity bias. Research participants may worry that the researcher will think they are un-progressive or "backward" if they answer questions in a way that does not fit *their* perception of *her* views. As a result, the researcher may withhold resources or finances that she may otherwise have been willing and able to impart. A white researcher may also be perceived as particularly naïve, not knowing much about the context of the research. The perception of naivete can lead respondents to explain cultural specifics as greater length. The presence of a female researcher may imply an expectation of wanting particularly to hear about women in power. The respondents, feeling that this is what the conspicuous researcher wants or needs, would adjust their participation accordingly.

In this experiment, I find evidence that the presence of the white female re-

searcher does affect the behavior of focus group participants. Specifically, I show that in the presence of a white, female researcher at focus groups in Tanzania, participants speak longer and more thoroughly than in a control group in the presence of a similar Tanzanian researcher.

I conducted focus groups in 79 villages in rural Tanzania in the summer of 2018. In half of those villages I, a white female from the United States, was present as a note-taker along with two Tanzanian enumerators. The other enumerators assisted in running the focus group: One facilitated the discussion, while the other wrote a rough transcript. In the other half of the focus groups, a female Tanzanian researcher around the same age took my place. To further understand the way that white researchers affect field research, I also varied the projected level of my cultural sensitivity. In half of the treatment conditions, I spoke the common language of Swahili, and wore a skirt demonstrating cultural competency. In the other half, the I spoke only in English and wore trousers, which, while not offensive, signaled my naivete to the context.

The presence of a white person in Tanzania, while not uncommon, is still conspicuous. The villages in this study were all rural, and some were quite remote. In some cases, I was the only white person participants could remember visiting the village. The strength of the manipulation can be further illustrated by the content of some of the focus groups. In one case, in which I was speaking only in English, a participant was discussing the benefits of having a white people in the village, citing white people's willingness to pay prices much higher than the market value for common items, mostly out of ignorance. At which point the respondent stopped short, motioned to me, and asked the Tanzanian facilitator, "She doesn't understand Swahili, right?"

The experimental results indicate that in the presence of a white, female foreigner, the focus group discussion lasted on average 45 seconds longer than when the role was being filled by a Tanzanian woman. These results are centered around on a particular subset of questions, namely those that asked about gender and those which asked about specifics of Tanzanian culture. These questions are ones for which the presence of a white female researcher may be most salient. Furthermore, I find that when I was present, male participants in focus group conversations were 17% less likely to dominate the conversation (se = .06), and participants provided more thorough information about village power structure.

The experiment extends existing research on the effect of foreign researcher presence on field work in two ways. First, it considers the effects of foreigner presence as part of a research team on qualitative data collection. Second, it provides quantitative evidence for the existence of social desirability bias in a group setting, as the focus group, not the individual, is the unit of analysis.

The paper will proceed as follows: First, I will address the existing literature on the subject, including social desirability bias, and focus group dynamics, from which I will build a theory. I will discuss how and in what ways I expect that the

presence of a white female researcher may affect the dynamics of a focus group in Tanzania. I will then explain the design I used to understand these effects and the types of data collected. I will discuss my findings that respondents spoke longer when I was present and were more thorough in their discussions of village dynamics. I will also show that men were less likely to dominate the conversation. I will then conclude with the implications of these results.

3.2 Theory

I base this research on two ideas: First, the content of focus groups is affected by the people facilitating the discussion. And second, that social desirability bias can affect the way research subjects respond to questions (Blair et al., 2020). In combination, this would suggest that focus groups respond differently to researchers depending on how the respondents perceive the preferences of the researcher. For our purposes, the change in the foreignness of the researcher signals either a change in the preferences of the researcher or a change in the relative cost of not complying with perceived social desirability. While the gender of the researcher was held constant, one can imagine that there is interaction between the whiteness and femaleness of the researcher which makes the gender of a white researcher salient (Muriaas et al., 2019).

Previous research has been conducted about how research participants respond differently to the various types of enumerators. Much of this research has focused on survey respondents. Race in the American South (Weinreb, 2006), religiosity in Egypt (Blaydes and Gillum, 2013), gender (Catania et al., 1996; Huddy et al., 1997; Kane and Macaulay, 1993) and ethnicity in Africa (Adida et al., 2016) have all been found to have an effect on how respondents answer survey questions. Several other excellent studies, (Feldman et al., 1951; Cotter et al., 1982; Davis, 1997; Hatchett and Schuman, 1975) all randomly controlled, have found that the race of the interviewer in a US context affects the responses of survey participants.

Somewhat less research has been done to understand the effect of conspicuously foreign (white, in this context) researchers on other types of data collection. Cillers, Dube, and Siddiqui (2015) find that the presence of a white male researcher during dictator games in Sierra Leone increase the player contributions in the game. Fern (1982) finds that the content of focus groups is similar with and without any facilitator at all, suggesting that the Hawthorne effect is null in these cases. Both observational (Hollander, 2004) and experimental (Humphreys et al., 2006) research has shown the that the dynamics within focus groups change significantly depending on the authority figures within or facilitating the focus group.

It is clear that participants respond differently depending on the identity of the person conducting the research. In this design, I never acted as a facilitator for the focus groups, in order to isolate the effects of foreigner presence from facilitator

effects. In this experiment, I was largely a passive observer during the focus groups, in most cases only speaking briefly to introduce myself in either English or Swahili. I only took notes on a tablet, rarely speaking otherwise. The Tanzanian researcher in the control treatment did likewise, though entirely in Swahili. The design of this experiment allowed me to better isolate the change in the focus group as a result of the reaction of the participants to the researcher.

Social desirability bias is a result of "the desire of respondents to avoid embarrassment and project a favorable image to others"(Fisher, 1993, 303). In this way, respondents attempt to improve their image based on how they believe others think. Goffman (1959) suggests that people have a mental conception of how others perceive them and a desire to change that perception by appealing to the values that they believe the other person holds.

A more precise conceptualization of this bias is offered by Blair, Coppock, and Moor (2020). Social desirability bias is one manifestation of sensitivity bias. The authors suggest considering four questions to determine when sensitivity bias could be a concern: 1) Is there a social referent who respondents have in mind when answering? 2) Do respondents believe the social referent can infer their answers exactly or approximately? 3) Do respondents perceive that the social referent prefers a particular answer to the question? 4) Do respondents believe they (or others) will suffer costs if that preferred response is not provided? (ibid).The randomization of the presence of a white researcher or an additional Tanzanian researcher may have different effects on these aspects.

Regarding the first element, all of the members of the research team are potential referents. The conspicuousness of myself as a white person in a village will mean that the respondents may consider me a more salient referent. However, it may also be that because I am an outsider, my opinions will be less important to the participants, so I am less of a valid social referent than a fellow Tanzanian. Varying the identity of the referent can identify my relative effects on the focus group.

In the second element, the experiment randomizes the degree to which the the answers can be inferred by the referent. While the Tanzanian researchers will clearly be able to understand the responses, it is less obvious that I would be able to understand the content of the focus group. If I do not signal that I speak Swahili, I weaken my perceived ability to infer their answers. As I will explain below, this element is randomized in the experiment: I signal my understanding of Swahili in half of the cases in which I am present.

The third element considers whether the respondents perceived that the referent has a preference for a particular response. The variety of topics in the focus groups allow observation of the element on a number of responses. The perceptions of my preferences as a westerner may be different than the perceptions of the preferences of my Tanzanian counterpart.

My foreignness my signal a particular set of values and preferences on the topics that are discussed. My Tanzanian counterpart may also be signaling preferences,

though perhaps different ones. While it is difficult to know what is assumed by other people, I may assume that I would signal that I would prefer opinions that are more progressive and more permissive of women's economic and social empowerment. What my Tanzanian counterparts signal is perhaps more complex. Though several female researchers acted as control conditions, all of them were from urban area of Tanzania, which may be signaled to the participants, perhaps only in subtle ways. At the very least, my counterparts were women who were working as researchers and traveling for such. Because of the heterogeneous culture in Tanzania, participants may assume that the Tanzanaian researcher has similar preference to them, or they may think that she also has preferences for more progressive views. However, by varying the identity of the researcher, the experiment can observe the relative effect of my presence in this regard.

The last element, the belief that non-conformity with expectations will be costly, may be the most salient in this context. As I will discuss below, white foreigners are often associated with financial resources. The presence of a white foreigner will likely signal to the respondents that answering in accordance with the perceived preferences of the referent may have financial benefits to the respondent (or in answering contrary, financial loses.) However, many NGOs are run by Tanzanians, and some elements of the situation may signal similar potential gains, even when I am not present. All research teams will be using tablets for some data collection, and, per ethics standards, all teams tell the participants that they are working with researchers from NGOs and academic institutions. Both of these elements may signal access to funds. My own conspicuous foreignness might not be a significant additional signal of potential gains, though the exogenous variation of the experiment will enable the observation of these causal effects.

3.2.1 Tanzanian Context

Tanzania has a long history of the presence of white people. From early missionaries to white colonialists, to contemporary business people, the presence of white people is not uncommon, particularly in neighborhoods near the city center. Tourism from white-majority countries is a large contributor to the economy (Wade et al., 2001), as are foreign NGOs that work out of the large cities (Kiondo and Mtatifikolo, 1999; Shivji, 2004; Ngowi and Jani, 2018). Still, the presence of a white person is conspicuous. Even in cities, where white people are seen regularly, it is common that young children (and some adults) will call out to white people when they see them on the street, without any other purpose than to comment that the person is white ('Mzungu'). White presence is not just noticeable, but often overtly noted.

In more rural, and especially remote areas, the presence of white foreigner, would be rarer and therefore more even more noticeable (Manley et al., 2016). In these cases, during the research, it was not uncommon that small crowds of young children would gather to observe the 'mzungu.' My femaleness only amplifies

the effect, as a white female traveling in Tanzania (without other white people) is additionally uncommon. In some cases, the focus group participants insisted that a white person had not visited that particular village, at least in living memory.

It is not mere novelty that causes a change in response from participants. In fact, it is the racial dynamic likely fuels the effects that I find in this research. Because white people in Tanzania often come with disposable income to spend, or financial backing for projects, whiteness is associated with comparative wealth (Manley et al., 2016). A white researcher, in contrast to an equivalent Tanzanian researcher, would signal to respondents that there is an increased financial incentive in answering in a way that they think would appeal to the researcher. In this case, no such financial incentive was offered or given, but sensitivity bias can be effective even if it is subconscious.

Furthermore, a white woman in Tanzania would likely be perceived as particularly uninhibited, signaling that she would may be interested in the empowerment of other women and would want to know about women in particular. Many NGOs and projects in Tanzania visibly seek to involve women in politics and education, bringing women's issues to relevance in the mind of the general public(Osirim, 2001).

