



Can Teacher Training Programs Influence Gender Norms? Mixed-Methods Experimental Evidence from Northern Uganda

Author(s): Marjorie Chinen, Andrea Coombes, Thomas De Hoop, Rosa Castro-Zarzur, and Mohammed Elmeski

Source: *Journal on Education in Emergencies*, Vol 3, No 1 (July 2017), pp 44 - 78

Published by: Inter-Agency Network for Education in Emergencies

Stable URL: <http://hdl.handle.net/2451/39662>

DOI: <https://doi.org/10.17609/N81T0D>

REFERENCES:

This is an open-source publication. Distribution is free of charge. All credit must be given to authors as follows:

Chinen, Marjorie, Andrea Coombes, Thomas De Hoop, Rosa Castro-Zarzur, and Mohammed Elmeski. 2017. "Can Teacher Training Programs Influence Gender Norms? Mixed-Methods Experimental Evidence from Northern Uganda." *Journal on Education in Emergencies* 3(1): 44-78.

The ***Journal on Education in Emergencies (JEiE)*** publishes groundbreaking and outstanding scholarly and practitioner work on education in emergencies (EiE), defined broadly as quality learning opportunities for all ages in situations of crisis, including early childhood development, primary, secondary, non-formal, technical, vocation, higher and adult education.

Copyright © 2017, Inter-Agency Network for Education in Emergencies. All rights reserved.



The Journal on Education in Emergencies published by the [Inter-Agency Network for Education in Emergencies](#) is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](#).

CAN TEACHER-TRAINING PROGRAMS INFLUENCE GENDER NORMS? MIXED- METHODS EXPERIMENTAL EVIDENCE FROM NORTHERN UGANDA

MARJORIE CHINEN, ANDREA COOMBES, THOMAS DE HOOP,
ROSA CASTRO-ZARZUR, AND MOHAMMED ELMESKI

ABSTRACT

This mixed-methods cluster-randomized controlled trial examines the impact of a teacher-training program that aimed to promote positive gender socialization in the conflict-affected region of Karamoja, Uganda. The theory of change suggests that the education system and teachers can play critical roles in promoting positive gender roles and gender equality, which has important implications for peacebuilding. Our study found evidence that the program positively influenced teachers' knowledge about the difference between gender and sex, and their attitudes toward gender roles and gender identity. We found no quantitative evidence for any short-term change in teachers' practices as a result of the program, nor did we find quantitative evidence of effects from a complementary, randomly assigned text-message intervention meant to reinforce the information delivered during the training. Qualitative research suggested that, while teachers adopted basic practices taught in the training, they were unready or unable to adopt more complex practices. The main implication is that training can influence teachers' knowledge and attitudes on gender equality, but traditional gender norms can be a barrier to changing behavior in the short term. A further implication is the importance of involving the community to create enabling environments in which new ideas about gender equality can be accepted and translated into practice.

Received August 30, 2016; revised February 27, 2017, and May 26, 2017; accepted June 7, 2017; electronically published July 2017.

Journal on Education in Emergencies, Vol. 3, No. 1

Copyright © 2017 by the Inter-Agency Network for Education in Emergencies (INEE). All rights reserved.
ISSN 2518-6833

INTRODUCTION

Gender equality is a fundamental human right. When women and men do not have equal access to resources or equal opportunities, there are direct economic and social costs. These costs largely affect women, and also have consequences for their children, communities, and countries. For instance, disadvantages in education translate into a lack of skills and limited access to opportunities in the labor market, which in turn affect social progress (Sustainable Development Goals 2015). The education system not only has the potential to build children's capacities, it also can play a vital role in shaping children's understanding of gender roles and stereotypes and in internalizing positive gender norms during childhood and into adolescence. Conversely, education that legitimizes harmful gender stereotypes and provides inequitable services, biased textbooks, and biased teaching methods can reinforce exclusion and stereotypes and threaten access to education and education quality, thereby undermining children's ability to contribute to peacebuilding.

Although research suggests that education can contribute to gender equality in conflict-affected environments and fragile states (Baranyi and Powell 2005, 2; Winthrop and Kirk 2008, 647), there is only limited evidence of what works to promote gender equality in education in conflict-affected settings. To address this evidence gap, more rigorous experimental and quasi-experimental evidence is needed on the topic. Furthermore, programs that aim to improve gender equality often focus on results that are hard to quantify, such as gender norms and women's empowerment (Burde, Guven, Kelcey, Lahmann, and Al-Abbadi 2015, 3). Therefore, it is crucial to supplement quantitative research with in-depth qualitative research in order to gain a better understanding of how programs work and affect these outcomes.

This paper helps to reduce this knowledge gap by focusing on the impact of an eight-month gender-socialization training program for teachers in Karamoja, Uganda. We conducted a cluster-randomized controlled trial (RCT) that randomly assigned 35 schools to a control group, 35 schools to the UNICEF-supported teacher-training program (Treatment 1), and 35 schools to the same program but with the addition of complementary text messages that reinforced training information (Treatment 2). In line with our theory of change, we estimated the impact the program would have on teachers' knowledge, attitudes, and self-reported practices around gender equality. We administered structured teacher surveys during baseline (March 2015) and endline (November 2015) data

collection. To increase our understanding of the program's intangible results, we supplemented the quantitative analysis with the collection and analysis of qualitative data.

We found evidence that the program had positive effects on teachers' knowledge about the difference between gender and sex, and on their attitudes toward gender roles and gender identity. We found no quantitative evidence for any short-term change in teachers' self-reported practices as a result of the program, or of any additional impact for the text-message component. Qualitative research showed that attitudes supportive of gender equality did not always align with the traditional ideas of gender roles in the larger community, which was a challenge to gaining broader acceptance of the concepts. These findings suggest that, while teacher training can influence knowledge and attitudes toward gender equality, traditional gender norms can be a barrier to changing behavior in the short term.

Such behavioral changes are important in Uganda, where sharp education and gender inequalities persist, particularly in the northeastern region of Karamoja. The primary school completion rate in Uganda is 64 percent, enrollment in lower secondary school is 34.9 percent, and enrollment in upper secondary school is 15.1 percent. Girls are more likely than boys to drop out of school at the higher levels (Pham, Vinck, and Gibbons 2015, 21). Karamoja is overrepresented in the country's lowest development indicators, particularly in education. It has the highest proportion of girls who are not in school or have never been to school, and the highest child mortality and poverty rates—75 percent of households in the region live below the official poverty line (Ministry of Gender, Labour, and Social Development and UNICEF 2015). Statistics show that average years of schooling in Karamoja is as low as three, that there is a 37:1 pupil-teacher ratio, and a 108:1 pupil-classroom ratio (Pham et al. 2015, 21).

The context of our study, therefore, is one in which improvements in gender equality are urgently needed. Through our rigorous experimental mixed-methods research, we contribute to the knowledge on what works to improve gender equality in education and peacebuilding, with special reference to conflict-affected settings. To the best of our knowledge, this is the first cluster-RCT study of a program for gender socialization in schools within a conflict-affected setting, one that investigates whether gender-sensitive approaches can be introduced into teacher training to reduce gender stereotypes, improve gender equity, and promote peacebuilding. It is also one of the first RCTs that uses a mixed-methods design to determine the effects of development programs in such a setting.

CONTEXT

Girls and women in Karamoja endure a hostile environment, due in part to the 20-year conflict between the rebel Lord's Resistance Army (LRA) and the Ugandan government, which lasted until 2006, when the LRA's power declined and peace talks began. During this conflict, tens of thousands of youth were abducted and forced to serve as soldiers, and many girls were victims of forced marriage to combatants, which resulted in deep social and psychological trauma. Children who were abducted often did not return to school, their wealth status later in life was lower, and they reported more symptoms of emotional distress than peers of the same age and location who were not abducted (Annan, Blattman, Mazurana, and Carlson 2011, 889; Blattman and Annan 2010, 882; Opinia and Bubenzer 2011, 5).

In addition to this national conflict, Karamoja has a history of recurring conflict between ethnic groups. The region includes seven districts inhabited by at least ten different ethnic groups. Their conflicts are the result of internal economic and social tensions, which often revolve around livestock, particularly cattle. Cattle ownership is a determinant of both social and economic status (Anderson and Broch-Due 1999), and cattle raiding has therefore been prevalent in Karamojong communities (Mkutu 2008). The conflicts are also the result of the region's predominantly pastoral lifestyle. During the dry season, communities tend to migrate to neighboring districts in search of pastures and water for their livestock, which sometimes escalates into border conflicts between tribes that are exacerbated by the different groups' proximity to one another.

Land disputes are also one of the most common sources of conflict in the region and one of the most difficult to resolve. A 2010 study conducted in four districts in northern Uganda with a representative sample of adults in those four districts reported that 63 percent of all disputes were related to land, and only 48 percent had been resolved by the time of the survey, compared to more than 75 percent of other disputes (Pham and Vinck 2010, 28).

