2014

Evaluation of health and education impacts of a girls' safe spaces program in Addis Ababa, Ethiopia

Annabel Erulkar
Population Council

Girmay Medhin

report

EVALUATION OF HEALTH AND EDUCATION IMPACTS OF A GIRLS' SAFE SPACES PROGRAM IN ADDIS ABABA, ETHIOPIA

Annabel Erulkar PhD
Girmay Medhin PhD

POPULATION COUNCIL
Ideas. Evidence. Impact.

POPULATION COUNCIL
Ideas. Evidence. Impact.

The Population Council confronts critical health and development issues—from stopping the spread of HIV to improving reproductive health and ensuring that young people lead full and productive lives. Through biomedical, social science, and public health research in 50 countries, we work with our partners to deliver solutions that lead to more effective policies, programs, and technologies that improve lives around the world. Established in 1952 and headquartered in New York, the Council is a nongovernmental, nonprofit organization governed by an international board of trustees.

Population Council
Heritage Plaza, 4th Floor
Bole Medhalialem Road
P.O. Box 25562, code 1000
Addis Ababa, ETHIOPIA
Tel: +251-116-631-712/4/6
email: info.ethiopia@popcouncil.org

popcouncil.org

Annabel Erulkar PhD is Country Director in the Population Council's office in Ethiopia and Senior Associate in the Poverty, Gender and Youth Program of the Population Council.

Girmay Medhin PhD is Associate Professor at Aklilu Lemma Institute of Pathobiology at Addis Ababa University.

Suggested citation: Erulkar A, Medhin G. 2014. "Evaluation of Health and Education Impacts of a Girls' Safe Spaces Program in Addis Ababa, Ethiopia,"
Addis Ababa: Population Council.

Photo credits: Dhaniel Negatu

Table of Contents

Acknowledgments

The project was made possible by funding from the Italian government through the CHYAO Trust Fund administered by the World Bank. The contents are the responsibility of the authors and do not necessarily reflect the views of the Italian government or the World Bank. We are grateful to our collaborators at the Ethiopian Ministry of Women, Children & Youth Affairs, the Addis Ababa Bureau of Women, Children & Youth Affairs and Kebele officials. Judith Bruce and Sarah Engebretsen made very helpful comments on earlier versions of this paper, as did multiple colleagues at the World Bank. Ruth Naylor provided very valuable input on measuring educational outcomes and analysis. We are grateful to Population Council staff who helped coordinate the research, including Tigest Tamrat, Lemi Negeri, Hanan Nourhussein and Habtamu Demele.

Introduction

Approximately one in eight of the world's population is a girl or young woman aged 10 to 24 and attention is increasingly focusing on to the central role of adolescent girls' in achieving global health and development goals (Levine, Lloyd, Greene and Grown, 2008; Temin and Levine, 2009, Plan International, 2007). Areas of focus by the development field include girls' education, their health, child marriage, and the experience of violence.

The focus on girls as key actors in development has resulted in an increasing number of initiatives developed to support girls in developing countries. Many of the programs for developing country girls employ a 'safe spaces' approach. A 'safe spaces' approach has been broadly described as "*a place where anyone can relax and be able to fully express, without fear of being made to feel uncomfortable, unwelcome, or unsafe on account of biological sex, race/ethnicity, sexual orientation, gender identity or expression, cultural background, religious affiliation, age, or physical or mental ability… A place where the rules guard each person's self-respect and dignity and strongly encourage everyone to respect others*" (Safespacenetwork.com). As female-only gatherings, girls' safe spaces programs are characterized as maintaining a private and confidential atmosphere, free from physical or psychological threats; conveniently located and safe to access, and not subject to intrusion from those in authority or those who would otherwise disrupt the environment of self-expression and exchange (Brady, 2005). Safe spaces are not necessarily dedicated, physical spaces, but can be created by applying these principles to meetings in a variety of locally available settings, including gatherings under a tree or in community halls. Safe spaces programs in both developed and developing countries are frequently targeted to subgroups of girls who are relatively more marginalized or isolated.

One of the earliest safe spaces programs for developing country girls was the Ishraq (Sunrise) program in rural Upper Egypt. Launched in 2001, the program mobilized out-of-school girls aged 12 to 15 into safe spaces groups where they received basic literacy, life skills, and financial education. In addition, Ishraq participants played sports, an uncommon opportunity for adolescent girls in this very traditional, conservative setting. The program ultimately reached over 3,000 girls. While adolescent girls in traditional settings often face increased restrictions on their movements and confinement to the domestic sphere, Ishraq demonstrated that safe spaces programs can overcome those restrictions and provide girls with opportunities to expand social networks, gain skills, and pursue personal development (Selim, et. al. 2013).