Considering these elements together, I would expect that focus group respondents, because of sensitivity bias due to the change in referent, will respond differently to questions when I am present than when a Tanzanian (black) female researcher takes my place. I, as the referent, would be seen as naïve, therefore questions that are specific to Tanzanian culture would illicit a more thorough response and take more time. Questions that are perceived as being of particular interest to me because I am white and female will also illicit a more thorough response, so I expect that questions about gender will have a longer discussion. Male participants may feel that they are less welcome to dominate the conversation, and female participants may feel more willing to respond, so I will see a decreased incidence of men dominating the conversation as a whole and on each question. Participants will also respond to my perceived naivete by being more thorough in their discussion of village power dynamics. They will also intuit that I would like to see women ranked more highly regarding power dynamics and adjust their responses accordingly.

Table 3.1 summarizes the hypotheses and the variables with which will be used to measure the result.

In the next section I will go into further detail about how these data were obtained and the details of each measurement.

Hypothesis	Measure
1. Participants will speak more during the research groups with a white female researcher.	Total length of the focus group discussion.
2. Participants will speak more about topics that are most salient to the observable identity of the white female researcher, specifically her foreignness and her femaleness.	*Foreignness:* Length of discussion of culturally-specific topic of ancestor worship. *Femaleness:* Length of discussion of women in leadership positions in the village.
3. Female participants will be more likely to speak equally with male participants, as both male and female participants will see this as the desire of the white researcher.	Binary measure of whether or not men dominated each question and the focus group as a whole.
4. Participants will be more thorough in their responses to better inform a naïve listener.	Total number of people ranked in the opinion-ranking exercise.
5. Women will be reported to be more important in village decision-making, per sensitivity bias.	Ranking of old women and young women in opinion-ranking exercise.

Table 3.1: Outcomes and measures of interest.

3.3 Design

To understand the effect of white researchers on focus groups conducted in Tanzania, I conducted a randomized control trial in 79 villages across Tanzania, in six regions, shown in Figure 3.1
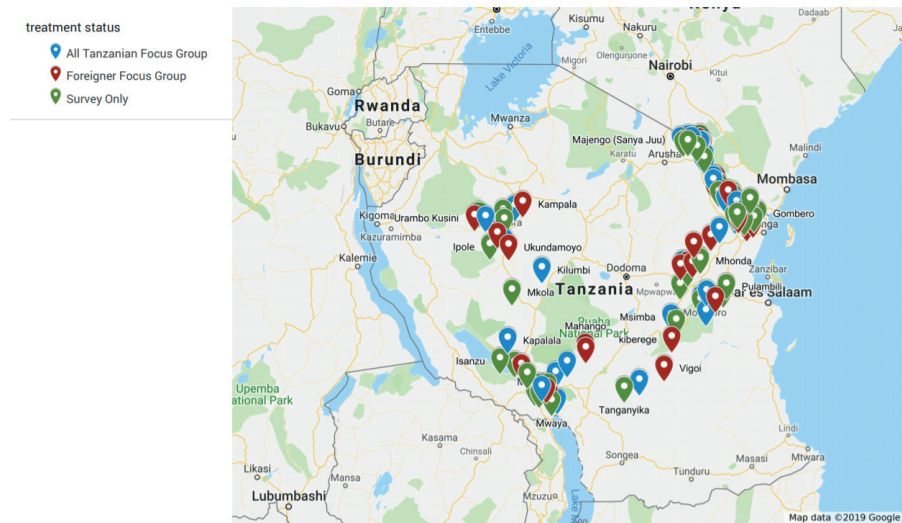

Figure 3.1: Focus Group location by treatment

The villages were all rural and some were remote, accessible only via motorcycle or by foot. Some villages, while small, were located on main highways. Focus group participants were recruited two days before the focus group by a separate team of recruiters, who were blind to the treatment condition. Focus groups were made up of up to ten respondents, depending on recruitment compliance, which varied. These respondents were selected via random walk from an approximate census of houses in the village. In some cases, a sub village unit was used if the village was too large to conduct the random walk in a single day. Respondents were compensated for their time at the end of the focus group.

I randomly exposed villages to one of three treatment arms, outlined below. The treatment was randomly assigned by village, blocking on region. The focus group discussion was led by a facilitator, who was always Tanzanian. A second researcher took substantive notes on the content of the focus group, and this role was also always filled by a Tanzanian researcher. A second note taker took notes on the dynamics of the focus group, including whether or not men dominated the conversation. In the treatment group, this role was filled by myself. In the control group, this role was filled by a Tanzanian, female researcher, creating randomized variation on the 3rd and 4th elements of the social desirability framework offered by Blair, Coppock, and Moor (2020).

Among cases in which the researcher was a white female, in half of the focus

groups, I signaled cultural competency by only speaking Swahili and wearing a long skirt. In the other condition (19 villages) I spoke only in English and wore trousers. While not offensive, a woman wearing pants and only speaking in a foreign language suggests a minor disconnect with the culture in the area, creating randomized variation on the 2nd element of the framework. The lack of symmetry (19 vs 20 villages) is due to logistical challenges in the field that resulted in one village being dropped from the sample. I pooled the results of the cultural competency variable for purposes of power. The regression table that includes this condition can be found in the Appendix.

A summary of these three treatment arms is found in Table 3.3, here.

Tanzanian Female Researcher	White Female Researcher
40 Villages/Focus Groups	*Cultural competency:* 20 Villages/Focus Groups *Cultural naivete:* 19 Villages/Focus Groups

The length of discussion for each question was recorded via a timer on a note-taking tablet, as was the total length of the focus group. Additionally, at the end of each question, enumerators recorded whether or not the conversation was dominated by a man and whether or not it was dominated by an older person. Similarly, at the end of the entire focus group, those questions were asked about the focus group in its entirety.

Each focus group was asked eleven open-ended questions about various topics of interest. The full list of questions can be found in the Appendix. Ten of the questions called for unstructured conversation of the topic, with some minimal prompting from the facilitator. The exception is Question 7, which asks the group to participate in a ranking task in which they were given eight pieces of paper, each with a category of person that would be in the village (clan leaders, young women, old men, etc.) Participants were then asked to collectively rank those groups by ordering the papers based on "whose opinions carry the most influence in ... decisions in this community." Once the papers were thus ranked, the rank order was recorded. The focus group was then given the opportunity to add to the list of groups that were provided by writing omitted groups on separate pieces of paper. The new groups were then sorted with the original ranking by influence. The total number of groups, as well as the ranking (and time elapsed) was recorded. Each focus group submitted one original ranking and the possibly a second, appended ranking.

Topics of the questions are listed below, with their regression table references included parenthetically:

1. Historic interactions with foreigners (History of foreigners).

2. Harm or benefit of those historic interactions (Harm/Benefit of foreigners).

3. Recent interactions with foreigners (two questions).

4. Harm or benefit of those recent interactions (Harm/Benefit of recent).

5. Modern methods of resolving disputes in the community (Land Disputes).

6. Historic methods of resolving land disputes in the community (Historical land disputes).

7. Ranking of influence of community groups (Opinion ranking)

8. Community cooperation for provision of public goods (Public goods projects).

9. *Cultural traditions surrounding honoring dead ancestors* (Celebrating ancestors).

10. *Women's role in village leadership.* (Women in leadership).

11. Preferences for priorities of elected representatives (MP type preference).

The questions about honoring dead ancestors and women's role in leadership are of most interest here because they draw attention to two salient aspects of the white female researcher: My naivete due to foreignness, and my femaleness, which interacts with the perception my whiteness. I will also consider the answers to the ranking question, including the relative rank given to women in the village, and the total number of groups listed, as a measure of thoroughness.

All the research activities were implemented by IPA-Tanzania in June-August of 2018. The author served as the treatment condition in all cases. The research assistants filling the other roles in the focus group were rotated regularly. Both male and female research assistants served as the facilitator and primary note-taker. The secondary note-taker role was filled only by myself or a (black) Tanzanian female counterpart. All of the research team was approximately the same age (25-32) and all had post-secondary education[2].

The unit of analysis is at the focus group level. In some cases, I analyze the focus group discussion as a whole, but I also explore the results from individual questions. The responses of the individual respondents are not included in the analysis. I am interested in the dynamics of the focus group as a whole, not the reactions of participants *per se.*

---

[2]I also compared the my personality traits to those of my Tanzanian counter parts, as check. I sought to demonstrate not that I was similar in personality to the Tanzanian researchers, but to show that I was different from them in a way that was similar to ways in which average Tanzanians differ from average Americans. To do this, I administered the same measure of personality (IPIP Big-Five Factor Markers) to both myself and the Tanzanian counterparts. The Big-Five test has relevance in psychology and existing literature, which uses this particular test to compare average scores from many countries in the world, including the US and Tanzania (Schmitt et al., 2007). While my scores differed from those of the Tanzanian researchers, it did so in ways that were consistent with existing research (ibid.), showing that in this measure I am different from Tanzanian researchers in ways that are similar to average differences between Americans and Tanzanians (see Appendix).

3.4 Results

The purpose of this paper is to understand the effect of the presence of a white female researcher on the length and content of focus groups conducted in rural Tanzania. In the theory section above, I proposed five outcomes that I would expect to see based on the theory and previous literature (see Table 3.1, above). Using data collected on the length and quality of each focus group and questions in each focus group, I will discuss these outcomes in turn.

The framing of this paper also referenced the work of Blair, Coppock and Moor (2020) to understand when a researcher should be concerned with sensitivity bias. The presence of the researcher randomizes the first of their four elements (presence of a referent), allowing for the possibility of the other three. Randomizing the signals of understanding the language and other cultural competencies could show my relative effect on the second element, the ability of the referent to infer the answers of the respondents.

Because the latter of these arms split the sample, it was difficult to infer a causal relationship in most cases. The regression tables for the variation across that treatment can be found in the appendix. By pooling the results of these two groups, I can measure the effect of my presence (the first element of sensitivity bias) on the five outcomes of interest.

3.4.1 Outcome 1: Focus groups will be longer in the presence of a white female researcher

An OLS regression, clustering on the focus group, shows that that the average question length was about 6.5 minutes long (395 seconds, clustered *se* = 14.43) when I was not present. When I was present, the questions took on average 45 seconds longer, with a standard error of 22.5 (See Figure 3.2, below). Regarding outcome 1, wherein focus groups talk longer when I was present, I can reject the null hypothesis at conventional confidence levels.

There is no significant difference between the culturally competent and culturally naïve sub-treatments. Even when the focus group participants had no reason to believe that I understood the content of the focus group discussion, they still had longer discussions when I was there. A full regression table is found in the Appendix.

3.4.2 Outcome 2: Focus groups will speak longer, particularly on topics that are culturally relevant or address gender issues.