Traditional views on gender roles in Karamojong society help to perpetuate violence. Faced with a lack of resources and influenced by traditional ideas about the male breadwinner, men often feel pressured to demonstrate their masculinity by raiding cattle, which fuels violent fighting between clans (Instituto da Defesa Nacional 2013). Men's failure to fulfill their traditional role as provider has also led to psychosocial problems and tension between men and women. These tensions can result in alcoholism, violence against intimate partners, and increased violence

against women in general (Specht 2013). In addition, the high number of cattle required to win a bride encourages cattle raiding (Vaughn and Stewart 2011).

There is a disjuncture between Uganda's formal education system and Karamojong norms. Some Karamojong simply reject formal education (Saminsky 2010). They believe that sending boys to school prevents them from gaining intimate knowledge of their herds, which is where most boys will earn their livelihoods. Girls are similarly expected to perform housework, which has little perceived correlation with what is taught in the classroom. In circumstances where parents must decide whether to send their sons or daughters to school, the girls are often left at home to learn domestic work, marry, and have children.

LITERATURE REVIEW

Studies of conflict and peacebuilding theorize that education can contribute to gender equality (Winthrop and Kirk 2008). School lesson plans and classroom conduct may play a key role in the “transmission or elimination of discrimination” (Duncan 2004, 21). For example, teacher training may promote peace by discouraging hostility, and the curriculum can provide positive models of masculinity and femininity that prevent the exacerbation of inequality (Knutzen and Smith 2012). However, education also can undermine gender equality through several mechanisms. Aikman and Unterhalter (2005), for example, state that a lack of adequate toilet facilities and running water can create a barrier to attending school, especially for girls during menstruation. They report further that schools with a limited number of female teachers can create a barrier to girls' attendance, as the presence of female teachers tends to be associated with more girl-friendly environments.

The nature of teachers' pedagogy and teacher-pupil relations can also play a critical role in promoting gender-equitable attitudes and behaviors in their students. For example, teachers and schools can transmit negative gender stereotypes by giving boys more attention in the classroom. Teachers also can undermine the learning experience by using biased language in the classroom, which reinforces gender differences and inequalities. National curricula and textbooks sometimes promote gender stereotypes that lead to gender inequality. Textbooks with stereotypical images of men and women—for example, with women depicted as mothers and housewives while men are portrayed in adventurous and influential roles—are still common in many countries (Aikman and Unterhalter 2005, 42). Thus, the challenge is not only to change the curriculum content to be more gender sensitive

but also to improve teacher training so teachers are adequately prepared to deliver it (Aikman and Unterhalter 2005, 42).

The literature indicates that a first step toward gender equity and peacebuilding should be to identify and transform the widely held norms underlying gender identities and the relations between women and men that reinforce damaging gender and sexual stereotypes (Strickland and Duvvury 2003, 7). However, the peacebuilding community is uncertain how to design programs that address discriminatory gender norms and practices that disadvantage women (Strickland and Duvvury 2003, 8).

The rigorous evidence base on which education programs are effective in conflict-affected settings is very limited, particularly on those that reduce gender inequality. Most claims about the relationship between education and gender equality are based on correlational studies whose designs do not allow for addressing counterfactual questions. A systematic review of the literature on education in crisis-affected contexts identified five cluster-RCTs and eight quasi-experimental impact studies, none of which focused on outcome measures associated with gender equality (Burde et al. 2015). For example, an RCT in Afghanistan found strong evidence that introducing village schools in Afghanistan resulted in a 52-percentage-point increase in girls' enrollment in education and an increase in average test scores of 0.65 standard deviations (Burde and Linden 2013). Another study in Uganda found evidence that community monitoring of the education service provision increased teacher and pupil attendance (Barr, Mugisha, Serneels, and Zeitlin 2012), while a study in Burkina Faso showed that providing schools with "girl-friendly" water, sanitation, and hygiene facilities increased enrollment by 13 percentage points (Kazianga, Levy, Linden, and Sloan 2013).

Most impact evaluations of education programs in conflict-affected settings rely on quantitative designs, and few triangulate the results with qualitative methods (Burde et al. 2015). It thus remains unclear how and why effective programs influence education outcomes. This constraint limits the lessons to be learned from impact evaluations. The lack of mixed-methods research is problematic, because education programs in conflict-affected settings often focus on intangible results that are hard to measure using quantitative research alone (Puri, Aladysheva, Iversen, Ghorpade, and Brück 2015).

THE PROGRAM

The Gender Socialization in Schools pilot program was developed by UNICEF and the Ugandan Ministry of Education, Science, Technology and Sports (MoESTS), which implemented it in partnership with Development Research and Training and the Forum for African Women Educationalists. The pilot was part of UNICEF's Learning for Peace program, which is founded on the idea that education and other social services have strong potential to foster social cohesion and enhance human security in countries affected by and emerging from violent conflict.

The training was organized in two stages. First, the implementing partners provided a three-day training of trainers in March 2015 for the coordinating center tutors (CCTs), district school inspectors, and MoESTS personnel. The training explained theoretical concepts of gender, conflict, and peacebuilding using participatory approaches, such as role playing, discussions, and storytelling, that incorporated familiar examples from Karamoja. Second, trained CCTs and inspectors delivered a three-day training for one thousand teachers at central locations in five districts of Karamoja.¹ The training aimed to empower primary teachers as agents of change, promote positive models of masculinity and femininity, and redress gender biases and question social norms. Moreover, the training aimed to create awareness of alternative norms and practices related to gender equality, build skills to help engage pupils in constructive dialogue, and provide materials to foster a shift in gender attitudes and beliefs and promote gender-sensitive practices in the classroom (Development Research and Training 2015). In August/September and November 2015, teachers received refresher trainings to reinforce content.

Between April and November 2015, a subset of 276 trained teachers received 13 text messages from UNICEF via the SMS platform GenderTrac. Each text message contained reinforcing reminders for teachers about certain content covered in the training and provided examples of good practices, such as promoting an equitable school environment, mechanisms for conflict resolution in school, positive discipline, and gender-responsive leadership and management.

¹ The training modules covered topics such as key gender concepts (e.g., the definitions of gender, sex, gender socialization, gender identities, gender roles, gender equity); gender-responsive teaching approaches (e.g., gender-responsive classroom set-up, gender-responsive language used in classrooms, gender-responsive content delivered by the teacher); gender-responsive learning materials; and gender-responsive classroom interaction.

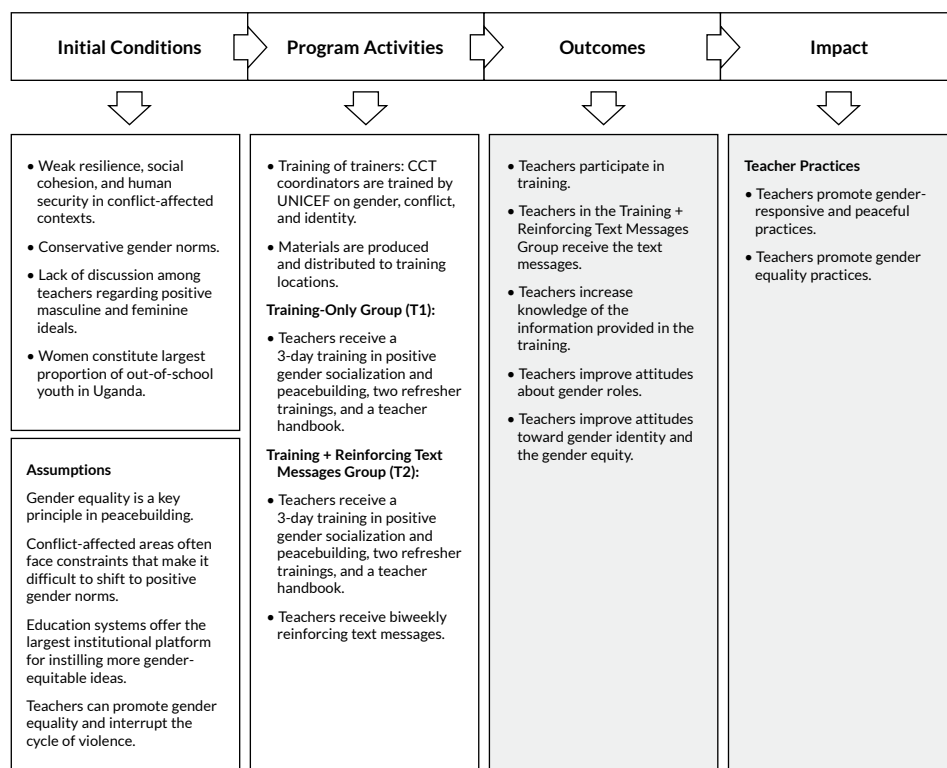
CONCEPTUAL FRAMEWORK

We defined our hypotheses and research questions based on a theory of change that we co-constructed with UNICEF, which maps out the causal chain among activities, outputs, outcomes, and impact, as well as the initial contextual conditions and the assumptions underlying the theory of change (White 2009). The theory of change hypothesizes that generating training materials, conducting teacher trainings, and sending reinforcing text messages about the importance of gender-sensitive education would result in greater knowledge about the difference between gender and sex, and about the relationship between gender, identity, and conflict, which could in turn improve teachers' attitudes toward issues of gender equality. These changes could increase the likelihood that teachers would use gender-responsive and peaceful practices.