In Uganda, Empowerment and Livelihoods for Adolescents (ELA) was a two-year program for girls aged 14 to 20, which provided life skills, financial literacy and vocational training through girls' groups led by female mentors. Participating girls were given the choice of vocational training, including hairdressing, tailoring, computer skills, and poultry rearing, among others. Life skills topics included sexual and reproductive health, pregnancy, sexually transmitted infections, HIV/AIDS, family planning and rape. ELA included a randomized control trial evaluation design, where the intervention was randomized to 100 treatment and 50 control communities. The evaluation demonstrated positive impacts on girls residing in the ELA communities including increases in employment, increases in condom use, decreased fertility and reduction in reports of nonconsensual sex (Bandiera O, Buehren N, Burgess R, et. al, 2012).

In rural Ethiopia, Towards Economic and Sexual Reproductive Health Outcomes for Adolescent Girls (TESFA) focused on providing support to girls who were married as children, below the age of 18. TESFA was implemented in two rural districts over a three year period. Married girls aged 10 to 19 were recruited into groups and provided with either: 1) economic empowerment information and skills, 2) sexual and reproductive health (SRH) information, or 3) both economic empowerment and SRH information. At

endline, while the proportion of girls with savings increased in all groups, the increase was dramatic and significant in the economic empowerment group and the combined group. In the SRH group, use of family planning increased by 27 percentage points from baseline to endline, a statistically significant difference from the other arms of the study (Edmeades J, Hayes R, Gaynair G, 2014).

Another program for married girls in Ethiopia was Meserete Hiwott (Base for Life). Married girls aged 10 to 24 were formed into safe spaces groups and provided with HIV and SRH information and financial literacy, as well as skills related to communication, self-esteem, equitable gender roles and non-violence. In parallel, a program for husbands promoted men's support of women's reproductive health, family health, and non-violent gender equitable relationships. Despite early skepticism that married girls would be allowed to join groups, over 230,000 married girls and 130,000 husbands participated in the program over five years. The post-test only evaluation suggested improvements in men's assistance with housework, accompaniment of wives to the clinic and couples undergoing voluntary counselling and testing (VCT) (Erulkar and Tamrat, 2014).

In urban Burkina Faso, Filles Eveilles (Girls Awakened) is an 8-month intervention for migrant girls in domestic service. Domestic workers aged 11 to 18 meet in safe spaces groups with adult female mentors. Meetings are held once a week, in locally available spaces such as community halls or schools. The interactive group meetings include life skills, health and hygiene, SRH and financial education. The pre- and post-test evaluation showed improvements in health seeking behavior, SRH knowledge, and savings behavior (Engebretsen, 2013).

Safe spaces approaches have been used for very marginalized and isolated groups such as married adolescents, domestic workers and out-of-school girls in poor rural areas. The aim of this study is to analyze the educational and health impacts of a program for slum-dwelling girls, including child domestic workers and rural-urban migrants.

'BIRUH TESFA' (BRIGHT FUTURE), ETHIOPIA

Established in 2006, Biruh Tesfa (Bright Future, in Amharic) aimed to increase social networks and support for the most marginalized girls in the poorest urban areas of Ethiopia. The program was implemented by the Ethiopia Ministry of Women, Children and Youth Affairs and their regional bureaus in Addis Ababa, Amhara and Tigray regions. Through the program, adult female mentors were recruited from low income urban areas and trained on the project curriculum, facilitation skills and recruitment. Once trained, female mentors paid house-to-house visits to mobilize out-of-school girls into safe spaces groups. During these house-to-house visits, mentors were able to negotiate for girls' participation if there was resistance from domestic workers' employers or other guardians. This recruitment strategy resulted in mobilization of the most marginalized groups of girls, including child domestic workers, daily laborers, and rural-urban migrants (Erulkar A, Gebru H, Mekonnen G, 2011, Erulkar, 2014). Once in groups, girls received basic literacy training and life skills, as well as wellness check-ups. In an earlier evaluation of the program which used a quasi-experimental research design, the Biruh Tesfa program demonstrated positive changes among girls in the project areas compared to those in the control areas in terms of social support, HIV knowledge, and knowledge and demand for VCT (Erulkar A, Ferede A, Girma W, Ambelu W, 2012).

In 2013, Powering Up 'Biruh Tesfa' was the name given to a new initiative to expand the Biruh Tesfa model in Addis Ababa, Ethiopia's capital, and focusing attention on measuring learning and health outcomes. The project was expanded to five subcities[1] and 17 woredas/districts of Addis Ababa, especially to low-income areas on the outskirts of the city which were emerging because of rapid urbanization. Eligibility for the program is being an out-of-school girl aged 7 to 18, residing in the project area. In all, 20 new sites for

[1] Subcities in Addis Ababa were Gulele, Arada, Yeka, Chirkos, and Lideta.

girls' safe spaces groups were established, managed by 80 female mentors. The program uses existing sites within the communities, mainly primary schools (used after school hours), youth centers, and community halls run by the local administration (kebele). Modest renovations were undertaken in these facilities including partitioning rooms, painting, and constructing or repairing roofs.