Figure 1 also shows my average effect on each question individually. While the average for all questions was significantly higher, I find that this effect is largely clustered around three questions. Although there is a marginal significance on
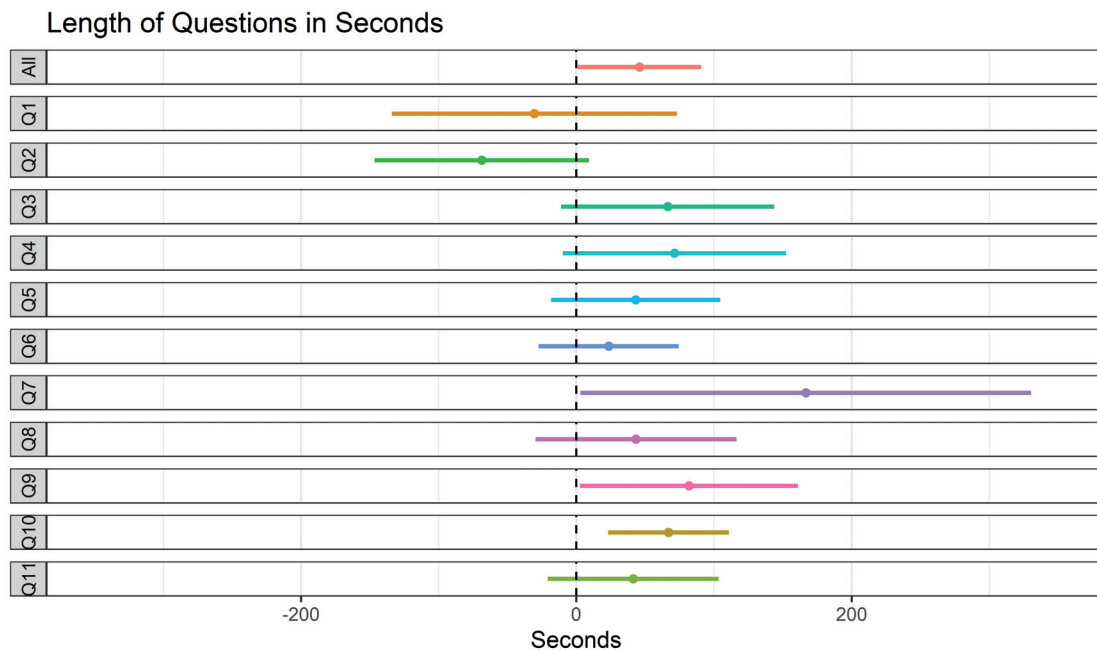
Figure 3.2: Difference in topic duration by question

some questions, the effect is mostly centered around questions 7, 9, and especially 10. Question 7, as discussed before, was a ranking exercise, which I discuss below. A full regression table is found in the Appendix.

As mentioned above, question 9 asks the participants about cultural practices surrounding the honoring or worshiping of ancestors. In this question the naivete of the listener would be particularly apparent. As expected, focus group discussion lasted significantly longer (81 seconds, standard error = 39.8) when I was present.

Question 10 asks about the role and challenges of female leaders in the village. When I was not present, discussion of this topic lasted on average about 5 minutes. When I was there, the discussion lasted over a full minute longer on average (67 seconds), with a standard error of 22 seconds. These comparisons can also be seen on Figure 3.2.

3.4.3   Outcome 3: Men are less likely to dominate the conversation with a white female researcher present.

Not only did participants speak longer when the foreigner was present, but the group dynamics of the focus groups changed somewhat as well. At the end of each question in the focus group, some dynamics were recorded about the discussion during the question. A binary, subjective measure was taken noting whether or not

men dominated the conversation (speaking approximately more than 75% of the time) on that question, and again at the end, concerning the focus group discussion as a whole.

Figure 3.3 shows that when I was present, men were reported to dominate the discussion significantly less, by 17 percentage points. A full regression table is in the Appendix. Consistent with outcome 3, I find that the presence of a foreign female reduces male dominance in the conversation.

Rate of male dominance
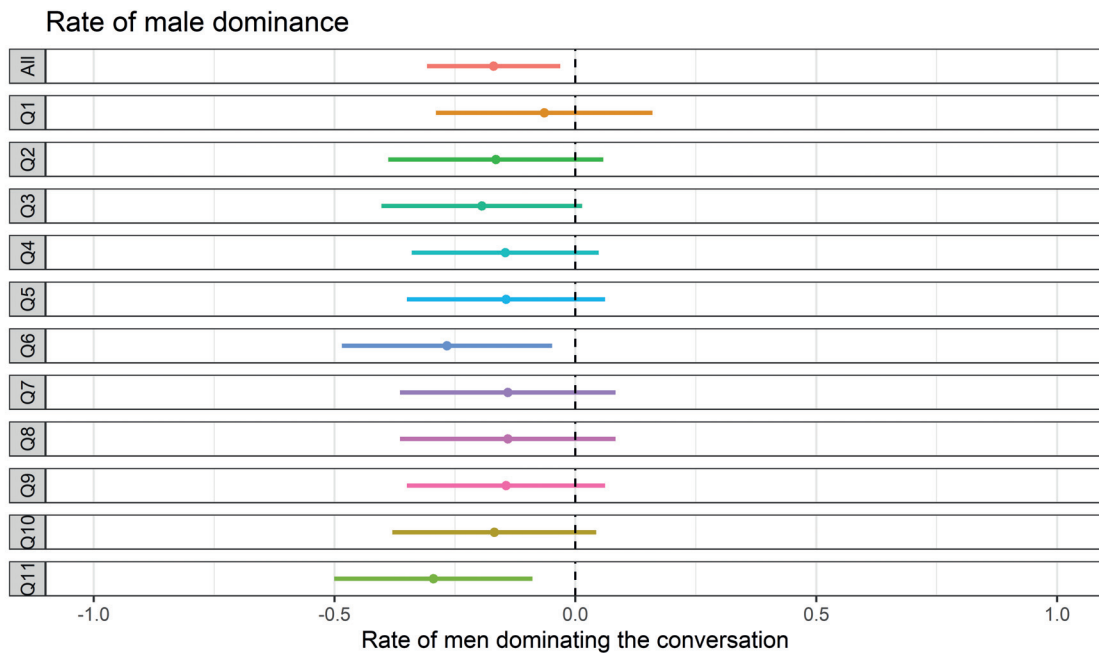
Figure 3.3: Male dominance of by topic

The data does not show the same pattern concerning the question of whether elderly people dominate the conversation. They were not more or less likely to dominate the conversation when I was present, suggesting that males were dominating less, instead of everyone else speaking more. If the latter had been the case, I would have expected to see similar results in other measures of conversation dominance. However, it may also be that I am less sensitive to man dominating the conversation than the Tanzanian researchers, so recorded what was actually similar rates of dominance less often.

3.4.4 Outcome 4: Participants will be more thorough in responses with a white female researcher

In question 7, which was also significantly longer when I was present, the focus group was asked to rank several types of people by their importance in community decision making. They were given eight types of community members, and after an initial ranking, were also invited to add other members of the community that were not on the list. Taking this as a measure of thoroughness for the sake of a naïve listener, or for the sake of social desirability, I hypothesized that when I was present, focus groups would collectively list more people in this second round of the ranking.

Table 4 shows that when I was present, groups produced an average of .63 additional names for the ranking, which is statistically significant at the conventional .05 level and follows the hypothesized outcome 4. The full regression results are in the Appendix

3.4.5 Outcome 5: Women will be reported to be more high-ranking when a white female researcher is present.

In that same exercise, and by similar logic of social desirability bias, I would expect to see that women would be ranked relatively higher when I was present, as the participants may belief that a I will think better of them if they valued the opinion of women. However, young women were most often ranked in the eighth position in this exercise. Interestingly, I find that I had a negative effect on the ranking of young girls in this exercise. When I was present, young women were on average 1 position lower (p-value of .06) than when I was absent (see Appendix). The ranks were relative to the total number of positions ranked, so the addition of new positions is accounted for.

This effect is not seen in the rankings of elderly women, who were ranked about on average two positions higher than the young women, but who saw no significant change in rank when I was present. The finding is contrary to the hypothesis that social desirability bias would lead to a higher ranking of women[3]. Figure 3.4, below, shows the shift for each group.

I suspect that the reason for this effect is two-fold. First, the participants also took about 3 minutes longer to complete this task when I am present than when I am not. It seems reasonable, (though undemonstrated) that participants are being more careful to explain this topic to a naïve listener, especially when the rank order can be so easily understood, even to a non-Swahili speaker. Secondly, this seems to be indicating evidence for a reference point effect. Young women might have some

---

[3]In one case, a young woman in the group tried to suggest that young women should be in the lowest ranked spot, since no one listens to them. Unfortunately, the group was talking over her input and placed the "young women" card near the middle of the ranking.

say in community decisions, but compared to the perceived responsibility of the foreigner, this is diminished.
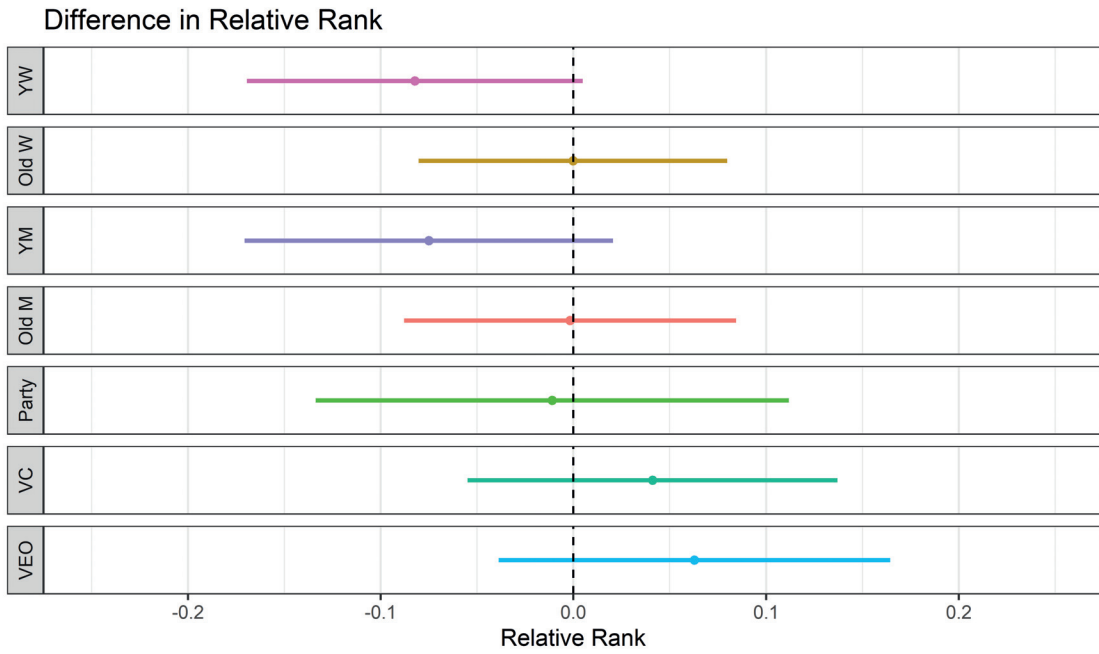

Figure 3.4: Relative ranking of each group

3.5 Conclusion

This research explores the effects of a white female researcher on focus groups in Tanzania. Using experimental methods, I randomly assign my own presence to 39 focus groups, leaving 40 additional focus groups as the controlled condition. I found that, on average, focus groups lasted longer and were less likely to be dominated by male participants when I was present. I also found that focus group participants were more likely to provide additional information. These effects were localized around questions in which the attributes of gender or foreignness of the researcher would be most salient, i.e. questions about gender or cultural practices. Respondents not only discussed gender-related questions at greater length, but they discussed all topics with greater gender equality, as measured by the diminished number of cases in which men dominated the conversation.