It is important to note that the program focuses on both men and women, which makes the program a relational rather than a single-sex exercise. We hypothesized that this approach would make the program more sustainable, for three reasons. First, including gendered attitudes toward men and masculinity can mitigate male alienation and backlash (e.g., Barker and Schulte 2010; de Hoop, van Kempen, Linssen, and van Eerdewijk 2014; Dworkin, Dunbar, Krishnan, Hatcher, and Sawires 2011). Second, the program's emphasis on positive models of masculinity (as well as femininity) could help to engage men as partners in women's empowerment trajectories. And third, because men are also disadvantaged by norms of negative masculinity (e.g., expectations of participation in raids, fighting, violence), focusing on positive models of masculinity may benefit men and make the communities they live in more peaceful.

Several assumptions underlie the theory of change. First, gender equality is a key principle of building sustainable peace. Second, limited socioeconomic and political progress constrain positive shifts in gender norms in conflict-affected areas. Third, education systems offer an institutional platform for instilling gender-equitable ideas and exposing children to positive gender norms. Fourth, teachers, who themselves may be affected by gender bias or perpetuate it, have the capacity to become agents of change by promoting positive visions of masculinity and femininity. The theory of change for the Gender Socialization in Schools program is depicted in Figure 1.

Figure 1: Theory of Change for the Gender Socialization in Schools Pilot Program



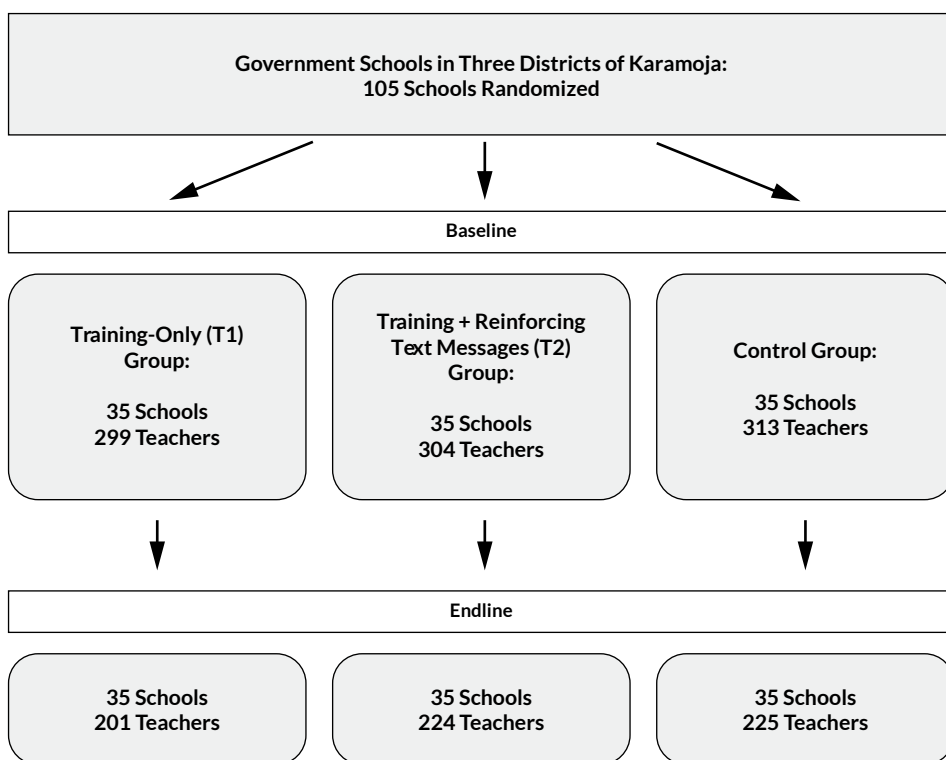
METHODS

RESEARCH DESIGN

The evaluation employed a cluster-RCT in which schools were randomly assigned to the teacher training alone (Treatment 1), the teacher training and reinforcing text messages (Treatment 2), and the control group. We stratified the randomization by the CCTs’ catchment areas, thus ensuring that each tutor’s area included each treatment and control condition. Increasing the geographic proximity of the schools assigned to the study groups was important in accounting for key social norms related to the outcomes of interest and to increase comparability across the groups. To minimize spillovers and contamination, the implementers encouraged local education authorities to minimize information-sharing with the control schools.

The sample was limited to 105 government schools in three of the seven districts in Karamoja, which were evenly assigned to each of the three groups (35 schools in each group).² All teachers in the selected schools were invited to participate in the study. This randomization process led to 304, 299, and 313 teachers, respectively, from whom we collected baseline data. Power calculations that took into consideration the nested structure of the evaluation suggest that the study had 80 percent power to detect a minimum effect size of 0.28 standard deviations.³ Figure 2 presents the flow diagram of the RCT design.

Figure 2: Flow Diagram of the Randomized Controlled Trial Design



² These three districts were selected as follows. First, two of the seven districts were excluded by implementers because they were already benefiting from several other education programs. Second, to minimize the length and cost of data collection, the research sample was selected from the three districts where training was scheduled to happen first. Within the CCTs' catchment areas in each of the three districts, we first randomly sampled the maximum number of schools that was a multiple of three, and then we randomly assigned those schools to each of the three study groups to have a balanced design.

³ The intra-class correlation for the different outcome measures was approximately 0.05. We conducted all power calculations using Optimal Design software.

QUANTITATIVE TEACHER SURVEY

We developed the survey based on a comprehensive literature review and a review of best practices in measuring outcomes depicted in the theory of change. The items are consistent with other inventories used to measure gender attitudes.⁴ The instrument was then refined on the basis of several consultations with UNICEF, MoESTS staff, and key stakeholders, followed by three pilots of the instrument in the field. During these pilots, we paid close attention to the length of the survey, the respondents' understanding of the survey questions, and some basic psychometric properties of the instruments. These pilots enabled us to create an instrument informed by the specific contextual characteristics of Karamoja. During this time, we also added vignettes to the questionnaire that we developed in collaboration with local Karamojong people (e.g., teachers, education officials, local data collectors). We did not rely on previously validated tests because they were not available for the context of Karamoja or the specific goals of the program. The final survey included questions related to gender norms, the division of household and labor duties between men and women, and differences in boys' and girls' educational opportunities and experiences.

We used vignettes to measure teachers' attitudes toward gender norms on topics such as gender roles in the household and sexual and physical violence. These vignettes described a fictional scenario and were typically used to determine the ways people make judgments and decisions about sensitive topics. Using vignettes can reduce the likelihood of courtesy and social desirability bias (White and Phillips 2012). Finally, survey items were worded to be consistent with Ugandan proverbs, folklore, and literature on the roles of men and women in Karamoja.

In the final version of the questionnaire, we used two groups of questions to measure teachers' knowledge of the information provided in the training. The first questions captured whether teachers understood the difference between gender—a social construction regarding the roles of women, girls, men, and boys—and sex, which refers to the biological characteristics of being female or male. The second group of questions focused on knowledge about other topics covered in the training, including gender-sensitive lesson planning, the legal framework for equal access to education, and the relationships among gender equality, peacebuilding, and social cohesion.

⁴ Examples of these inventories include Ashmore, Del Boca, and Bilder (1995); Glick and Fiske (1997); Baber and Tucker (2006); and Pulerwitz and Barker (2008).

We created several index scores to measure teachers' attitudes toward gender roles, gender identity, gender equality in schools, and several dummy indicators to measure sexual harassment of girls by boys, violence by boys against girls, and violence by girls against boys. The gender-role indexes measured the shared expectations of behavior based on a person's gender, while the gender-identity index aimed to capture how masculine and feminine teachers see themselves in terms of what it means to be a man or a woman in their society. The items in the gender-equality construct captured attitudes toward gender equality in school and whether teachers' expectations for girls and boys were similar. The binary variables measured whether teachers punished boys for sexually harassing girls, whether teachers punished boys or girls for behaving violently toward the opposite sex, and whether teachers punished boys and girls equally for behaving violently toward the opposite sex.

Finally, we generated two indexes to approximate teachers' practices. The first index measured teachers' gender responsiveness when planning and implementing activities and exercising discipline in the school, while the second measured teachers' practices associated with gender equality (e.g., we asked teachers whether they assigned more difficult tasks to boys or easier tasks to girls, and whether they used the same strategies to teach girls and boys).

The survey also included questions related to the school's cultural practices, such as relationships between teachers and students, relationships between boys and girls, and school clubs. We also included several questions to capture teachers' demographic and teaching backgrounds and basic school characteristics, such as the number of teachers and students, and information on infrastructure and available services. Table 1 summarizes the contents of the teacher questionnaire.⁵

⁵ To construct the outcomes of interest, we generated an index and a scale. Correct knowledge answers or progressive attitudes were coded as 1, and the index was computed as the summation of these answers. Practices were coded using a scale of 1 to 4, where responsive and peaceful practices received higher scores. The scale was created through a factor analysis in which we constructed weights from the matrix of factor loadings. The Cronbach's alpha coefficients, which measured the internal consistency of the outcome indexes, ranged from 0.62 to 0.73 for the knowledge indexes, from 0.64 to 0.88 for the attitudes indexes, and from 0.68 to 0.72 for the practice indexes. Finally, for the descriptive indexes, the Cronbach's alpha coefficients ranged from 0.71 to 0.82. Evidence suggests that Cronbach's alpha coefficients smaller than 0.7 may be an indication of less reliable scales. We did some robustness checks in which we conducted the same impact analysis for individual items. These analyses suggest that the results of the study are robust to the use of individual items as opposed to indices or scales. These results are included in the report created for UNICEF (Chinen et al. 2016). In this paper, we present the program's impact on the index because the interpretation of these results is more intuitive. Nonetheless, the results are robust to the use of scales as outcome measures.