In each site, two or three separate rooms are used for the girls' groups. Groups are segmented by age, with separate groups for those aged 7 to 11 and 12 to 18. Biruh Tesfa safe spaces groups meet in the late afternoon, and last for two hours, meeting five days per week. The timing of group meetings was selected to coincide with the availability of school classrooms as well as the most likely time when girls in domestic service are free to attend the classes. Two mentors facilitate each class. Once in the groups, girls receive non-formal education during four out of five meetings in a week, using the Ethiopia Ministry of Education non-formal education curriculum. The fifth day's meeting is devoted to life skills that cover topics such as self-esteem and communication, financial literacy, hygiene and menstruation, disabilities, and violence. Each participating girl receives school materials including exercise books, pens, pencils, text books and a book bag. They also receive some basic clothing, including underwear, and sanitary napkins. In all, participants receive materials worth about 800 Ethiopian Birr, or about USD 42.

Girls in need of health services are offered a voucher to subsidize the cost of basic health services. To set up the system, clinics located in the catchment areas of the safe spaces sites were inventoried and information collected on services offered, cost of services and hours of operation. Selected clinics from both the private and public sectors were invited to join the network of providers to support marginalized girls. Ultimately, 12 partner clinics (11 private and 1 governmental) enrolled in the voucher scheme and received orientation. Girls who require basic medical services can request a voucher from her mentor. Mentors also offer to accompany girls to the service, if they are uneasy to go alone, do not know where to go or have misgivings about seeing a medical professional, for many, their first time. The voucher provides girls with free medical consultations, services and necessary medications. Vouchers are used for a range of basic medical issues including gastro-intestinal problems, respiratory problems, skin infections, injuries, reproductive tract infections and services related to HIV. Providers submit vouchers at month's end to the project, for reimbursement of services and medications provided.

The Powering Up Biruh Tesfa expansion was implemented for six months from September 2013 to February 2014, after which the program continued through support of another donor. In all, 3,159 girls took part in the program expansion, across 17 woredas/districts. The average age of participating girls was 12.8 years. Forty-one percent of participants were child domestic workers. During the six-month project period, 487 vouchers for medical service were issued to 320 participants. Seventy percent of the voucher users had never visited a health facility before receiving the voucher, suggesting a high proportion of girls were introduced to the formal health system through the scheme.

Methods

DATA COLLECTION

This is a longitudinal study of girls residing in the expansion sites of Biruh Tesfa and in comparable areas where Biruh Tesfa was not implemented. The baseline survey was conducted prior to the establishment of the girls' safe spaces groups; the endline survey took place among the same respondents, six months after establishment of the groups. The study was conducted in 27 woredas/districts of the five sub-cities of Addis Ababa; 17 were experimental areas and 10 served as waitlisted controls. The intervention districts were purposefully selected in collaboration with the Addis Ababa Bureau of Women, Children and Youth Affairs. Partners selected low-income areas of Addis Ababa, especially those emerging on the outskirts of the city, for establishment of Biruh Tesfa safe spaces sites. Comparable control districts were selected based on available population and school attendance data, as well as the researchers' observation of socio-economic similarities with the experimental areas.

Girls aged 12 to 18, who were out-of-school during the 2012-13 academic year, and interested and allowed to participate in a girls' club were considered eligible for the survey in both the experimental and control sites. Study participants were identified through house-to-house visits made by program mentors and data collectors. In the intervention area, mentors went house-to-house with local guides to identify eligible girls and introduce the Biruh Tesfa program to girls, household heads, and other gatekeepers. The mentors were provided a script for introducing the program to ensure uniformity in recruitment. Girls who wanted to join the program and who were given permission to do so were included in the sampling frame for the survey.

In the control areas, interviewers were coupled with local guides to make similar house-to-house visits. Local guides were selected to have similar characteristics as the mentors. During these visits, interviewers undertook similar recruitment procedures and were similarly scripted, as in the intervention areas, introducing the program to both eligible girls and gatekeepers. However, instead of being invited into the program, girls were asked if they would be interested and allowed to join a girls' program, if it were to be expanded to their areas. Gatekeepers were asked if they would allow eligible girls in their household to join, should an expansion occur. This measure was undertaken in an attempt to control for selectivity, or the bias that results when those joining programs have different characteristics than those not joining. After introducing the program and obtaining the approval and intention to enroll the girl into the program, the interviewer recorded the household information on the listing form to enter the girl into the sampling frame. Using SPSS statistical software, 2,000 girls (1,000, each, from intervention and control areas) were randomly selected from the sampling frame of out-of-school girls 12 to 18 who were interested in joining girls' groups,.

The survey instrument was a structured questionnaire covering a range of issues including demographic characteristics, educational experience, migration, work experience, time-use, social networks, reproductive health, sexual history and the experience of violence. In addition, respondents were administered reading and math tests in order to measure literacy and numeracy levels. The same study instrument was used at both rounds of survey. However, at endline, an additional section measuring exposure to the intervention was added to the end of the questionnaire. 'Biruh Tesfa' targets out-of-school girls aged 7 to 18 years. However, the survey was restricted to those aged 12 to 18, for ethical reasons. In addition, younger adolescents aged 12 to 14 received an abbreviated questionnaire, omitting sensitive questions on sexual activity and violence.