Surprisingly, on one measure – the rank of women in the village – the hypothesized social desirability bias was not found, but rather the reverse: young women were ranked lower in significance.

Though I have attributed the change in length in discussion to my effect on the focus group participants, another explanation for the extended discussions would consider the effect that I, as a PI on the project, on the Tanzanian researchers who are conducting the focus group. It is possible that some amount of shirking takes place when I, who has a position of authority on the project, am not present. While the Tanzanian focus group team was made up of qualified and capable individuals, it would not be unusual for even a dedicated worker to put in a little extra effort to push for more information when an authority figure is in the room. When a researcher is considering whether or not to be present during field work, they consider the trade-off between having additional supervisory control over the work and having unnecessary influence on the outcomes of the research.

However, while there is an effect on the total length of the focus group, much of that difference was driven by specific questions. If it were the facilitators driving the interaction, I would not expect to see particular patterns in shirking from the research team. The change then, seems to be due to the effect on focus group participants, rather than on the research team.

When considering how much contact to have with research subjects in a focus group setting, researchers should consider this evidence for the sensitivity bias. This evidence does not suggest that researchers will have less reliable content in focus groups if they are in attendance. A bias would mean that the information received is further from the parameter of interest than it would have otherwise been. In fact, depending on the topic, the researcher might expect to get more thorough information when they are present, as is observed when asking about relative influence of opinion in the village. However, as in the case of women as leaders in politics, some topics may have a moral weight to them that, in which case, respondents are likely to shape their answers to their perception of the values

of the researcher. In deciding whether or not to participate be present during data collection in the field, it would behoove researchers to understand local conditions and perceptions of the topic. Understanding the context in which the data is collected is key[4].

While these treatment effects are particular to myself, researchers are likely to experience similar patterns in their own field work. My study provides evidence that the presence of a white researcher can have substantial effect on the qualitative and quantitative data collection processes. This particular type of Hawthorn effect is especially important to understand as many social scientists conduct field work in places inn which they are conspicuous foreigners. The decision of how to balance the effect they may have on their respondents and the need to supervise data collection is a complicated and perennial one. The findings also reiterate the connection between the measurement and the measurer. Even in experimental research, it is essential for the researcher to consider the their own effect on the variables of interest.

---

[4]Gaining that understanding, of course, likely takes some amount of data collection and field work. While this may seem like a "turtles all the way down" problem, it is more likely an iterative process that takes repeated asking in a variety of contexts and ways.

Appendix A

Appendices

Appendices

A.1 Focus group questions- Female representation

1. General culture/ warm up

   (a) What are some of the most important elements of Tanzanian culture?
   (b) In what ways, if any, should men and women act differently in Tanzanian culture?

2. Politics and Representation

   (a) The Bunge has Special Seats for women to increase women's representation in parliament. Do you think that this has been good for Tanzania as a whole, bad for Tanzania, or neither all good nor all bad? Why?
   (b) Do you think that these special seats are good for WOMEN in Tanzania, bad for women, or neither all good nor all bad? Why?

3. Educational investment

   (a) Do you think that it as important for a girl to get an education as it is for a boy to get an education? Why do you think so?

4. Aspirations and Ideals

   (a) What would you hope your daughter might do in her life [note that this term should include nieces in translation]? What are some of the challenges that she might face?
   (b) Itf things went really well for a typical girl from a typical family in Tanzania, when she grew up, what would her like be like?

A.2 Focus group questions-Placebo

1. General culture/ warm up

   (a) What are some of the most important elements of Tanzanian culture?

2. Politics and representation

   (a) The Bunge seeks to represent the people of Tanzania. Do you think that they do a good job? What could they do to it improve?
   (b) Do you participate in politics a lot, a little, or not at all? Why?

3. Educational investment

   (a) What role does education play in Tanzanian culture? Has it been important in your life? What might you change about education in Tanzania?

4. Aspirations and ideals

    (a) What were some of your goals or dreams when you were a child? Did you achieve those goals? What were some of the challenges you faced?

    (b) If things went really well for a normal child from a typical family in Tanzania, what would their life look when they grew up [Note that the pronoun will be naturally ungendered in Swahili.]

A.3 Construction of the metrics

The outcomes were sorted into bundles of outcomes that were then combined into single metrics based on a common measure. The bundles are:

1. Behavioral outcomes

2. An index of personal prospective political participation

3. An index of perceived norms of women in politics

4. An index of perception of broad gender norms society

5. An index of aspirational norms for self or female child in the family

6. Measures of household bargaining power

The behavioral outcome was an expressed interest in receiving more information about an upcoming Gender Festival in Dar es Salaam put on annually by the Tanzanian Gender Networking Programme. The respondents were told briefly about the nature of the festival and then asked if they would like to provide contact information to learn more about the festival. They were also asked if they knew anyone else who would like to know about the festival. Those for whom we received contact information were sent more information in advance of the festival which took place in September of 2019. The respondents willingness to provide their contact information, as well as the total number contacts that the respondents provided (including themselves) is the behavioral index. More liberal views would be reflected in providing contact information for more people.

Personal prospective political participation is a measure of the respondents' reported willingness to run for public office of any kind, or other otherwise participate in politics. It is also a reflection of their view of their own fitness for office. These expressed views were measured via the following outcome measures on the survey:

1. How likely are you to do each of the following? [5-point Likert]

   (a) Hold a position in your local political party

   (b) Run for local office

   (c) Run for national office

2. Do you think someone like you could be elected to the village council in this village? (Likert)

3. How much do you agree with each of the following statements? [5-point Likert]

(a) Someone like me could do a good job on the local council

(b) Someone like me should run for local office

As these measure are well correlated, and have significant variation these outcomes will be combined into an index for personal willingness to participate in politics. A higher score on this index, as with all others, would correlate with more liberal attitudes towards gender

The measure of perceived norms around women's participation in politics will consist of two outcomes. The outcomes for this index had two sets of two questions each. The sets correlated well within the set, but not with each other. Each pair of outcomes will be averaged to create two measures of norms around women in politics. Questions from the second pair are taken from a long-standing measure of sexism that has been used in the General Social Survey (GSS) in the United States for several decades (Charles et al., 2018 (Working paper)).

1. For each of the following, please tell me whether you personally believe that it is morally acceptable, morally wrong, or is it not a moral issue?

    (a) A woman spending somewhat less time at home or with her children so that she can pursue local political office.

    (b) A woman spending somewhat less time at home or with her children so that she can pursue national political office.

2. How much do you agree or disagree with each of the following statements? [Likert scale]

    (a) Women should take care of running their home and leave running the country up to men.

    (b) Men are better suited emotionally for politics than are most women.

These measures (each an index of two questions) are used to indicate perceived gender norms regarding women in politics, with a higher outcome indicating more liberal attitudes towards gender. Other outcomes which did not have high enough variation have been dropped.

The outcome measures for broader gender norms in society consist of the following outcome variables from the survey. These measures will not be average, but will be individual measures of progressive gender norms. These were also taken from the GSS. These outcomes will not be combined because there is insufficient correlation between them. One variable in this category without significant variation has been dropped.

1. How much do you agree or disagree with each of the following statements? [Likert scale]

(a) A working mother can establish just as warm and secure relationship with her children as a mother who does not work.

(b) It is much better for everyone involved if the man is the achiever outside of the home and the woman takes care of the home and family.

(c) A small child is likely to suffer if his or her mother works.

The following questions are the outcome variable for aspirational norms for self or female child in the family. They will not be combined into an average. Question 1 will be treated initially as a continuous variable. These measures are intended to reflect the opinion of the respondent on the potential for daughter or young child in the family and what the ideal would look like. As above, some variables with little variation have been dropped.

1. About what age would you expect your daughter to get married?

2. If you had several children and could only afford to send one child to school, which child would you send to school?

   (a) The oldest child.
   (b) The youngest child.
   (c) The oldest female child.
   (d) The oldest male child.
   (e) Use some other decider.

Finally, the index for domestic bargaining power will be constructed with the following outcomes from the survey. These are standard measures of bargaining power, and as they all are highly correlated, they will be combined into an average.

1. Who usually makes decisions about making purchases for daily household needs, like soap, sugar, and kerosene?

2. Who usually makes decisions about visits to your family or relatives?

3. Who usually makes decisions about major household purchases, like buying a goat or cookstove?

A.4 Results from all pre-registered regressions

Figure A.3 (below) illustrates the outcome measurements by each of the 10 metrics listed above. The figure shows the group mean of each index by treatment condition. The measurement of views of morality of political participation for women is labeled "Moral Poli. Part." Per the submitted pre-analysis plan, all 9 attitudinal outcomes are reported.
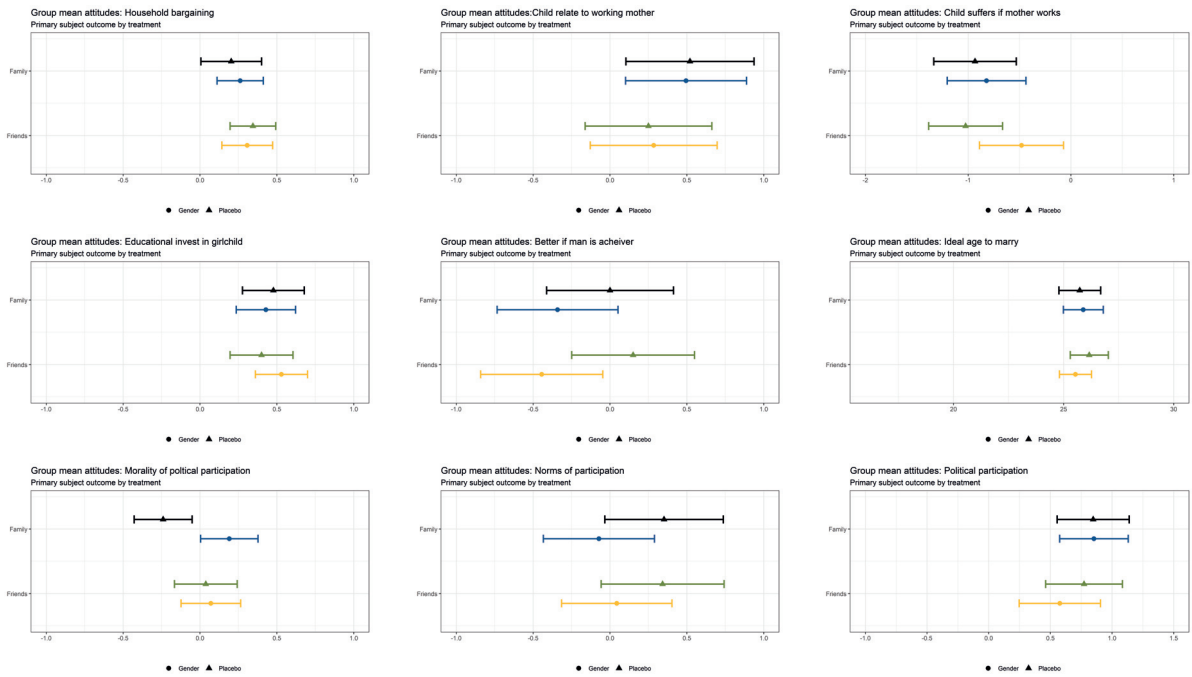
Table A.4 (below) shows the ATE, CATE for both the Family and Friends treatments, and the difference-in-differences for each of the regressions. Per the pre-analysis plan, all 10 regressions are reported with the interaction effect. Regression 3 shows the significant treatment effect on views of the morality of women participating in politics. Regression 10 shows the significant CATE among focus groups made up of friends on the behavioral outcome.