Table 1: List of Outcome Indexes and Indicators Created from the Teacher Survey

Outcome Indexes		Description
1	Knowledge about the difference between gender and sex	Teacher understands the difference between gender and sex.
2	Knowledge about gender, identity, and conflict	Teacher understands issues of gender, identity, and conflict.
3	Attitudes toward gender roles, Index 1	Teacher believes men and women are equally capable of conducting jobs that are traditionally associated with one gender.
4	Attitudes toward gender roles, Index 2	Teacher supports textbook image that shows a father in a caretaker role.
5	Attitudes toward gender roles, Index 3	Teacher does not oppose being seen conducting activities traditionally associated with women.
6	Attitudes toward gender identity	Teacher disagrees with statements describing very traditional masculine stereotypes.
7	Attitudes toward gender equality	Teacher agrees with statements associated with gender equality.
8	Gender-responsive and peaceful self-reported practices	Teacher conducts gender-responsive and peaceful practices in the classroom, when planning, implementing activities, managing behavior, and exercising discipline in the school.
9	Gender equality self-reported practices	Teacher conducts activities to promote gender equality in the classroom.
Additional Dummy Indicators		
10	Attitudes: Reactions to sexual harassment (vignettes)	Teacher intervenes in scenarios of sexual harassment. Teacher blames harasser for harassment. Teacher punishes harasser. Teacher opposes violent retaliation to sexual harassment.
11	Attitudes: Fair punishment to sexual harassment (vignettes)	Teacher punishes males and females correctly and equally.
12	Attitudes toward violence (vignettes)	Teacher intervenes in scenarios of classroom violence.
Secondary and Long-Term Indexes		
13	Gender and culture in schools	Teacher identifies positive gender culture in the school.
14	Problems in the school environment	Teacher identifies many gender-based problems in the school environment.
15	Teachers' sense of self-efficacy	Teacher feels capable of solving gender-based problems in the school environment.

Data collection in the treatment group took place in the teacher-training locations on the first day, before the teachers started to receive the training. All teachers who attended the training agreed to complete the survey. The data collectors organized the teachers in classrooms, explained the goals of the study, read the consent agreement aloud, and stayed in the classroom to answer questions or clarify aspects of the survey. Data collection in the control group took place in schools from Monday to Thursday during the same weeks the intervention group data was collected. All teachers in the control schools agreed to participate in the survey. As with the treatment group, the data collectors gathered the teachers in classrooms and followed the same protocols, thus standardizing the data-collection process for the two groups to the extent possible.⁶

QUALITATIVE INSTRUMENTS

We collected qualitative data from the CCTs who were present during the training and from head teachers whom the implementing partners identified as “teacher leaders,” based on their participation in the training and their interactions with other participants. We conducted one-hour semi-structured interviews at midline data collection with 15 intervention CCTs and 8 head teachers. The purpose of these interviews was to understand leaders’ experiences with (a) implementation of the training, (b) school-level follow-up on the training, and (c) the level of understanding of training concepts.

In order to facilitate a deeper understanding of the factors that enabled, or inhibited, the uptake of gender-equitable practices and peaceful conflict resolution as a result of this program, we collected qualitative data from teachers and students in a random selection of six intervention and two control schools across the three districts and the three study groups. Endline focus group discussions (FGDs) with the treatment teachers built on midline discussions, primarily about the training itself and about the specific challenges the teachers had experienced in applying practices. The endline FGDs specifically targeted knowledge, attitudes, and practices that might have changed as a result of the program. The FGDs with two control schools allowed us to find contrasts between treatment teachers and control teachers. The interviewers facilitated the approximately two-hour FGDs using a guide, which included a flexible set of questions and probes intended to invite the participants to steer the discussion toward the issues that interested them while ensuring that they remained focused on relevant topics. We organized separate focus groups for male and female teachers so participants would feel

⁶ The study team received consent from all adult participants and handled the quantitative and qualitative data according to procedures and protocols approved by the institutional review board.

free to express their authentic experiences. We targeted six participants in each focus group, although some groups were smaller because of the low number of female teachers.

ANALYSES

We examined the main effects the interventions had on the outcomes of interest using an analysis of covariance (ANCOVA) model. The ANCOVA model allows for an estimate of the causal effect of the program by comparing outcomes in the intervention schools with outcomes in the control schools, and controlling for the value of the outcome variable at baseline. The main advantage of the ANCOVA model, as opposed to difference-in-difference analysis, is that the ANCOVA model increases statistical power, particularly when outcomes are not strongly autocorrelated (McKenzie 2012, 211). The model was particularly appropriate for this study because we changed a few items and the wording of a few questions between the baseline and the endline, based on feedback and analyses of the baseline data. Using an ANCOVA model enabled us to use the original index as control variables, despite the changes in the wording. This would not have been possible using a difference-in-difference model. The ANCOVA model used cluster robust standard errors at the school level to account for the nested structure of the data.

We checked the robustness of the treatment estimates against several different model specifications that included a different set of covariates, and found that the impact estimates were robust to the specification of the regression.⁷ The impact results for the statistical model that only controlled for the pretest score are presented in this paper. We also investigated possible selection bias due to missing data, because 29 percent of the teachers who participated in the baseline were not available to complete the endline survey. For this reason, we examined whether the percentage of teachers with missing data was similar or different across the three study groups; whether teachers with complete baseline and endline data were equivalent in terms of observed covariates collected at baseline across the three study groups; and whether teachers with missing data were similar at baseline to those with complete data. These analyses revealed that the percentage of teachers with missing data were similar across the three groups, and that teachers with completed data at both baseline and endline were equivalent in the vast majority of

⁷ We specified a total of six model specifications for each outcome variable. The first model included only the treatment indicator; the second one added the pretest score; the third one added the district fixed effects; the fourth included teacher characteristics such as gender, religion, and education; the fifth model added school characteristics, such as the number of female teachers and enrollment size; and the sixth combined treatments 1 and 2, in light of the lack of differential treatment effects across the two interventions.

observed characteristics gathered in the teacher survey.⁸ We also investigated the possibility of spillovers (or the possibility that the interventions affected control teachers), and found limited evidence of spillover effects.⁹

Finally, we explored the possibility of heterogeneous treatment effects by teachers' sex. The robustness analyses and complete impact regression results are presented in the endline report we developed for UNICEF (Chinen et al. 2016). To analyze the qualitative data, we used methodological triangulation and triangulation among raters to ensure data trustworthiness and credibility of findings. The research team utilized grounded theory (Glaser and Strauss 2009) to guide analysts trained in NVivo qualitative data analysis Software (QSR International Pty Ltd., Version 10, 2012). Grounded theory utilizes qualitative data to deduct a new theory about the findings, rather than testing an existing one. Combining the grounded theory with a rigorous impact-evaluation design enabled us to triangulate the research findings through mixed methods. The quantitative research served to test predefined hypotheses, while the qualitative research enabled us to gain a better understanding of why the program positively influenced some, but not all, outcomes of interest.

Three raters separately coded the text data to independently identify the themes in the discussion. These themes formed the coding structure used to categorize raw data from the interviews and focus groups and identify themes about the primary findings. An inter-rater reliability test showed that the three primary coders had an overall average of 99.1 percent agreement, indicating a high level of consistency among the researchers in interpreting the data and clarity of the coding scheme.

8 We called absent teachers and inquired why they missed the endline survey. The three main reasons were the following: (1) teachers were not informed about the teacher training or came late to training (in the case of intervention teachers); (2) teachers were sick during the data collection; and (3) teachers were busy grading primary leaving exams. A list of all the reasons and all the missing analyses are presented in the endline report (Chinen et al. 2016).

9 Approximately 33 percent of control teachers (or 76 teachers) indicated receiving some training in gender, conflict, and peacebuilding, and 20 percent reported receiving another training in gender issues. Moreover, 18 percent of control teachers reported receiving some coaching on gender, conflict, and peacebuilding, and 7 percent indicated receiving some text messages about gender, conflict, and peacebuilding. These results suggested the presence of spillovers from treatment to control teachers, which may have resulted in an underestimation of the impact of the intervention. However, we did not encounter additional evidence that control teachers attended the trainings. Besides, it is possible that some teachers confused the different gender trainings. Approximately 49 percent of the 76 control teachers who reported attending the Gender, Conflict, and Peacebuilding training also reported attending another training in gender issues.