The study is longitudinal. In order to control for loss to follow-up, interviewers requested a phone number through which respondents could be contacted for re-interview at endline. At endline, respondents were also offered a phone card with the value of 50 Ethiopian Birr (ETB; approximately USD 3) to encourage continued contact and for re-interview at endline. Informed consent was obtained from all survey participants and parental/guardian consent was provided in cases when the respondent was under age 18. The study protocol was approved by the local ethical review board at the Ethiopia Ministry of Science and Technology and the sponsoring institution's review board.

MEASURES

The expansion phase, 'Powering Up Biruh Tesfa' aimed to increase literacy and numeracy capabilities among out-of-school girls living in the poorest and most marginalized urban areas. It also aimed to improve girls' health, including utilization of health services.

Literacy and numeracy acquisition was measured using reading and math cards based on those developed for the UWEZO[2] program in East Africa, measuring progressive skill levels in reading and math. The UWEZO tests were adapted for use in Ethiopia. For example, the reading cards measured identification of Amharic letters, followed by reading of single words and sentences. Finally, the test measures reading for comprehension. Math cards follow a similar progression of numeracy acquisition, including counting dots, identifying numbers, identifying the larger of two numbers, adding, subtracting and, finally, word problems. Two tests were adapted at baseline and assigned at random to respondents. At endline, two different tests were adapted and similarly assigned at random.

Literacy was measured through four items while math was measured through nine items. Each item was coded as to whether or not the respondent correctly answered it. An overall score was constructed to reflect acquisition of literacy in the aggregate (ranging from 0 to 4) and well as overall numeracy (ranging from 0 to 9). A combined literacy-numeracy score was constructed, with scores ranging from 0 to 1. Because the number of numeracy items outnumbered literacy items, literacy items were weighted by a factor of two (scores ranging from 0 to 8), added to the numeracy score, and divided by 17, the total amount of items in the test.

Health seeking behavior was another primary outcome. Respondents were read a list of places and asked if they had frequented the facility in the last six months. Two of the locations were private clinics and government hospitals or health centers. If the respondent answered that she had been to either a private clinic or government hospital / health center, she was marked as having used a health service.