Table A.1: regression table for each of the 10 outcomes

	outcome	ATE	CATE family	CATE friends	DID
1	Political participation	-0.02 (0.14)	0.07 (0.18)	-0.11 (0.21)	-0.18 (0.28)
2	Norms of participation	-0.31 (0.18)	-0.30 (0.25)	-0.32 (0.25)	-0.02 (0.35)
3	Morality of participation	0.24 (0.09)	0.44 (0.13)	0.05 (0.13)	-0.39 (0.18)
4	Working mother relate to child	0.14 (0.19)	0.02 (0.27)	0.27 (0.27)	0.25 (0.38)
5	Man as achiever	-0.32 (0.19)	-0.11 (0.27)	-0.53 (0.27)	-0.42 (0.38)
6	Child suffers if mother works	0.18 (0.19)	0.04 (0.27)	0.31 (0.26)	0.28 (0.37)
7	Ideal marry age	-0.11 (0.41)	0.17 (0.62)	-0.39 (0.53)	-0.56 (0.82)
8	Educational aspiration	0.03 (0.09)	-0.02 (0.13)	0.09 (0.13)	0.11 (0.18)
9	Household bargaining	-0.02 (0.07)	0.00 (0.11)	-0.05 (0.10)	-0.05 (0.15)
10	Behavior outcomes	0.36 (0.23)	0.06 (0.31)	0.77 (0.33)	0.72 (0.45)

rotating

A.5 Regression table: Behavior outcomes, subjects

	Behavior outcomes	Behavior w/ interaction
Intercept	1.71*	2.00*
	$[1.33; 2.08]$	$[1.62; 2.38]$
Friends	0.06	−0.50
	$[-0.35; 0.47]$	$[-1.12; 0.12]$
Gender Group	0.57*	0.12
	$[0.14; 0.99]$	$[-0.45; 0.68]$
Gender x Friends		0.88*
		$[0.10; 1.66]$
R^2	0.14	0.22
Adj. R^2	0.10	0.17
Num. obs.	52	52
RMSE	0.71	0.68

* p-value .05

Behavior outcomes

Figure A.1: Participation rate by Question

Full participation rate: Question 1`
Mean of Focus Groups that had full participation on question 1

● Gender ▲ Placebo

Full participation rate: Question 2
Mean of Focus Groups that had full participation on question 2

● Gender ▲ Placebo

Full participation rate: Question 3
Mean of Focus Groups that had full participation on question 3

● Gender ▲ Placebo

Full participation rate: Question 4
Mean of Focus Groups that had full participation on question 4

● Gender ▲ Placebo

Full participation rate: Question 5
Mean of Focus Groups that had full participation on question 5

● Gender ▲ Placebo

Full participation rate: Question 6
Mean of Focus Groups that had full participation on question 6

● Gender ▲ Placebo

A.6 Group Means: Rate of Participation

A.7 Group means: Attitudes of all participants

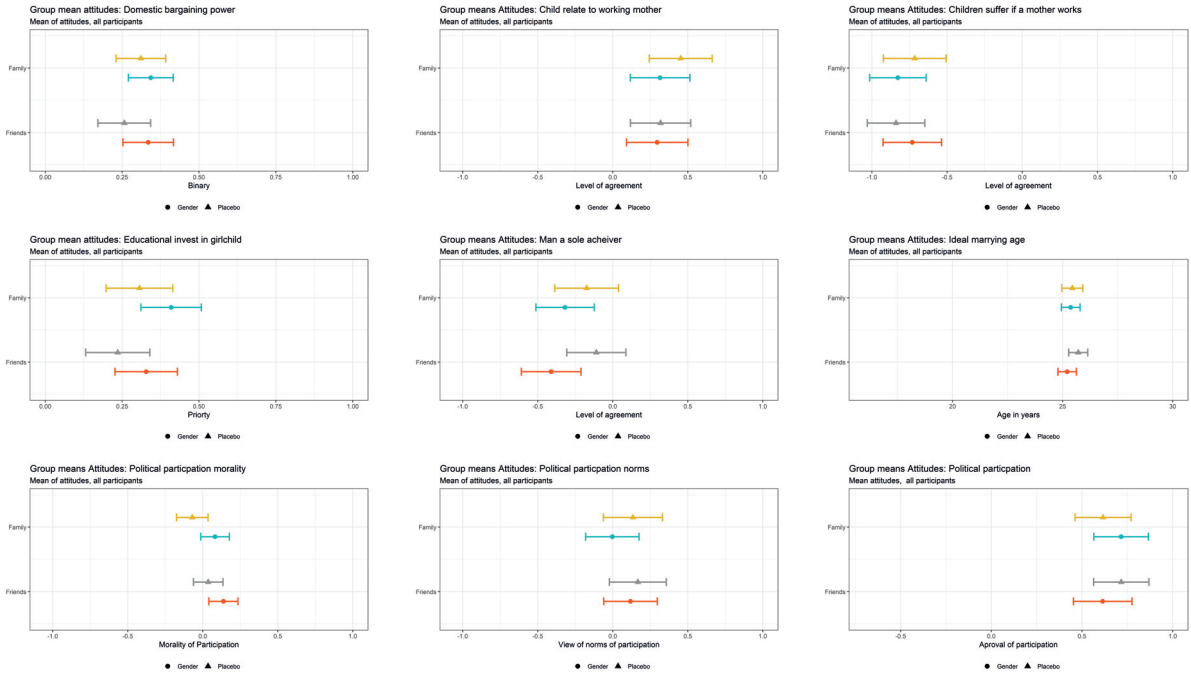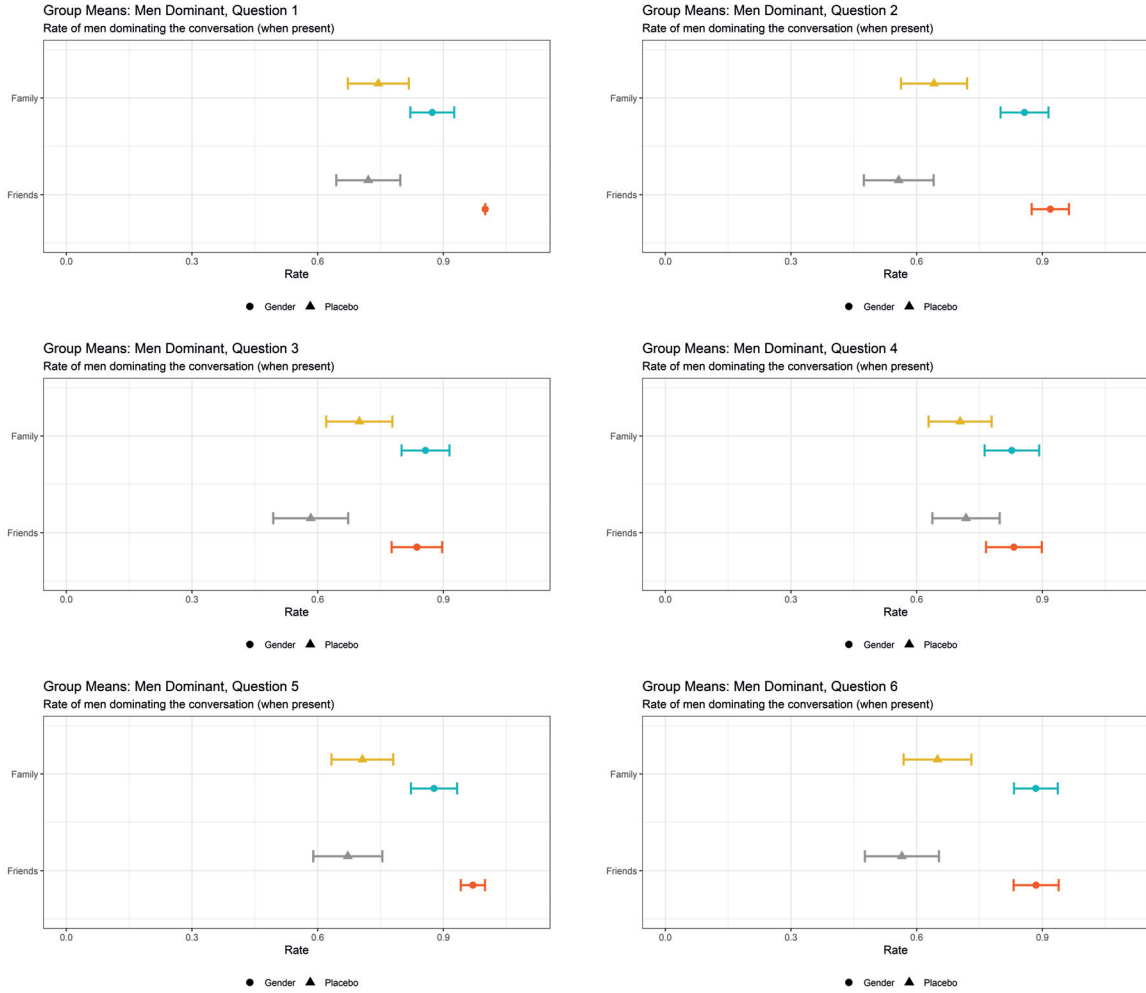Figure A.2: Some outcome variables with no variation across treatment for all FG participants

Figure A.3: Male Dominance by Question

A.8 Group means: Male dominance of Question

A.9 Wordclouds for placebo treatment

Figure A.4: Wordcloud: Family placebo group

Figure A.5: Wordcloud: Friends placebo group

A.10 Regression table: Men Dominating the Conversation

.1 Male Dominance of Discussion

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Overall	Q1 interaction	Q2 interaction	Q3 interaction	Q4 interaction	Q5 interaction	Q6 interaction	Q7 interaction	Overall interaction
rcept	0.71*	0.60*	0.69*	0.70*	0.68*	0.65*	1.01*	0.40*	0.71*	0.59*	0.67*	0.66*	0.68*	0.64*	1.00*	0.45*
nds	0.02	−0.03	−0.08	−0.01	−0.00	−0.05	−0.04	−0.07	0.03	−0.02	−0.05	0.09	0.01	−0.02	0.00	−0.18
der Group	0.18*	0.29*	0.19*	0.12	0.22*	0.23*	−0.10	0.07	0.19	0.30*	0.22*	0.22*	0.24*	0.26*	−0.09	−0.03
der x Friends									−0.01	−0.02	−0.07	−0.19	−0.03	−0.05	−0.04	0.21
	0.06	0.11	0.05	0.02	0.07	0.07	0.01	0.01	0.06	0.11	0.06	0.03	0.07	0.07	0.01	0.02
R^2	0.04	0.09	0.04	0.01	0.06	0.06	−0.02	−0.00	0.04	0.09	0.03	0.01	0.05	0.05	−0.03	−0.00
n. obs.	140	143	135	130	137	139	74	144	140	143	135	130	137	139	74	144
SE	0.38	0.42	0.43	0.43	0.39	0.42	0.31	0.49	0.38	0.42	0.43	0.43	0.39	0.42	0.32	0.49

hypothesis value outside the confidence interval.

dominate conversation outcomes

A.10.2   All Focus Groups Questions, in order

A.11   Focus Group Questions

Question 1: Can you tell us about the history of foreigners visiting this community? (i.e. Where were they from? When did they first come? What did they do here?)