RESULTS

BALANCE AT BASELINE

At baseline, we found that the treatment and control groups were statistically equivalent in the vast majority of observable characteristics. In other words, the randomization process successfully created equivalent groups before the intervention started. We found no statistically significant differences between treatment and control teachers. Tables 2 and 3 present descriptive statistics for teachers' demographic and background characteristics, as well as teachers' education characteristics. The descriptive statistics for all the survey covariates, including for the baseline outcomes, can be found in the "Baseline Report" generated for the study (Chinen et al. 2016).

Table 2: Baseline Teacher Demographics and Background Characteristics

Variable	Control		Training-Only (T1)		Training + Text (T2)		p-values		Diff (SD)	
	N	Mean	N	Mean	N	Mean	C = T1	C = T2	T1 - C	T2 - C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Gender: Female	312	0.25	299	0.24	303	0.27	0.89	0.63	-0.01	0.04
Ethnicity:										
Ethur	313	0.35	297	0.43	303	0.37	0.45	0.84	0.16	0.04
Other	313	0.32	297	0.26	303	0.29	0.20	0.48	-0.13	-0.07
Bokora Karimojong	313	0.11	297	0.09	303	0.13	0.69	0.74	-0.07	0.06
Religion:										
Catholic	313	0.61	299	0.67	304	0.62	0.26	0.97	0.11	0.00
Protestant	313	0.28	299	0.23	304	0.27	0.21	0.71	-0.11	-0.03
Other	313	0.11	299	0.10	304	0.12	0.95	0.63	-0.01	0.04
Native language:										
Ngakaramojong	313	0.23	299	0.25	304	0.25	0.73	0.74	0.06	0.05
Thur	313	0.18	299	0.20	304	0.16	0.64	0.84	0.07	-0.03
Ateso	313	0.17	299	0.14	304	0.17	0.29	0.96	-0.08	0.00
Luo	313	0.14	299	0.13	304	0.11	0.78	0.52	-0.03	-0.08
Marital status:										
Officially married	310	0.67	296	0.64	303	0.63	0.43	0.42	-0.07	-0.09
Unofficially married	310	0.21	296	0.26	303	0.31	0.18	0.03	0.12	0.21
Other main source of income: None										
Growing crops	308	0.25	292	0.28	302	0.24	0.50	0.74	0.07	-0.03
Growing crops	308	0.69	292	0.67	302	0.70	0.79	0.77	-0.03	0.03

CAN TEACHER-TRAINING PROGRAMS INFLUENCE GENDER NORMS?

(Table 2 cont.)

	Control		Training-Only (T1)		Training + Text (T2)		p-values		Diff (SD)	
	N	Mean	N	Mean	N	Mean	C = T1	C = T2	T1 - C	T2 - C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teacher's age	313	36.57	297	37.84	301	38.14	0.06	0.03	0.15	0.18
Years teacher lived in the district where born	312	30.61	296	31.92	299	32.25	0.18	0.09	0.11	0.14
Number of women married to	235	1.48	227	1.25	223	1.33	0.17	0.35	-0.17	-0.11
Total number of children	295	4.97	285	4.92	288	5.15	0.88	0.51	-0.01	0.05
Female children	296	2.35	287	2.48	292	2.52	0.51	0.29	0.06	0.08
Male children	295	2.62	288	2.42	288	2.63	0.24	0.92	-0.09	0.01

NOTES: Robust t-statistics clustered at the school level.

Table 3: Baseline Teacher Educational Characteristics

Variable	Control		Training-Only (T1)		Training + Text (T2)		p-values		Diff (SD)	
	N	Mean	N	Mean	N	Mean	C = T1	C = T2	T1 - C	T2 - C
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Teaching language:										
English	313	0.91	299	0.92	304	0.94	0.47	0.22	0.06	0.11
Swahili	313	0.02	298	0.00	304	0.01	0.21	0.37	-0.14	-0.10
Local language	313	0.42	298	0.46	304	0.39	0.39	0.54	0.08	-0.05
Opinion matters in the community where teaching										
Meet with pupils' parents:										
At least once a month	313	0.28	296	0.31	300	0.38	0.53	0.03	0.06	0.20
Once per term	313	0.52	296	0.45	300	0.42	0.07	0.03	-0.14	-0.20
Received any previous training on gender issues										
Years teaching	313	11.43	298	12.21	302	12.72	0.27	0.11	0.09	0.15
Years teaching in that school	300	4.17	290	4.86	292	5.20	0.02	0.02	0.17	0.25
Total grades taught	313	2.30	298	2.39	304	2.21	0.49	0.50	0.08	-0.08
Minutes to go from home to school	302	16.41	274	19.62	273	18.29	0.17	0.36	0.16	0.09

NOTES: Robust t-statistics clustered at the school level.

IMPACT ON KNOWLEDGE

Quantitative evidence indicated that the program resulted in an increase in teachers' knowledge of the information provided in the training. We found evidence for a positive and statistically significant program impact on teachers' knowledge about the difference between gender and sex. The point estimates were 0.35 ($p < 0.1$) and 0.60 ($p < 0.01$) for the training-only and training-plus-texting groups, respectively. This finding indicates that the effect for the training-plus-texting group was almost twice the effect for the training-only group, but this difference was not statistically significant ($p < 0.22$). We found no positive effects on the second indicator of knowledge about gender, identity, and conflict. Table 4 presents the impact on knowledge.

Qualitative evidence indicated similarly that, at endline, teachers were better able to use the concepts of gender and sex consistently with the training program's definitions (e.g., gender sensitive, gender socialization). Teachers also understood that classroom discrimination based on gender identity could affect social interactions, girls' self-confidence, and their feeling that they need to skip school during menstruation. In describing the need to address the "stigma" of menstruation, one teacher said, "I even talk to them about fear; they should not have fear when they are undergoing their menstruation period." Finally, teachers identified ways to ensure a more gender-sensitive environment, including the classroom set-up, equal participation and representation in lessons, and shared responsibility.

However, qualitative data also suggested that challenges remained. One participant pointed to the difficulty of learning the concepts: "There were certain concepts that were . . . difficult to define or explain; [for example], when it came to things like gender disparity, gender equality, and all those concepts." The same participant also pointed out the value of refresher trainings in helping to reinforce concepts and ensure understanding of basic concepts: "When we first started with the first training, it was like [we] were trying to understand those concepts about gender, but the second one was so much [more] interesting and most participants [expressed interest in having] another time of really going through those concepts again."

Table 4: Impact on Knowledge Outcomes

Knowledge	Control Mean (Endline)	Training-Only (T1)			Training + Reinforcing Text Messages (T2)		
		Program Impact	SE	Effect Size	Program Impact	SE	Effect Size
Understands the difference between gender and sex	4.94	0.35*	0.18	0.17	0.60***	0.17	0.30
Understands issues of gender, identity, and conflict	15.77	0.13	0.22	0.05	-0.04	0.21	-0.02

NOTES: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Adjusted regression results using ANCOVA OLS controlling for pretest score. Cluster robust standard errors at the school level. Sample size ranges between 566 and 630, depending on outcome.

IMPACT ON ATTITUDES

Overall, the quantitative and qualitative evidence indicated that the program resulted in teachers having more positive attitudes toward gender roles and gender identity. We found positive and statistically significant program effects on all three indicators of teachers' attitudes toward gender roles, measuring shared expectations of behavior given one's gender. The results for all attitudes indexes are presented in Tables 5 and 6.

The first gender-role index was created by asking teachers whether women or men are more or equally capable of doing jobs that are traditionally associated with one gender. Intervention teachers were more likely to agree with statements that implied relatively progressive attitudes toward gender roles. For example, intervention teachers were more likely to agree with statements suggesting that men and women are equally capable of doing jobs traditionally associated with one gender (e.g., engineer, mechanic, nurse, politician). The ANCOVA point estimates on the full scale (which ranged from 0 to 10 points) were 0.83 ($p < 0.01$) for the training-only group and 0.48 ($p < 0.05$) for the training-plus-texting group, both of which were statistically significant.

The second gender-role index was created by presenting a hypothetical vignette in the proposed new English textbook. It showed a father cooking dinner and looking after his baby to capture teachers' attitudes toward traditional gender roles (see Figure 3). The teachers were then asked whether they would support having such a picture in the textbook, whether this was an example of gender equality that they would promote, and whether they would think this man's wife was treating him well if he were their brother. We found that intervention

teachers were also more likely to show more progressive attitudes toward gender roles depicted in the hypothetical situation. The ANCOVA point estimate for the index (which ranged from 0 to 3) was 0.44 ($p < 0.01$) for the training-only and the training-plus-texting groups, which were both positive and statistically significant.

Figure 3: Hypothetical Situation Depicted in a Vignette in the Proposed New English Textbook Showing a Father Cooking Dinner and Looking after His Baby



The third gender-role index measured teachers' attitudes toward conducting activities traditionally associated with women. This index was considered relevant, given that most of the teachers (75 percent) were men. For example, intervention teachers were more likely to disagree with statements such as, "I would not want my friends to see me washing women's clothes" and "I would not want my friends to see my spouse correcting me in public." The pattern of responses for other items was more similar among the three groups. The regression point estimates for the index scale (which ranged from 6 to 24 points) were 0.59 ($p < 0.1$) for the training-only group and 0.57 ($p < 0.01$) for the training-plus-texting group, both of which were positive and statistically significant.