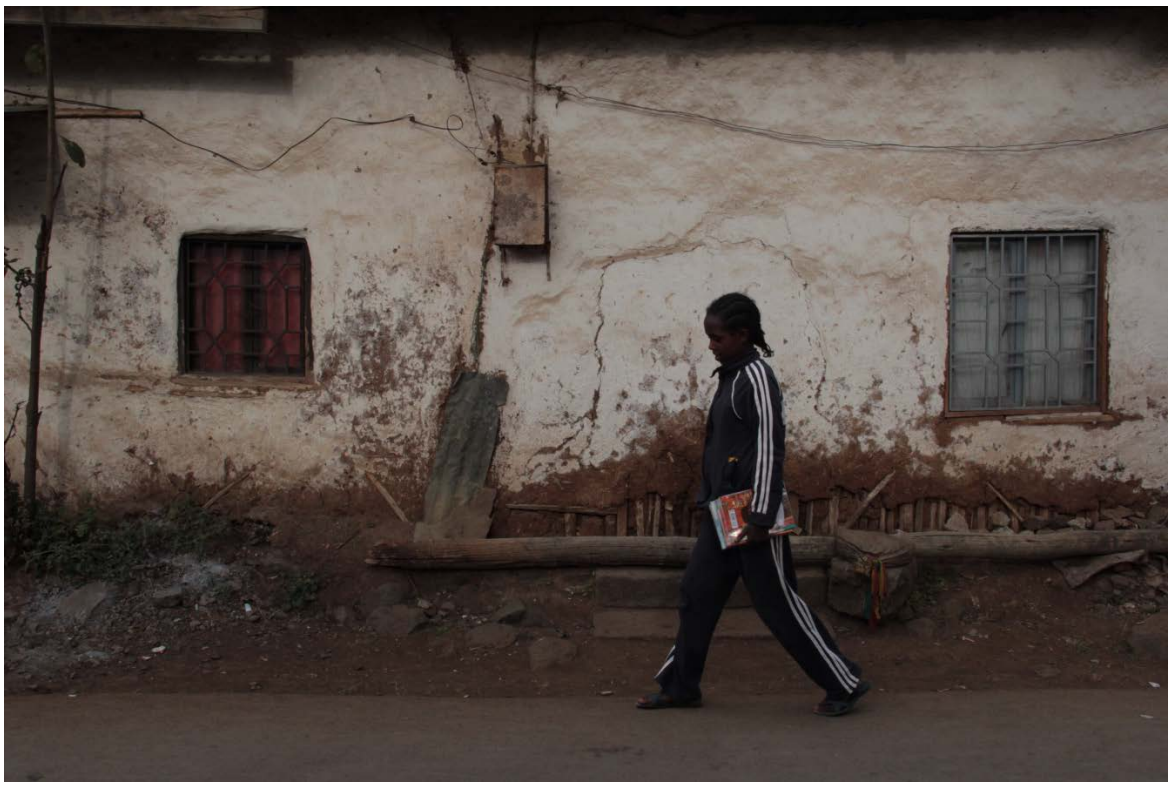
ANALYSIS

The data was entered using EpiData and analysis was undertaken in STATA and SPSS. We examined response rates during both rounds of survey. Because of loss to follow-up in the second round of survey, we stratified baseline respondents by completion status (completed versus lost to follow-up) and examined background characteristics of each group. Educational outcomes were examined, item by item and in aggregate score, for experimental and control groups at both baseline and endline. There was significant school enrollment between baseline and endline surveys. Therefore, we created an interaction variable between treatment arm and having a history of attending formal school, in order to examine whether girls who had never been to school and girls who had ever been to school were differentially impacted in the experimental and control sites. Generalized linear models were used to compare mean aggregate literacy and numeracy scores between the two study arms at endline, after taking account of an individual's

[2] *www.uwezo.net*

baseline results and background characteristics including age, migration history and relationship to household head. After obtaining a significant interaction between the history of formal schooling and treatment arms, we conducted this analysis on the sub-group of girls who had never been to formal schooling by the time of endline survey. Coefficients are reported for the effect of being in the experimental group, after testing the other factors, both individually and in a fully adjusted model.

For the dichotomous outcome, health service utilization in the last six months, we used McNemar's chi-squared test to examine the differences between baseline and endline, followed by logistic regression, controlling for baseline estimates of the outcome and background variables. Age was included in the models as a continuous variable. Migration status was coded as a dichotomous variable reflecting whether the respondent had migrated to Addis Ababa from another place or was native to Addis Ababa. Respondents' relationship to the household head was coded as being 1) a daughter or granddaughter to the household head, 2) another relative (other than daughter or granddaughter), or 3) a non-relative or employee of the household head.



Results

SAMPLE CHARACTERISTICS

At baseline, 1,876 girls were interviewed out of the 2,000 sampled, amounting to a 94 percent response rate (Table 1). At baseline, samples from experimental and control sites differed with respect to age, schooling, and their relationship to the household head. Respondents from the control group were significantly more likely than those in the experimental group to have been to school and to have attained higher levels of schooling. For example, 25 percent of girls in the experimental group had achieved at least five years of schooling, while 35 percent of girls in the control group had achieved at least five years. Likewise, respondents in the control group were significantly more likely to be employees or nonrelatives of the household head (50 percent), compared to respondents in the experimental area (39 percent).

The majority of respondents were migrants to Addis Ababa (92 percent). Few girls reported media exposure through radio, the most accessible form of media. Sixty-four percent of the experimental group and 59 percent of the control group reported never listening to radio, which partly reflects the level of isolation and marginalization of this group of girls.

TABLE 1: Baseline sample characteristic, by treatment group, and endline characteristics by completion status

	Experimental (n=913)	Control (n=963)	p value	Completed at endline (n=1,276)	Lost to follow up (n=600)	p value
Age group			<0.001			<0.05
12 to 14	39.2	31.3		31.3	36.9	
15 to 18	60.8	68.7		68.7	63.1	
Ever been to school (yes)	67.0	78.7	<0.001	72.7	73.2	NS
Highest year of education			<0.001			NS
None	35.7	25.1		30.0	30.4	
1 to 4 year	39.1	39.7		40.8	38.7	
5 or more years	25.2	35.2		29.2	30.9	
Relationship to household head			<0.001			<0.001
Daughter or grand-daughter	12.3	8.7		14.5	1.8	
Other relative	48.7	41.5		49.1	36.2	
Employee or nonrelative	39.0	49.8		36.4	62.0	
Migrant to the area (yes)	91.2	92.1	NS	88.5	98.3	<0.001
Radio listenership			NS			NS
Never listens	63.5	59.2		61.5	61.2	
Listens daily or periodically	36.5	40.8		38.5	38.8	
Treatment group						NS
Experimental	-	-		47.2	49.4	
Control	-	-		52.8	50.6	

*Differences between groups significant at *p<0.05 **p<0.01 ***p<0.001*

At endline, the study achieved a 68 percent response rate, or 32 percent loss to follow-up. There was no difference in levels of loss-to-follow-up between treatment and control groups. The reasons for loss-to-follow-up were migrating out of Addis Ababa (59 percent), moving to another unknown location (36 percent), refusal (3 percent), migrating to another country (2 percent) and deceased (< 1 percent). Schooling participation and attainment was not associated with loss to follow-up. However, respondents who were an employee or nonrelative of the household head and those who were migrants, were

significantly more likely to be lost to follow-up compared to girls who were related to the household head or native to the area.

EXPOSURE TO BIRUH TESFA

Among girls in the experimental areas, 47 percent ultimately reported that they participated in Biruh Tesfa. Girls who ultimately participated in the groups were significantly more likely to be in the younger age group and less likely to be domestic workers. Whether or not the girls was a migrant was not associated with participation. Three girls in the control area reported participating, reflecting minimal contamination in the control area. When girls in the experimental area were asked the reason for not participating, the most common reasons were that they did not know enough about the project (34 percent), they had no time to attend the meetings (32 percent) and employer or family disapproval of the project (14 percent) (*analysis not shown*).