Question 2: Would you say these historic interactions with foreigners brought benefits, problems or a bit of both to the community? [Ask speakers to elaborate.]

Question 3: Can you now tell us about any foreigners visiting this community in the past five years? (i.e. Where were they from? In what activities were they engaging?)

Question 4: Would you say these recent interactions with foreigners brought benefits, problems or a bit of both to the community? [Ask speakers to elaborate.]

Question 5 If two people in this community have a dispute over land, livestock or other resources that they could not resolve on their own, how would they try to resolve it?

Question 6: In the past – think back to the time before independence – if two people in this community had a conflict over land, livestock or other resources that they could not resolve on their own, how would they try to resolve it?

Question 7: When it comes to making decisions about things like water access, land use, school building or other community issues, there are many people and groups who have opinions and interests. We would like you to think about whose opinions carry the most influence in these types of decisions in this community. *(1)* Consider the following individuals and groups. How would you rank them in terms of influence? Village chairperson, village executive officer, party leaders, clan elders, elderly men, elderly women, young men, young women. [Have them move labeled cards around] *(2)* Is there any other individual or group who carries a lot of influence? [Write on new card]

Question 8: In cases where this community needs a new project – like a school, well, road, or agricultural project – how would you go about trying to accomplish it? [Follow-up, as necessary: Who would you expect to provide funds and resources? Most community members? Government? International donors and organizations?]

Question 9: In this community, what kind of things do people do to celebrate and honor their ancestors? [Follow up: What kind of things could happen if you did not do these things? Do all people do these things?]

Question 10: Do women play any leadership roles in this village? What roles? What challenges are involved in women taking on leadership roles?

Question 11: Do you think it is better to vote for a Member of Parliament who promises to deliver development projects to this community or to vote for a Member of Parliament who promises to make laws and policies to benefit the country? Why?

A.12 Regression table:With cultural competency

	Length (all)	Length Q7 (Ranks)	Length Q9 (Ancestors)	Length Q10 (Women)
Intercept	444.465***	1046.278***	446.333***	366.611***
	(21.684)	(74.189)	(33.600)	(23.159)
White female researcher				
Cultural competency	−6.014	100.246	74.762	−5.421
	(34.360)	(124.235)	(62.597)	(33.609)
R^2	0.000	0.016	0.034	0.001
Adj. R^2	−0.002	−0.010	0.008	−0.026
DF Resid.	17.000	17.000	17.000	17.000
N	429	39	39	39

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$

Table A.2: Discussion Length

A.13 Big Five Personality Index

How I am in general

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who likes to spend time with others? Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

1 Disagree Strongly 2 Disagree a little 3 Neither agree nor disagree 4 Agree a little 5 Agree strongly

I am someone who. . .

1. Is talkative
2. Tends to find fault with others
3. Does a thorough job
4. Is depressed, blue
5. Is original, comes up with new ideas
6. Is reserved
7. Is helpful and unselfish with others
8. Can be somewhat careless
9. Is relaxed, handles stress well
10. Is curious about many different things
11. Is full of energy
12. Starts quarrels with others
13. Is a reliable worker
14. Can be tense
15. Is ingenious, a deep thinker
16. Generates a lot of enthusiasm
17. Has a forgiving nature
18. Tends to be disorganized
19. Worries a lot
20. Has an active imagination
21. Tends to be quiet
22. Is generally trusting
23. Tends to be lazy
24. Is emotionally stable, not easily upset
25. Is inventive
26. Has an assertive personality
27. Can be cold and aloof
28. Perseveres until the task is finished
29. Can be moody
30. Values artistic, aesthetic experiences
31. Is sometimes shy, inhibited
32. Is considerate and kind to almost everyone
33. Does things efficiently
34. Remains calm in tense situations

35. Prefers work that is routine

36. Is outgoing, sociable

37. Is sometimes rude to others

38. Makes plans and follows through with them

39. Gets nervous easily

40. Likes to reflect, play with ideas

41. Has few artistic interests

42. Likes to cooperate with others

43. Is easily distracted

44. Is sophisticated in art, music, or literature

SCORING INSTRUCTIONS

To score the BFI, you'll first need to reverse-score all negatively-keyed items:

Extraversion: 6, 21, 31

Agreeableness: 2, 12, 27, 37

Conscientiousness: 8, 18, 23, 43

Neuroticism: 9, 24, 34

Openness: 35, 41

To recode these items, you should subtract your score for all reverse-scored items from 6. For example, if you gave yourself a 5, compute 6 minus 5 and your recoded score is 1. That is, a score of 1 becomes 5, 2 becomes 4, 3 remains 3, 4 becomes 2, and 5 becomes 1.

Next, you will create scale scores by averaging the following items for each B5 domain (where R indicates using the reverse-scored item).

Extraversion: 1, 6R 11, 16, 21R, 26, 31R, 36

Agreeableness: 2R, 7, 12R, 17, 22, 27R, 32, 37R, 42

Conscientiousness: 3, 8R, 13, 18R, 23R, 28, 33, 38, 43R

Neuroticism: 4, 9R, 14, 19, 24R, 29, 34R, 39

Openness: 5, 10, 15, 20, 25, 30, 35R, 40, 41R, 44

Enumerator	Extraversion	Agreeableness	Conscientiousness	Neuroticism (Negative emotionality)	Openness (open-mindedness
Neema (f)	4.5	4.7	4.6	2.1	4.1
Sarah (f)	3.8	4.4	4.6	2.3	3.4
Josephine (f)	4.2	4.7	4.8	1.7	4.1
Justine (m)	3.3	4.0	4.5	2.6	4.3
Elieza (m)	4.5	4.8	5.0	1.5	4.3
Liz	4.2	4.1	3.0	1.5	4.2
TZ average compared to US average.	=	=	+.5	-.2	-.1

A.14 Regression Table: Focus group length

	Total Focus Group Length	Q1: History of foreigners	Q2: Harm/ benefit of foreigners	Q3: Recent foreigners	Q4: Harm/ benefit of recent	Q5: Land disputes	Q6: Historical land disputes	Q7: Opinion ranking	Q8: Public goods project	
Intercept	395.43*	475.37*	321.35*	277.05*	212.65*	297.27*	295.40*	933.60*	482.37*	4
White Female Researcher	45.80*	−30.53	−68.76	66.33	71.22	43.11	23.45	166.66*	43.29	
R^2	0.01	0.00	0.04	0.04	0.04	0.02	0.01	0.05	0.02	
Adj. R^2	0.01	−0.01	0.03	0.02	0.03	0.01	−0.00	0.04	0.00	
Num. obs.	868	79	79	79	79	79	79	79	79	
RMSE	278.98	231.10	175.15	172.53	179.82	136.91	113.32	364.67	163.81	

* Null hypothesis value outside the confidence interval.

Table A.3: Discussion Length

A.15 Regression Table: Men dominating the conversation

	Total Focus Group Length	Q1: History of foreigners	Q2: Harm/ benefit of foreigners	Q3: Recent foreigners	Q4: Harm/ benefit of recent	Q5: Land disputes	Q6: Historical land disputes	Q7: Opinion ranking	Q8: Public goods project	
Intercept	0.464*	0.475*	0.550*	0.425*	0.325*	0.375*	0.600*	0.525*	0.525*	(
White Female Researcher	−0.170*	−0.065	−0.165	−0.194	−0.146	−0.144	−0.267*	−0.140	−0.140	−
R^2	0.031	0.004	0.027	0.043	0.028	0.025	0.071	0.020	0.020	
Adj. R^2	0.030	−0.009	0.015	0.030	0.015	0.012	0.059	0.007	0.007	
Num. obs.	869	79	79	79	79	79	79	79	79	
RMSE	0.478	0.502	0.498	0.466	0.434	0.460	0.487	0.499	0.499	

* Null hypothesis value outside the confidence interval.

Table A.4: Did men dominate the discussion

A.16 Regression Table: Relative ranking of groups

	Total number of positions	Village Chair Ranking	Village Exec Officer	Party Leaders	Old Men	Old Women	Young Men	Young Women
Intercept	9.639*	0.306*	0.384*	0.529*	0.567*	0.674*	0.653*	0.781*
White Female Researcher	0.633	−0.041	−0.063	0.011	0.002	0.000	0.075	0.082
R^2	0.042	0.009	0.019	0.000	0.000	0.000	0.030	0.045
Adj. R^2	0.041	−0.004	0.006	−0.013	−0.013	−0.013	0.018	0.032
Num. obs.	769	79	79	79	79	80	80	78
RMSE	1.506	0.214	0.227	0.274	0.191	0.182	0.215	0.193

* Indicates p .05

Table A.5: Relative rank by foreginer presence

Bibliography

Muzdalifat Abeid, Projestine Muganyizi, Pia Olsson, Elisabeth Darj, and Pia Axemo. Community perceptions of rape and child sexual abuse: a qualitative study in rural tanzania. *BMC International Health and Human Rights*, 14, 2014.

Amitav Acharya. Who are the norm makers? the asian-african conference in bandung and the evolution of norms. 20(3):405 – 417, 19 Aug. 2014. doi: https:// doi.org/10.1163/19426720-02003006. URL https://brill.com/view/journals/ gg/20/3/article-p405_6.xml.

Amitav Acharya. How ideas spread: Whose norms matter? norm localization and institutional change in asian regionalism. *International organization*, 58(2):239–275, 2004.