Positive quantitative effects were also found for teachers' attitudes toward gender identity, which refers to how male or female teachers see themselves relative to what it means to be a man or a woman in their society. Intervention teachers were more likely to disagree with very traditional masculine stereotypes, including such statements as, "Some women need to be beaten," "Educated women make unruly wives," "When you beat boys, you raise disciplined men," "When men are speaking, serious women are not supposed to talk." The ANCOVA point estimates for the full index scale (which ranged from 22 to 52 points) were 1.26 ($p < 0.01$) for the training-only group and 0.62 ($p < 0.1$) for the training-plus-texting group.

We found limited and no evidence for positive program effects on the index we used to measure attitudes toward gender equality, and on attitudes toward sexual harassment and on punishment for sexual harassment or violence. We used vignettes to understand the action teachers might take in a hypothetical situation about sexual and physical violence, and in a conflict between boys and girls in the classroom. The vignettes included situations in which a boy inappropriately touched a girl, a girl was physically violent, and men were showing behavior usually associated with women. Although the direction of the point estimates was generally positive, the results were generally not statistically significant. Thus, the findings suggested little evidence of changes in reactions to sexual harassment or in the punishment of students for sexual harassment or violence.

Table 5: Impact on Attitudes toward Gender Roles, Gender Identity, and Gender Equality

Attitudes	Control Mean (Endline)	Training-Only (T1)			Training + Reinforcing Text Messages (T2)		
		Program Impact	SE	Effect Size	Program Impact	SE	Effect Size
Gender Roles Index 1: Believes women and men are equally capable of doing traditionally gendered jobs	6.15	0.83***	0.20	0.34	0.48**	0.20	0.20
Gender Roles Index 2: Supports textbook image of father in a caretaker role	1.48	0.45***	0.11	0.35	0.44***	0.11	0.35
Gender Roles Index 3: Does not oppose being seen conducting activities traditionally associated with women	19.18	0.60*	0.30	0.20	0.57**	0.27	0.20
Gender Identity: Disagrees with statements of masculine stereotypes	38.90	1.26***	0.33	0.24	0.62*	0.33	0.12
Gender Equality: Agrees with statements of gender equality	46.52	0.702*	0.36	0.20	0.36	0.39	0.10

NOTES: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Adjusted regression results using ANCOVA OLS controlling for pretest score. Cluster robust standard errors at the school level. Sample size ranges between 548 and 613, depending on outcome.

Table 6: Impact on Attitudes toward Sexual Harassment, Punishment for Sexual Harassment or Violence

Attitudes: Additional Dummy Indicators	Control Mean (Endline)	Training-Only (T1)			Training + Reinforcing Text Messages (T2)		
		Program Impact	SE	Effect Size	Program Impact	SE	Effect Size
Reactions to Sexual Harassment:							
Intervenes in scenarios of sexual harassment	0.92	-0.01	0.03	-0.04	0.01	0.02	0.02
Blames harasser for harassment	0.81	0.03	0.03	0.08	0.03	0.03	0.06
Punishes harasser	0.58	-0.11**	0.05	-0.22	-0.08	0.05	-0.16
Opposes violent retaliation to sexual harassment	0.77	0.05	0.04	0.13	0.05	0.04	0.12
Fair Punishment to Sexual Harassment:							
Punishes females and males correctly and equally	0.75	0.01	0.04	0.03	0.08**	0.04	0.18
Attitudes Toward Violence:							
Intervenes in scenarios of classroom violence	3.64	-0.33**	0.14	-0.18	-0.21	0.18	-0.12

NOTES: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Adjusted regression results using ANCOVA OLS controlling for pretest score. Cluster robust standard errors at the school level. Sample size ranges between 636 and 648, depending on outcome.

The qualitative data yielded similarly mixed results on attitudes. Teachers' basic attitudes changed, as they reported that boys and girls should be equal in responsibilities, work, and their futures. The majority of teachers said children should not be encouraged to participate in girl-only or boy-only activities and should share responsibilities. Several teachers observed that the trainings expanded their ideas about what girls could do in the classroom. One teacher said, "I went through a girl's school throughout my education . . . we used to say that boys do more work than girls, but with training and these techniques, we have come to learn that all these people are equal." Several also noted that, after encouraging girls in math, they saw the girls' performance improve—sometimes beyond that of the boys.

However, some of the teachers indicated that traditional attitudes continue to shape some of their approaches in the classroom. For example, one head teacher said, "Of course you know culture is part of the community, [and] as far as our community is concerned . . . there are certain responsibilities which are more of men than of ladies."

In sum, the program meaningfully changed teachers' attitudes toward gender roles, but these changes also created challenges for them, given an environment in which traditional gender norms heavily dictate children's roles and responsibilities.

RESULTS ON SELF-REPORTED PRACTICES

The program did not appear to influence overall teacher practices in the short term. No consistent quantitative evidence was found for positive effects on the two overall indexes of self-reported practices that were created from the survey. The intervention and control teachers gave similar answers to most of the questions included in the two indexes. These results are presented in Table 7.

The first index of self-reported practices, Gender-Responsive and Peaceful Practices, measured teachers' gender responsiveness when implementing activities, managing behavior, and disciplining boys and girls separately. For example, we asked teachers whether they rewarded girls for behaving appropriately, whether they ensured that girls used peaceful means to resolve conflicts with their peers, how teachers disciplined girls who misbehaved, whether teachers helped girls catch up on the lessons when they were absent for genuine reasons, etc. The same questions were also asked to inquire about their practices with boys. For the Gender-Responsive and Peaceful Practices index (which ranged from 12 to 48 points), the ANCOVA point estimate for the training-only group was -0.12, but it was not statistically significant. The point estimate for the training-plus-texting group was 0.57 ($p < 0.1$), which was statistically significant at the 10 percent level.

The second index of self-reported practices, Gender-Equality Practices, measured teachers' practices related to gender equality. For example, we asked teachers whether they assigned more difficult tasks to boys or easier tasks to girls, whether they discussed strategies for providing a safe learning environment for girls and boys with other teachers, whether they used the same strategies to teach girls and boys, whether they made sure girls and boys had equal opportunities to participate, etc. For the Gender-Equality Practices index (which ranged from 15 to 60 points), the ANCOVA point estimates were 0.07 for the training-only group and 0.39 for the training-plus-texting group. These effects were not statistically significant.

Table 7: Impact on Practice Outcomes

Practices	Control Mean (Endline)	Training-Only (T1)			Training + Reinforcing Text Messages (T2)		
		Program Impact	SE	Effect Size	Program Impact	SE	Effect Size
Conducts gender-responsive and peaceful practices in the classroom	36.45	-0.12	0.35	-0.30	0.57*	0.32	0.14
Conducts activities to promote gender equality in the classroom	53.25	0.07	0.34	0.02	0.39	0.29	0.09

NOTES: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Adjusted regression results using ANCOVA OLS controlling for pretest score. Cluster robust standard errors at the school level. Sample size ranges between 567 and 587, depending on outcome.

Qualitative data indicated that intervention teachers adopted some practices taught in the training. The teachers were best able to grasp more pragmatic training concepts, such as equitable classroom set-up, which mainly involved a mixed-gender seating arrangement. One teacher said, “If you have 10 boys and 10 girls, they should sit in [an] arrangement whereby a boy and a girl sit [in] the reading corners.” Some teachers reported fostering equal representation by dividing resources (such as textbooks) equitably between boys and girls, and encouraging equitable participation in class activities. Other teachers reported dividing classroom responsibilities between boys and girls, including leadership roles and classroom duties.

Several teachers explained that gender-sensitive lesson planning meant creating lessons that had objectives, activities, and examples that incorporated both boys and girls. One teacher said, “The technique that I am now applying in the classroom situation is . . . considering both boys and girls equally,” while another said, “Both boys and girls have to participate in the lesson.” Some teachers mentioned that they should tailor lessons to boys and girls, saying, for example, “I also learnt about friendly methods which can make a child really participate in an activity, and also [about] the instructional materials, which should be child friendly. The activity which is given should cater for all without any gender discrimination.” However, most teachers did not explain how they tailored lessons to male and female needs or connected gender-equitable practices to peacebuilding. In addition, data indicated that traditional practices such as corporal punishment are still used in the classroom. The short timeline of the program may have limited its ability to promote more complex changes in ideas and practices.

The teacher training promoted practices by drafting action plans that encourage a peaceful school environment for students. Although teachers seemed to have a basic understanding of the purpose and use of action plans, not all were aware of how each element was linked to gender. They also expressed difficulty meeting the goals of the plans because of larger issues related to poverty in their schools. One teacher said, for example, “I talked of the [seating] arrangement in our action plan, but the challenge we have faced with this is inadequate [seating] facilities.”

RESULTS ON THE EFFECT OF COMPLEMENTARY TEXT MESSAGES

We did not find consistent evidence across the different outcome measures that teachers who received reinforcing text messages in addition to the training activities earned more positive scores than teachers who only received the training. These results could be partially explained by the fact that 28 percent of the teachers in this group reported not receiving any text messages. Nonetheless, teachers on average reported receiving 13 messages related to the program, which is the number of reinforcing messages sent by UNICEF. The finding suggests that, at least in the first eight months of implementation in which they received three trainings, the SMS program did not bring additional benefits to the teacher training.