On average, participating girls attended 39 sessions. Comparing migrants and non-migrants, as well as groups based on relationship to the household head (daughter/granddaughter, other relative, nonrelative/employee), there was no statistically significant difference in the number of meetings attended between groups. However, younger adolescents aged 12 to 14 attended an average of 43 meetings compared to older adolescents (aged 15 to 18) who attended an average of 34 meetings, a difference that was statistically significant. There were high levels of exposure to all the topics covered in the curriculum, in particular hygiene (90 percent), numeracy (82 percent) and Amharic literacy (82 percent). Among the topics in the curriculum, family planning was mentioned the least as having been covered. This may be due to mentors' discomfort with the topic and resulting tendency to avoid it, or to beneficiaries' reluctance to report having received the information. In addition, it appeared that younger girls were slightly less likely to be exposed to the topic than older girls; 50 percent of girls 12 to 14 reported having received the topic compared to 58 percent of girls 15 to 18 ($p<0.10$). Nearly one third of participating girls (31 percent) used the medical voucher for services *(analysis not shown)*.

EDUCATIONAL OUTCOMES

Between baseline and endline, participation in schooling increased dramatically (Table 2). Eligibility for inclusion in the study was being out of formal school at baseline. Between baseline and endline, participation in formal schooling among formerly out-of-school girls increased from 0 to 38 percent. The increase was similar in the experimental and control groups which has implications for inference. For example, between rounds of survey, the percent of girls who had never attended school declined in the experimental area by 12 percentage points and in the control area by 10 percentage points. The enrollment is likely attributable to the Ministry of Education campaign which was launched in September, 2013 to get children back to school. Under the campaign, parents, teachers, communities and community leaders are given the responsibility to target out-of-school children and support their re-entry into school (UNICEF, 2013).

At the same time, attendance of non-formal schooling also increased dramatically in the experimental site. Ever attendance in non-formal schooling increased in the experimental site from 6 percent to 49 percent and from 5 percent to 12 percent in the control arm. The dramatic increase in non-formal education in the experimental site seems attributable to Biruh Tesfa. Among those reporting attendance of non-formal education at endline, 86 percent were participants in Biruh Tesfa.

TABLE 2: Educational participation in formal and nonformal school by treatment group and time of survey

	Experimental (n=630)		Control (n=646)	
	Baseline	Endline	Baseline	Endline
Formal schooling				
Never attended	32.4	20.3***	21.4	11.5***
Formerly attending	67.6	42.2***	78.6	50.3***
Currently attending	0.0	37.5***	0.0	38.2***
Mean years formal education	2.6	3.2	3.3	3.8***
Nonformal education				
Never attended	94.0	51.4***	95.2	88.5***
Formerly attending	5.6	14.4***	4.3	5.6
Currently attending	0.5	34.1***	0.5	5.9***

*Differences between survey rounds significant at *p<0.05 **p<0.01 ***p<0.001*

Between baseline and endline, respondents in both experimental and control areas improved their scores on the literacy and numeracy tests (Table 3). On average, both experimental and control respondents improved their Amharic literacy by nearly half a point, on a scale of 0 to 4; numeracy scores increased by more than half a point, changes that were statistically significant for both groups.

Because of increases in school enrollment across study arms due to a Ministry of Education campaign, we focused attention on learning outcomes among girls who had never been in the formal schooling system at endline. This was an attempt to remove the effect of the Government of Ethiopia's campaign to reintegrate young people back into the formal schooling system. As expected, girls who had never been to school recorded relatively low scores on the literacy and numeracy tests, compared to the overall study population. At baseline, girls who had never been to school scored at average of 0.4 (out of a maximum 4.0) on the reading test and 2.6 (out of a maximum 9.0) on the math test. Between baseline and endline, girls in the experimental area who had never been in formal schooling demonstrated significant increases on the Amharic literacy score, numeracy score and composite score. For example, the mean score on the numeracy test increased from 2.6 to 3.5 between baseline and endline ($p<0.001$). Similar improvements in test scores were not seen among girls who had never been to school in the control group. Scores for these girls in the control group either remained constant or only increased slightly, in the case of the numeracy test.