Claire Adida, Karen Ferree, Daniel Posner, and Amanda Robinson. "who's asking? interviewer coethnicity effects in african survey data.". 49(12):1630–1660, 2016.

Afrobarometer. Afrobarometer round 6 tanzania, 2015. URL https:// afrobarometer.org/online-data-analysis/analyse-online.

Icek Ajzen and Martin Fishbein. Attitudes and the attitude-behavior relation: Reasoned and automatic processes. *European Review of Social Psychology*, 11(1): 1–33, 2000. doi: 10.1080/14792779943000116. URL https://doi.org/10.1080/ 14792779943000116.

Icek Ajzen and Martin Fishbein. The influence of attitudes on behavior. 2005.

Alberto Alesina, Paola Giuliano, and Nathan Nunn. On the origins of gender roles: Women and the plough *. *The Quarterly Journal of Economics*, 128(2):469–530, 2013. doi: 10.1093/qje/qjt005. URL +http://dx.doi.org/10.1093/qje/qjt005. /oup/backfile/content$_public/journal/qje/$128/2/10.1093$_qje_qjt$005/4/qjt005.pdf.

G. W. Allport. *Attitudes*, pages 798–844. Clark University Press, Worcester, MA, US, 1935. ID: 1935-19907-010.

Siwan Anderson and Mukesh Eswaran. What determines female autonomy? evidence from bangladesh. *Journal of Development Economics*, 90(2):179–191, November 2009 2009. doi: https://doi.org/10.1016/j.jdeveco.2008.10.004. ID: 271688.

Robert Axelrod and William Donald Hamilton. The evolution of cooperation. *science*, 211(4489):1390–1396, 1981.

Albert Bandura. Social foundations of thought and action. 1986(23-28), 1986.

Albert Bandura. Self-efficacy. pages 1–3, 2010.

Lori Beaman, Esther Duflo, Rohini Pande, and Petia Topalova. Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science*, 335(6068):582–586, 2012. URL http://www.jstor.org/stable/41487241.

Paul Allen Beck and M Kent Jennings. Family traditions, political periods, and the development of partisan orientations. 53(3):742–763, 1991.

Myra Betron. Gender-bassed violence in tanzania: An assessment of policies, services, and promising interventions. Technical Report Task Order 1, USAID Health Policy Initiative, 2008. URL http://pdf.usaid.gov/pdf_docs/Pnadn851.pdf.

Nita Bhalla. Tanzania lets pregnant girls attend school with world bank project. *Reuters*, April 2020.

G. Blair, A. Coppock, and M. Moor. "when to worry about sensitivity bias: A social reference theory and evidence from 30 years of list experiments." american political science review: 1-19. pages 1–19, 2020.

Lisa Blaydes and Rachel Gillum. Religiosity-of-interviewer effects: Assessing the impact of veiled enumerators on survey response in egypt. 6(3):459–485, 2013.

Lisa Blaydes and Drew A. Linzer. The political economy of women's support for fundamentalist islam. *World Politics*, 60(4):576–609, 2008. ISSN 00438871, 10863338. URL http://www.jstor.org/stable/40060212.

Rae Lesser Blumberg. Income under female versus male control: Hypotheses from a theory of gender stratification and data from the third world. *Journal of Family Issues*, 9(1):51–84, 03/01; 2018/06 1988. doi: 10.1177/019251388009001004. URL https://doi.org/10.1177/019251388009001004. doi: 10.1177/019251388009001004; 14.

]Boserup Ester Boserup. Woman's role in economic development., [1970].

John Bullock, Donald Green, and Shang Ha. Yes, but what's the mechanism? (don't expect an easy answer). *Journal of personality and social psychology*, 98:550–8, 04 2010. doi: 10.1037/a0018933.

Leonardo Bursztyn, Alessandra L. González, and David Yanagizawa-Drott. Misperceived social norms: Women working outside the home in saudi arabia. *American Economic Review*, 110(10):2997–3029, October 2020. doi: 10.1257/aer.20180975. URL https://www.aeaweb.org/articles?id=10.1257/aer.20180975.

Bruce Campbell. Race-of-interviewer effects among southern adolescents. 45(2):
    231–244, 1981.

Michael Carpini, X. Delli, Lomax Cook, and Lawrence Jacobs. Public deliberation,
    discursive participation, and citizen engagement. *Annual Review of Political Science*,
    7:315–344, 2004.

Katherine Casey, Rachel Glennerster, and Edward Miguel. Reshaping institutions:
    Evidence on aid impacts using a preanalysis plan. 127(4):1755–1812, 2012.

Joseph A. Catania, Diane Binson, Jesse Canchola, Lance M. Pollack, Walter Hauck,
    and Thomas J. Coates. Effects of interviewer gender, interviewer choice, and item
    wording on responses to questions concerning sexual behavior. 60(3):345–375, 1996.

Kerwin Kofi Charles, Jonathan Guryan, and Jessica Pan. The effects of sexism on
    american women: The role of norms vs. discrimination. *National Bureau of Economic
    Research*, 2018 (Working paper). URL http://www.nber.org/papers/w24904.

Raghabendra Chattopadhyay and Esther Duflo. Women as policy makers: Evidence
    from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443, 2004.
    URL http://www.jstor.org/stable/3598894.

Jacobus Cilliers, Oeindrila Dube, and Bilal Siddiqi. The white-man effect: How
    foreigner presence affects behavior in experiments. 118c:397–414, 2015.

Amanda Clayton. Women's political engagement under quota-mandated female rep-
    resentation: Evidence from a randomized policy experiment. *Comparative Political
    Studies*, 2015.

Karisa Cloward. False commitments: Local misrepresentation and the international
    norms against female genital mutilation and early marriage. 68(3):495–526, 2014.
    ISSN 00208183, 15315088. URL http://www.jstor.org/stable/43282117.

Patrick R. Cotter, Jeffrey Cohen, and Philip Coulter. Race of-interviewer effects in
    telephone interviews. 46(2):278–284, 1982.

Jens Olav Dahlgaard et al. Trickle-up political socialization: The causal effect on
    turnout of parenting a newly enfranchised voter. 112(3):698–705, 2018.

Darren W. Davis. The direction of race of interviewer effects among african-americans:
    Donning the black mask. 41(1):309–322, 1997.

Kenneth G DeMarree, S Christian Wheeler, and Richard E Petty. Priming a new
    identity: self-monitoring moderates the effects of nonself primes on self-judgments
    and behavior. 89(5):657, 2005.

Oeindrila Dube and Harish S.P. Queens. pages 2019–120, August 2019.

Alice H Eagly and Shelly Chaiken. *The psychology of attitudes.* Harcourt brace Jovanovich college publishers, 1993.

Alice H. Eagly and Maureen Crowley. Gender and helping behavior. a meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100(3):283–308, 1986.

Catherine C. Eckel and Philip J. Grossman. Are women less selfish than men?: Evidence from dictator experiments. *The Economic Journal*, 108(448):726–735, 1998. URL http://www.jstor.org/stable/2565789.

Mats Ekström and Johan Östman. Family talk, peer talk and young people's civic orientation. 28(3):294–308, 2013. doi: 10.1177/0267323113475410. URL https://doi.org/10.1177/0267323113475410.

RC Ellickson. The market for social norms. 3(1):1–49, 2001. ISSN 1465-7252. doi: 10.1093/aler/3.1.1. URL https://doi.org/10.1093/aler/3.1.1.

Seymour Epstein. The self-concept revisited: Or a theory of a theory. 28(5):404, 1973.

Russell H Fazio. How do attitudes guide behavior. *Handbook of motivation and cognition: Foundations of social behavior*, 1:204–243, 1986.

James D Fearon. Domestic political audiences and the escalation of international disputes. pages 577–592, 1994.

Jacob J. Feldman, Herbert Hyman, and Clyde Hart. A field study of interviewer effects on the quality of survey data. 15(4):734–761, 1951.

E. F. Fern. he use of focus groups for idea generation: The effects of group size, acquaintanceship, and moderator on response quantity and quality. 19(1):1–13, 1982.

Leon Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1957.

Leon Festinger. Cognitive dissonance. 207(4):93–106, 1962. ISSN 00368733, 19467087. URL http://www.jstor.org/stable/24936719.

Robert J. Fisher. Social desirability bias and the validity of indirect questioning. 20(2): 303–315, 1993.

Alan S. Gerber and Todd Rogers. Descriptive social norms and motivation to vote: Everybody's voting and so should you. *The Journal of Politics*, 71(1):178–191, 2009. doi: 10.1017/S0022381608090117. ID: 2010-08094-007.

Erving Goffman. *The Presentation of Self in Everyday Life.* Anchor Books, 1959.

Tanushree Goyal. How women mobilize women into politics: A natural experiment in india, 2020.

Eddie Harmon-Jones and Judson Mills. *An introduction to cognitive dissonance theory and an overview of current perspectives on the theory.*, pages 3–24. 2019. ISBN 1433830108. doi: 10.1037/0000135-001.

Shirley Hatchett and Howard Schuman. White respondents and race-of-interviewer effects. 39(4):523–528, 1975.

Jocelyn A Hollander. The social contexts of focus groups. *Journal of contemporary ethnography*, 33(5):602–637, 2004.

Robert J. Homant and Daniel B. Kennedy. Police perceptions of spouse abuse: A comparison of male and female officers, 1985 1985. ID: 271758.

MALA HTUN and S. LAUREL WELDON. The civic origins of progressive policy change: Combating violence against women in global perspective, 1975–2005. *The American Political Science Review*, 106(3):548–569, 2012. URL http://www.jstor.org/stable/23275433.

Leonie Huddy, Joshua Billig, John Bracciodieta, Lois Hoeffler, Patrick J. Moynihan, and Patricia Pugliani. The effect of interviewer gender on the survey response. 19 (3):197–220, 1997.

Macartan Humphreys, William A. Masters, and Martin E. Sandbu. The role of leaders in democratic deliberations: Results from a field experiment in são tomé and príncipe. *World Politics*, 58(4):583–622, 2006. doi: 10.1353/wp.2007.0008.

Institute for Democracy & Electoral Assistance International IDEA. Gender quotas database, 2020.

International Republican Institute IRI. *Tanzania National Elections Gender Assessment*. 1225 Eye Street NW, Suite 700 Washington, DC 20005, 10 2015.

Emily W. Kane and Laura J. Macaulay. Interviewer gender and gender attitudes. 57 (1):1–28, 1993.

Diana Kapiszewski, Lauren MacLean, and Benjamin Read. *Field Research in Political Science: Practices and Principles.* Cambridge University Press, 2015.

Christopher F Karpowitz and Tali Mendelberg. *The silent sex: Gender, deliberation, and institutions.* Princeton University Press, 2014.

Andrew Kiondo and Fidelis Mtatifikolo. Developing and sustaining ngos in tanzania. challenges and opportunities in the new millennium. In *OSSREA Workshop University of Dar es Salaam, Dar es Salaam*, 1999.