RESULTS ON SECONDARY AND LONG-TERM MEASURES

Our results showed no evidence of positive effects from the program on the measures Gender and Culture in Schools, Problems in the School, and Teachers’ Sense of Self-Efficacy to solve the most pressing problems of the school. The first outcome measured general aspects of the school culture. For example, we asked teachers whether they knew about their pupils’ families, whether they talked to children about their personal lives, whether they organized clubs, etc. The second outcome inquired about common problems in the school environment. For example, we asked whether hunger, absenteeism, or early marriage were common problems in their school. Finally, the third outcome attempted to measure whether teachers thought they were capable of resolving the problems listed in the previous outcome.

We did not expect any statistically significant effects of the program on these outcome measures because the program did not aim to affect them in the short term. Courtesy and social desirability bias might nonetheless have resulted in statistically significant differences between the treatment and control teachers. That we did not find these statistically significant differences suggests that courtesy

and social desirability bias did not create a high degree of bias in our positive impact estimates on knowledge and attitudes. These results are presented in Table 8.

Table 8: Impact on Secondary and Long-Term Outcomes

Secondary and Long-Term Measures	Control Mean (Endline)	Training-Only (T1)			Training + Reinforcing Text Messages (T2)		
		Program Impact	SE	Effect Size	Program Impact	SE	Effect Size
Identifies positive gender culture in school	40.97	0.75	-0.50	0.17	-0.02	0.41	-0.01
Identifies many gender-based problems in the school environment	5.33	0.07	0.26	0.03	0.03	0.24	0.01
Feels capable of solving gender-based problems in the school environment	30.80	0.18	0.53	0.04	0.54	0.61	0.11

NOTES: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. Adjusted regression results using ANCOVA OLS controlling for pretest score. Cluster robust standard errors at the school level. Sample size ranges between 574 and 576, depending on outcome.

The qualitative findings indicate that teachers recognized the need for greater support and found it difficult to independently obtain support from parents, politicians, and other community leaders. Although many teachers have made progress in their understanding of gender equality, the majority continue to have difficulty reconciling these concepts with traditional ideas of gender in the larger community. The disjunction between the training ideas and the deeply embedded community norms was evident throughout the data. Multiple teachers expressed difficulty promoting equal opportunity and sharing responsibility in a community where gender norms heavily dictated children's roles and responsibilities. One teacher described the difficulty with involving parents: "Problem comes when they disagree with you and I don't know which means we should use, because we can also not force them ... so the influence I think is on the ground."

Teachers said they should involve parents in gender and peacebuilding because, as one teacher noted, "it is from the community that [students] should learn first ... before they come to school." In response to the resistance teachers faced in implementing training concepts, some had already involved community members by holding PTA meetings and engaging school management.

Although few teachers seemed to have engaged students or the community in sensitization activities, several were able to describe avenues they could take

to sensitize stakeholders on particular issues. Several other teachers mentioned the importance of communicating rules and expectations about conflict in the classroom to the students.

RESULTS BY TEACHERS' SEX

Finally, we did not find evidence for statistically significant differences in the estimates of impact between male and female teachers. This lack of evidence for the heterogeneity of effects could be explained by the small number of female teachers in the schools (25 percent), which could have resulted in a lack of statistical power to detect subgroup effects.¹⁰

DISCUSSION

This paper offers evidence that teacher-training programs that emphasize gender socialization can have a positive short-term impact on knowledge about the difference between gender and sex, and on attitudes toward gender roles and gender identity. The study, however, found no quantitative evidence for a short-term positive impact on self-reported teacher practices. The triangulation of qualitative and quantitative findings indicated that, after eight months, the program equipped intervention teachers with new knowledge about gender, changed their basic attitudes about gender equality issues, and taught them about less traditional views on gender roles. However, the more in-depth qualitative data suggested that teachers still identified with traditional gender norms and beliefs about gender. These findings imply that more reinforcement, longer-term programming, or more community involvement is required to encourage teachers to successfully transfer what they learned to real-life situations in school and non-school settings.

Traditional ideas of gender in the community can be a barrier to changing complex behavior in the short term. The disjunction between the training ideas and deeply embedded community norms was evident throughout the data. Teachers reported that they had difficulty enforcing new ideas about gender norms that did not align with traditional conceptions of gender. The short timeline of the program may also have limited its ability to promote more complex changes in ideas and practices. The majority of the teachers in our sample continue to have difficulty reconciling views consistent with gender equality with traditional ideas of gender in the larger community.

¹⁰ Of the 916 teachers at baseline, 687 were male and 229 were female.

The lack of evidence for positive effects on self-reported teacher practices is consistent with research suggesting that teacher practices are difficult to change in the short term (Bonder 1992; Mukhopadhyay and Wong 2007; Stromquist 2007; Sullivan 2013; World Bank 2008). However, the qualitative component showed that intervention teachers appeared to be changing some simple classroom practices, such as seating arrangements. These positive changes may have resulted from concrete examples in the training sessions and training manual that explain how to operationalize the training concepts.

Although previous research indicates that sending reminder SMS messages can encourage people in developing countries to increase their financial savings (Karan, McConnell, Mullainathan, and Zinman 2010), we did not find evidence that the complementary text messages sent to reinforce the information communicated during the teacher training had a positive impact. Uganda has been using SMS successfully to improve communication between education stakeholders by communicating exam results from the National Examination Board (Ndiwalana 2011). It is possible, however, that messages about gender equality are too complex to communicate via SMS messages. It seems important to reconsider the content of these messages, how they are delivered, and any limitations on teachers' ability to access them.

The findings have several implications for policy and practice. Perhaps most importantly, they speak to the importance of community involvement by parents, politicians, and other community leaders in creating a more enabling environment in which new ideas can be welcomed, accepted, and translated into practice. In addition, the prevalence of qualitative data on the challenge of reconciling traditional norms enforced in the home with new ideas presented in school suggests that the more direct involvement of community and school governance bodies could enhance local buy-in for the content of the training. Finally, the training could benefit from providing teachers with regular coaching, on-site monitoring visits, and/or one-on-one reflection sessions. The literature on the effectiveness of teacher training demonstrates that one-time in-service trainings at a central location tend to be less effective than longer-term teacher-training strategies (Conn 2014; McEwan 2001; Showers and Joyce 1996). Such long-term strategies may be even more of an imperative in programs that aim to change social norms, such as the Gender Socialization in Schools pilot program.

We need to remain cautious in interpreting these findings because of several limitations. First, the program was evaluated during its first year of implementation. New programs may experience unexpected challenges, or they may not be

implemented as intended, both of which could weaken the effects of the program in the first years. Second, we were only able to estimate the short-term effects of the program (after eight months). Third, limited resources precluded our collecting data from teachers by visiting all the treatment schools. We mitigated this limitation by administering the survey the morning the teachers arrived for the training. This strategy reduced data-collection costs but prevented gathering data in exactly the same way across the three study groups (the research team did attempt to mimic the same conditions and procedures across groups). Fourth, the limited resources and short timeline limited our focus to teachers, who were the program's direct beneficiaries. Finally, quantitative data on the teachers were limited to self-reported surveys. Interviews and other more comprehensive data-collection methods were exclusively qualitative. To overcome the limitations of self-reported data—which may suffer from courtesy and social desirability bias—we used vignettes, minimized the use of leading questions, and included various types of questions. We also piloted the instrument multiple times and revised some items after considering comprehensive feedback from local experts.

Future research should focus on experimental designs to determine the impact of gender-socialization programs at the student level and the longer-term effects on teacher practices. We argue that the impact of these programs may be different in conflict-affected settings because psychosocial development and social cohesion may play a smaller role in other low- and middle-income countries. It therefore will also be important to compare the effectiveness of these programs between conflict-affected countries and other low- and middle-income countries. Such comparisons can provide lessons about the links between conflict, gender equality, and social cohesion. In addition, it will be crucial to improve the measurement of teacher practices related to gender and of peacebuilding indicators so we can more fully examine the relation between teachers' activities in the classroom and peacebuilding.

REFERENCES

- Aikman, Sheila, and Elaine Unterhalter. 2005. *Beyond Access: Transforming Policy and Practice for Gender Equality in Education*. Oxford, UK: Oxfam Publishing.
- Anderson, Davie, and Vigids Broch-Due. 1999. *The Poor Are Not Us: Poverty and Pastoralism in Eastern Africa*. Oxford, UK, and Athens, OH: James Curry and Ohio University Press.