TABLE 3: Mean scores (and standard deviation) on aggregate Amharic literacy and numeracy measures among all respondents and those who have never been to formal schooling, by treatment group and time of survey

	All respondents				Respondents who never attended formal school			
	Experimental (n=630)		Control (n=646)		Experimental (n=127)		Control (n=74)	
	Baseline	Endline	Baseline	Endline	Baseline	Endline	Baseline	Endline
Amharic literacy (0-4)	1.86 (1.82)	2.26*** (1.77)	2.07 (1.78)	2.45*** (1.66)	0.46 (0.10)	0.66* (0.10)	0.32 (0.85)	0.32 (0.72)
Numeracy (0 – 9)	5.20 (3.15)	5.98*** (2.85)	5.58 (3.00)	6.21*** (2.86)	2.63 (2.48)	3.50*** (2.78)	2.50 (2.25)	2.64 (2.16)
Literacy/numeracy score (0 – 1)	0.52 (0.37)	0.62*** (0.34)	0.57 (0.35)	0.65*** (0.33)	0.21 (0.25)	0.28*** (0.26)	0.19 (0.21)	0.19 (0.18)

*Paired T-test; Differences between means significant at *p<0.05 **p<0.01 ***p<0.001*

Because of the enrollment campaign that increased school enrollment between rounds of survey, we included an interaction term combining treatment arms and having ever attended formal school to understand if girls who had never attended school were impacted differently in the experimental and control sites (Table 4). The main effect of residing in the intervention area was not significantly associated with a greater mean score on the literacy / numeracy test. The effect of never having attended school was associated with an average of 0.52 points lower on the test, compared to those who had attended school. The interaction term was significant. After controlling for other background characteristics, age, migration history and relationship to household head, the intervention positively impacted on those who were had never attended formal schooling and who resided in the intervention site.

TABLE 4: Mean difference in aggregate literacy/numeracy score (range 0 to 1) (and 95% confidence intervals) by treatment group and history of attending formal schooling

Outcome	Mean difference (95% CI)	p-value
Unadjusted effect of being in intervention group	-0.01 (-0.05, 0.03)	0.52
Effect of being in intervention group after adjusting for ever having attended school and interaction term		
Being in the intervention	-0.01 (-0.05, 0.02)	0.55
Never attended school	-0.52 (-0.59, -0.45)	0.00
Interaction never attended school and intervention group	0.10 (0.01, 0.19)	0.03
Fully adjusted effect of being in intervention group[1]		
Being in the intervention	-0.01 (-0.04, 0.03)	0.75
Never attended school	-0.52 (-0.59, -0.45)	0.00
Interaction never attended school and intervention group	0.11 (0.01, 0.20)	0.02

[1] *Adjusting for age, migration history and relationship to household head*

Because, the interaction combining treatment group and schooling history was significant, we conducted sub-group analysis only on the girls who had never attended formal schooling. Generalized linear models were used to estimate the mean difference of never-schooled respondents' literacy/numeracy aggregate score, between treatment groups, after adjusting for one's baseline score, age, migration status and relationship to household head (Table 5). After adjusting for background factors and one's baseline performance on the test, being in the experimental group was associated with a significantly higher endline literacy/math score, compared to the control group (adjusted mean difference 0.08; 95% CI 0.02, 0.14; $p<0.05$). This suggests that residing in the Biruh Tesfa project site was associated with improvements in reading and math scores among girls outside of the formal schooling system.

TABLE 5: Mean difference in aggregate literacy/numeracy score (range 0 to 1) (and 95% confidence intervals) among girls who have never attended formal schooling, by treatment group, baseline score, and background characteristics

Outcome	Mean difference (95% CI)	p-value
Unadjusted effect of being in intervention group	0.09 (0.02, 0.16)	<0.01
Effect of being in intervention group after adjusting for baseline score	0.08 (0.02, 0.13)	<0.01
Effect of being in intervention group after adjusting for baseline score and each of the following factors, individually:		
Age	0.07 (0.02, 0.13)	<0.05
Migrant to the area	0.08 (0.02, 0.13)	<0.01
Relationship to household head	0.08 (0.02, 0.14)	<0.01
Fully adjusted effect of being in intervention group	0.08 (0.02, 0.14)	<0.05

HEALTH SERVICE UTILIZATION

Health service utilization in the last six months increased significantly among respondents in the experimental site between baseline and endline (from 31 to 41 percent; p<0.001). Among respondents in the control site, the increase in health service utilization was modest and not significant (26 to 30 percent) (*analysis not shown*).

At endline, the odds of health service utilization in the last six months was 1.60 times higher in the experimental group than the control group (CI: 1.27, 2.02; p <0.001) (Table 6). Although baseline health service utilization was a significant predictor of the current health service utilization, the observed excess health service utilization by the intervention group was not explained by baseline usage. Similarly, age of study participants, being migrant to the area, relationship of to the household head and having attended formal schooling did not explain the observed excess health service utilization by the intervention group over that of the control group (adjusted OR = 1.55; CI: 1.22, 1.98; p <0.001).