Maria Knight Lapinski and Rajiv N Rimal. An explication of social norms. *Communication theory*, 15(2):127–147, 2005.

Ioana M. Latu, Marianne Schmid Mast, Joris Lammers, and Dario Bombari. Successful female leaders empower women's behavior in leadership tasks, May 2013 2013. ID: 272387.

Pei-Te Lien. Does the gender gap in political attitudes and behavior vary across racial groups? *Political Research Quarterly*, 51(4):869–894, 1998.

Eugene Litwak and Ivan Szelenyi. Primary group structures and their functions: Kin, neighbors, and friends. *American Sociological Review*, 34:465, 08 1969. doi: 10.2307/2091957.

Kaveh Majlesi. Labor market opportunities and women's decision making power within households, March 2016 2016. ID: 271688.

Andrew Manley, Haydn Morgan, and Joshua Atkinson. 'mzungu!': implications of identity, role formation and programme delivery in the sport for development movement. 8(3):383–402, 2016. doi: 10.1080/19406940.2014.962072. URL https://doi.org/10.1080/19406940.2014.96207.

Antony SR Manstead. Attitudes and behaviour. *Applied social psychology*, pages 3–29, 1996.

Michael McDevitt and Steven Chaffee. From top-down to trickle-up influence: Revisiting assumptions about the family in political socialization. 19(3):281–301, 2002. doi: 10.1080/01957470290055501. URL https://doi.org/10.1080/01957470290055501.

Rachel I McDonald and Christian S Crandall. Social norms and social influence. 3:147–151, 2015. ISSN 2352-1546. doi: https://doi.org/10.1016/j.cobeha.2015.04.006. URL https://www.sciencedirect.com/science/article/pii/S2352154615000558. Social behavior.

Daniel C. Molden. Understanding priming effects in social psychology: What is "social priming" and how does it occur? 32(Supplement):1–11, 2014. doi: 10.1521/soco.2014.32.supp.1. URL https://doi.org/10.1521/soco.2014.32.supp.1.

Jana Morgan and Melissa Buice. Latin american attitudes toward women in politics: The influence of elite cues, female advancement, and individual characteristics. *American Political Science Review*, 107(4):644–662, 2013. doi: 10.1017/S0003055413000385.

Projestine Selestine Muganyizi, Lennarth Nyström, Gunilla Lindmark, Maria Emmelin, Siriel Massawe, and Pia Axemo. Effect of supporter characteristics on expression of negative social reactions toward rape survivors in dar es salaam,

tanzania. *Health care for women international*, 31(8):668–685, 07/16 2010. doi: 10.1080/07399331003629378. URL https://doi.org/10.1080/07399331003629378. doi: 10.1080/07399331003629378.

Ragnhild L. Muriaas, Vibeke Wang, Lindsay Benstead, Boniface Dulani, and Lise Rakner. Why the gender of traditional authorities matters: Intersectionality and women's rights advocacy in malawi. *Comparative Political Studies*, 52(12):1881–1924, 2019. doi: 10.1177/0010414018774369. URL https://doi.org/10.1177/0010414018774369.

Diana Mutz. The consequences of cross-cutting networks for political participation. *American Journal of Political Science*, 46(4):838–855, 2002.

Revocatus Evarist Ngowi and Dev Jani. Residents' perception of tourism and their satisfaction: Evidence from mount kilimanjaro, tanzania. 35(6):731–742, 2018. doi: 10.1080/0376835X.2018.1442712. URL https://doi.org/10.1080/0376835X.2018.1442712.

Mary Johnson Osirim. Making good on commitments to grassroots women: Ngos and empowerment for women in contemporary zimbabwe. 24(2):167–180, 2001. ISSN 0277-5395. doi: https://doi.org/10.1016/S0277-5395(01)00154-6. URL https://www.sciencedirect.com/science/article/pii/S0277539501001546.

Laura Paler, Leslie Marshall, and Sami Atallah. How cross-cutting discussion shapes support for ethnic politics: Evidence from and experiment in lebanon. *Quarterly Journal of Political Science*, 15:33–71, 2020.

Elizabeth Levy Paluck. The promising integration of field experimentation and qualitative methods. 628:9–71, 2010.

Fred Pampel. Cohort change, diffusion, and support for gender egalitarianism in cross-national perspective. *Demographic Research*, 25:667–694, 2011.

IPU Parline. *Global and Regional Averages*, 2020. URL https://data.ipu.org/women-averages.

Ellen Quintelier. Engaging adolescents in politics: The longitudinal effect of political socialization agents. 47(1):51–69, 2015. doi: 10.1177/0044118X13507295. URL https://doi.org/10.1177/0044118X13507295.

Michelle C. Reiss and Kaushik Mitra. The effects of individual difference factors on the acceptability of ethical and unethical workplace behaviors. *Journal of Business Ethics*, 17(14):1581–1593, 1998. URL http://www.jstor.org/stable/25073992.

Jennifer Rosen. The effects of political institutions on women's political representation: A comparative analysis of 168 countries from 1992 to 2010. *Political Research Quarterly*, 66(2):306–321, 2013. doi: 10.1177/1065912912449698. URL https://doi.org/10.1177/1065912912449698.

D. P. Schmitt, J. Allik, R. R. McCrae, and V. Benet-Martinez. The geographic distribution of big five personality traits: Patterns and profiles of human self-description across 56 nations. 38(2):173–212, 2007.

Aner Sela and Baba Shiv. Unraveling Priming: When Does the Same Prime Activate a Goal versus a Trait? 36(3):418–433, 2009-03. ISSN 0093-5301. doi: 10.1086/598612. URL https://doi.org/10.1086/598612.

Paschal Sheeran and Thomas L Webb. The intention–behavior gap. *Social and personality psychology compass*, 10(9):503–518, 2016.

Issa G Shivji. Reflections on ngos in tanzania: what we are, what we are not, and what we ought to be. 14(5):689–695, 2004.

Dirk Smeesters, S Christian Wheeler, and Aaron C Kay. The role of interpersonal perceptions in the prime-to-behavior pathway. 96(2):395, 2009a.

Dirk Smeesters, Vincent Y. Yzerbyt, Olivier Corneille, and Luk Warlop. When do primes prime? the moderating role of the self-concept in individuals' susceptibility to priming effects on social behavior. 45(1):211–216, 2009b. ISSN 0022-1031. doi: https://doi.org/10.1016/j.jesp.2008.09.002. URL https://www.sciencedirect.com/science/article/pii/S0022103108001790.

Judith G. Smetana. The role of parents in moral development: A social domain analysis. 28(3):311–321, 1999. doi: 10.1080/030572499103106. URL https://doi.org/10.1080/030572499103106.

Paul C. Stern, Thomas Dietz, Troy Abel, Gregory A. Guagnano, and Linda Kalof. A value-belief-norm theory of support for social movements: The case of environmentalism. 6(2):81–97, 1999. ISSN 10744827, 22040919. URL http://www.jstor.org/stable/24707060.

William B Swann jr and Michael D Buhrmester. Self-verification. 2011.

Margaret Tankard and Elizabeth Paluck. Norm perception as a vehicle for social change. *Social Issues and Policy Review*, 10:181–211, 01 2016. doi: 10.1111/sipr.12022.

Deborah J Terry, Michael A Hogg, and Katherine M White. The theory of planned behaviour: self-identity, social identity and group norms. 38(3):225–244, 1999.

Jacqui True. The political economy of violence against women /, c2012.

Andreas Tutić. Revealed norm obedience. *Social Choice and Welfare*, 44(2):301–318, 2015. ISSN 01761714, 1432217X. URL http://www.jstor.org/stable/43662592.

Christopher Udry. Fieldwork, economic theory, and research on institutions in developing countries. 93(2):107–111, 2003.

UNWomen. Women claim their space in tanzania's elections, October 2015.

Peter van der Windt, Macartan Humphreys, and Raul Sanchez de la Sierra. Gender quotas in development programming: Null results from a field experiment in congo, Available online 2 March 2018 2018. ID: 271688.

Rachel B. Vogelstein and Alexandra Bro. *Women's Power Index*, 2020. URL https://www.cfr.org/article/womens-power-index.

Thomas J. Volgy and Sandra Sue Volgy. Women and politics: Political correlates of sex-role acceptance. *Social Science Quarterly*, 55(4):967–974, 1975. ISSN 00384941, 15406237. URL http://www.jstor.org/stable/42859425.

Derek J Wade, B.C Mwasaga, and Paul F.J. Eagles. A history and market analysis of tourism in tanzania. 22(1):93–101, 2001. ISSN 0261-5177. doi: https://doi.org/10.1016/S0261-5177(00)00019-4. URL https://www.sciencedirect.com/science/article/pii/S0261517700000194.

David S. Wallace, René M. Paulson, Charles G. Lord, and Jr. Charles F. Bond. Which behaviors do attitudes predict? meta-analyzing the effects of social pressure and perceived difficulty. *Review of General Psychology*, 9(3):214–227, 2005. doi: 10.1037/1089-2680.9.3.214. URL https://doi.org/10.1037/1089-2680.9.3.214.

Christina Wegs, Andreea A Creanga, Christine Galavotti, and Emmanuel Wamalwa. Community dialogue to shift social norms and enable family planning: an evaluation of the family planning results initiative in kenya. 11(4):e0153907, 2016.

Alexander Weinreb. The limitations of stranger-interviews in rural kenya. 71(6):1014–1039, 2006.

S Christian Wheeler, Kimberly Rios Morrison, Kenneth G DeMarree, and Richard E Petty. Does self-consciousness increase or decrease priming effects? it depends. 44(3):882–889, 2008.

Christina Wolbrecht and David E. Campbell. Leading by example: Female members of parliament as political role models. *American Journal of Political Science*, 51(4):921–939, 2007. URL http://www.jstor.org/stable/4620108.

Christina Wolbrecht and David E. Campbell. Role models revisited: youth, novelty, and the impact of female candidates. *Politics, Groups, and Identities*, 5(3):418–434, 07/03 2017. doi: 10.1080/21565503.2016.1268179. URL https://doi.org/10.1080/21565503.2016.1268179. doi: 10.1080/21565503.2016.1268179.

Lena Wängnerud. Women in parliaments: Descriptive and substantive representation. *Annual Review of Political Science*, 12(1):51–69, 2009. doi: 10.1146/annurev.polisci.11.053106.123839. URL https://doi.org/10.1146/annurev.polisci.11.053106.123839.

Mi Yung Yoon. Special seats for women in the national legislature: The case of tanzania. *Africa Today*, 55(1):61–86, 2008. URL http://www.jstor.org/stable/27666951.

Chance York. Is it top-down, trickle-up, or reciprocal?: Testing longitudinal relationships between youth news use and parent and peer political discussion. 70 (4):377–393, 2019. doi: 10.1080/10510974.2019.1614965. URL https://doi.org/10.1080/10510974.2019.1614965.