- Annan, Jeannie, Christopher Blattman, Dyan Mazurana, and Kristopher Carlson. 2011. "Civil War, Reintegration, and Gender in Northern Uganda." *Journal of Conflict Resolution* 55 (6): 877-908.
- Ashmore, Richard, Frances Del Boca, and Scott Bilder. 1995. "Construction and Validation of the Gender Attitude Inventory: A Structured Inventory to Assess Multiple Dimensions of Gender Attitudes." *Sex Roles* 32 (11-12): 753-85.
- Baber, Kristine, and Corinna Tucker. 2006. "The Social Roles Questionnaire: A New Approach to Measuring Attitudes toward Gender." *Sex Roles* 54 (7-8): 459-67.
- Baranyi, Stephen, and Kristiana Powell. 2005. *Fragile States, Gender Equality and Aid Effectiveness: A Review of Donor Perspectives*. Ottawa: The North-South Institute.
- Barker, Gary, and Jennifer Schulte. 2010. *Engaging Men as Allies in Women's Economic Empowerment: Strategies and Recommendations for CARE Country Offices*. Washington, DC: CARE Norway by International Center for Research on Women.
- Barr, Abigail, Frederick Mugisha, Pieter Serneels, and Andrew Zeitlin. 2012. *Information and Collective Action in Community-Based Monitoring of Schools: Field and Lab Experimental Evidence from Uganda*. Unpublished draft, Dartmouth College, Hanover, NH.
- Blattman, Christopher, and Jeannie Annan. 2010. "The Consequences of Child Soldiering." *Review of Economics and Statistics* 92 (4): 882-98.
- Bonder, Gloria. 1992. "Altering Sexual Stereotypes through Teacher Training." In *Women and Education in Latin America: Knowledge, Power, and Change*, edited by Nelly P. Stromquist. Boulder, CO: Lynne Rienner, 222-49.
- Burde, Dana, Ozen Guven, Jo Kelcey, Heddy Lahmann, and Khaled Al-Abbadi. 2015. "What Works to Promote Children's Educational Access, Quality of Learning, and Wellbeing in Crisis-Affected Contexts." Education Rigorous Literature Review, Department for International Development. London: Department for International Development.

- Burde, Dana, and Leigh Linden. 2013. "Bringing Education to Afghan Girls: A Randomized Controlled Trial of Village-Based Schools." *American Economic Journal: Applied Economics* 5 (3): 27-40.
- Chinen, Marjorie, Andrea Coombes, Thomas De Hoop, and Mohammed Elmeski. 2016. *Final Report for the Impact Evaluation of the Transformative Potential of Positive Gender Socialization in Education for Peacebuilding in Uganda*. Washington, DC: American Institutes for Research. <http://www.ungei.org/Evaluation-of-the-Transformative-Potential-of-Positive-Gender-Socialization.pdf>.
- Conn, Katherine. 2014. *Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Rigorous Impact Evaluations*. New York: Columbia University Press.
- De Hoop, Thomas, Luuk van Kempen, Rik Linssen, and Anouka van Eerdewijk. 2014. "Women's Autonomy and Subjective Well-Being: How Gender Norms Shape the Impact of Self-Help Groups in Odisha, India." *Feminist Economics* 20 (3): 103-35.
- Development Research and Training. 2015. *Gender, Conflict, and Peacebuilding Teacher Handbook*. Kampala: UNICEF Uganda.
- Duncan, Greg J. 2004. *The Contributions of Hard Skills and Socio-Emotional Behavior to School Readiness*. Unpublished doctoral dissertation, Northwestern University, Chicago.
- Dworkin, Shari L., Megan S. Dunbar, Suneeta Krishnan, Abigail M. Hatcher, and Sharif Sawires. 2011. "Uncovering Tensions and Capitalizing on Synergies in HIV/AIDS and Antiviolence Programs." *American Journal of Public Health* 101 (6): 995-1003.
- Glaser, Barney, and Anslem Strauss. 2009. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Alpine.
- Glick, Peter, and Susan T. Fiske. 1997. "Hostile and Benevolent Sexism Measuring Ambivalent Sexist Attitudes toward Women." *Psychology of Women Quarterly* 21 (1): 119-35.

- Instituto da Defesa Nacional. 2013. *Gender Violence in Armed Conflicts* (Report no. 11). Lisbon, Portugal: Instituto da Defesa Nacional.
- Karlan, Dean, Margaret McConnell, Sendhil Mullainathan, and Jonathan Zinman. 2010. *Getting on the Top of Mind: How Reminders Increase Savings* (NBER working papers 16205). Cambridge, MA: National Bureau of Economic Research.
- Kazianga, Harounan, Dan Levy, Leigh Linden, and Matt Sloan. 2013. "The Effects of 'Girl-Friendly' Schools: Evidence from the BRIGHT School Construction Program in Burkina Faso." *American Economic Journal: Applied Economics* 5 (3): 41-62.
- Knutzen, Anna, and Alan Smith. 2012. *Uganda Conflict Analysis* (Internal document). New York: UNICEF.
- McEwan, Patrick J. 2014. "Improving Learning in Primary Schools of Developing Countries: A Meta-Analysis of Randomized Experiments." *Review of Educational Research* 85 (3): 353-94.
- McKenzie, David. 2012. "Beyond Baseline and Follow-Up: The Case for More T in Experiments." *Journal of Development Economics* 99 (2): 210-21.
- Ministry of Gender, Labour, and Social Development, and UNICEF. 2015. "Situation Analysis of Children in Uganda" (Report). Kampala: UNICEF Uganda.
- Mkutu, Kennedy. 2008. *Guns and Governance in the Rift Valley: Pastoralist Conflict and Small Arms, African Issues*. Bloomington: Indiana University Press.
- Mukhopadhyay, Maitrayee, and Franz Wong. 2007. *Revisiting Gender Training: The Making and Remaking of Gender Knowledge. A Global Sourcebook*. Oxford, UK: Oxfam Publishing.
- Ndiwalana, Ali. 2011. *Uganda Case Study: Transformation-Ready. The Strategic Application of Information and Communication Technologies in Africa* (Final education sector study). London: ICT Development Associates Ltd.

- Opinia, Sylvia, and Friederike Bubenzer. 2011. *Gender Justice and Reconciliation in Northern Uganda* (Policy brief no. 4.) Cape Town: Institute for Justice and Reconciliation.
- Pham, Phuong, and Patrick Vinck. 2010. *Transitioning to Peace: A Population-Based Survey on Attitudes about Social Reconstruction and Justice in Northern Uganda*. Berkeley, University of California, Human Rights Center.
- Pham, Phuong N., Patrick Vinck, and Niamh Gibbons. 2015. *Population-Based Survey on Peace and Education: Uganda*. New York: UNICEF, Learning for Peace, Harvard Humanitarian Initiative.
- Pulerwitz, Julie, and Gary Barker. 2008. "Measuring Attitudes toward Gender Norms among Young Men in Brazil: Development and Psychometric Evaluation of the GEM Scale." *Men and Masculinities* 10 (3): 322-38.
- Puri, Jyotsna, Anastasia Aladysheva, Vegard Iversen, Yoshodhan Ghorpade, and Tilman Brück. 2015. *What Methods May Be Used in Impact Evaluations of Humanitarian Assistance?* (Discussion paper no. 8755). Bonn: Institute of Labor Economics.
- Saminsky, Alina. 2010. "Burying the Pen: Tackling Resistance to Formal Education by the Karamojong in Uganda." *Inquiries Journal/Student Pulse* 2 (2). <http://www.inquiriesjournal.com/a?id=154>.
- Showers, Beverly, and Bruce Joyce. 1996. "The Evolution of Peer Coaching." *Educational Leadership* 53: 12-16.
- Specht, Irma. 2013. *Gender, Disarmament, Demobilization and Reintegration and Violent Masculinities* (Report no. 11). Lisboa: IDN Cadernos.
- Strickland, Richard, and Nata Duvvury. 2003. *Gender Equity and Peacebuilding: From Rhetoric to Reality. Finding the Way* (Discussion paper). Washington, DC: International Center for Research on Women.
- Stromquist, Nelly P. 2007. *The Gender Socialization Process in Schools: A Cross-National Comparison* (Background paper prepared for the Education for All Global Monitoring Report 2007). New York: UNESCO.

- Sullivan, Oriel. 2013. "What Do We Learn about Gender by Analyzing Housework Separately from Child Care? Some Considerations from Time-Use Evidence." *Journal of Family Theory & Review* 5 (2): 7-84.
- Sustainable Development Goals. 2015. "Goal 5: Achieve Gender Equality and Empower all Women and Girls." http://www.un.org/sustainabledevelopment/wp-content/uploads/2016/08/5_Why-it-Matters_GenderEquality_2p.pdf.
- Vaughan, Jenny, and Tim Stewart. 2011. *Cattle Raiding in Karamoja: A Conflict and Market Assessment*. Portland, OR: Mercy Corps.
- White, Howard. 2009. "Theory-Based Impact Evaluation: Principles and Practice." *Journal of Development Effectiveness* 1 (3): 271-84.
- White, Howard, and Daniel Phillips. 2012. *Addressing Attribution of Cause and Effect in Small n Impact Evaluations: Towards an Integrated Framework* (Working paper 15). New Delhi: International Initiative for Impact Evaluation.
- Winthrop, Rebecca, and Jackie Kirk. 2008. "Learning for a Bright Future: Schooling, Armed Conflict, and Children's Well-Being." *Comparative Education Review* 52 (4): 639-61.
- World Bank. 2008. *Gender in Agriculture Sourcebook*. Washington, DC: World Bank.