TABLE 6: Odds ratios (and 95% confidence intervals) comparing intervention and control groups' health service utilization in the last six months

Outcome	Odds Ratio (95% CI)	p-value
Unadjusted effect of being in intervention group	1.60 (1.27, 2.02)	<0.001
Effect of being in intervention group after adjusting for baseline health service use	1.55 (1.23, 1.97)	<0.001
Effect of being in intervention group after adjusting for baseline health service use and each of the following factors, individually:		
Age	1.59 (1.25, 2.01)	<0.001
Migrant to the area	1.55 (1.22, 1.96)	<0.001
Relationship to household head	1.52 (1.20, 1.94)	<0.001
Ever attended formal schooling	1.55 (1.22, 1.97)	<0.001
Fully adjusted effect of being in intervention group	1.55 (1.22, 1.98)	<0.001

DISCUSSION

This study has limitations. At baseline, the sample of respondents in the two groups was not comparable in terms of age, schooling and their relationship to the household head. These variables were included in multivariate models in order to adjust for differences in the two groups. At the same time, the differences in the samples may be an indicator for other differences in the groups that are unobserved or unmeasured, which may have influenced the results. In addition, nearly a third of respondents were lost to follow-up at endline. Respondents who were lost to follow-up were more likely to be employees or nonrelatives of the household head and also more likely to be migrants, compared to those who were successfully interviewed. While attempts were made to minimize this loss, the results suggest that longitudinal designs may be less appropriate for highly mobile urban populations, such as domestic workers and other girls living away from parents. Indeed, our level of loss-to-follow up is consistent with other longitudinal studies of adolescents in low income urban areas of Africa (Marston, Beguy, Kabiru and Clelland, 2013, Engebretsen, 2013). Moreover, program managers should take into account the mobility of these urban populations at the point of program design, devising more powerful strategies to maintain contact, even among mobile populations. Finally, we examined literacy and numeracy levels of girls who had never been to school at endline. While the success of the government's enrolment campaign appears to be similar across study sites, selectivity in who enrolled versus who did not may have influenced our results.

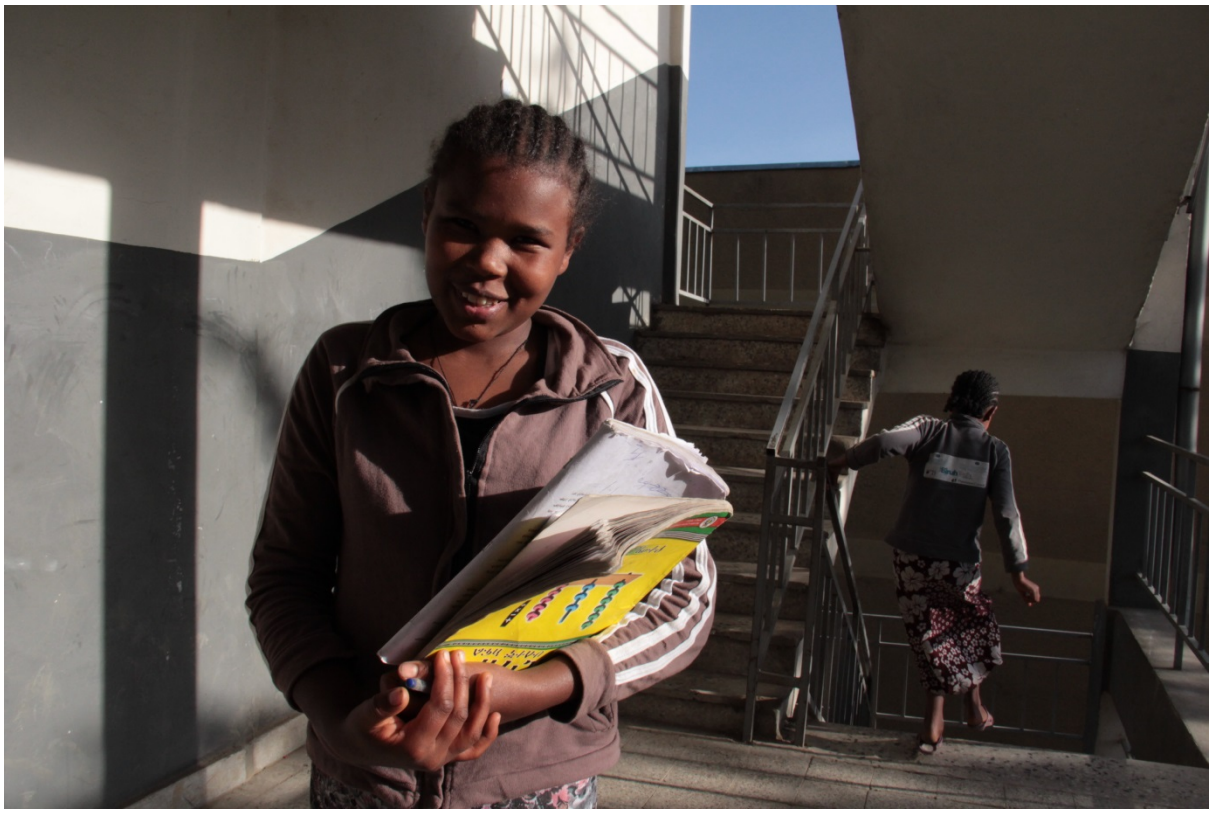
In just six months of implementation, the expansion phase of Biruh Tesfa extending into emerging low-income areas (entitled Powering Up 'Biruh Tesfa'), engaged over 3,000 of Addis Ababa's poorest and most marginalized adolescent girls. Service statistic data from this expansion show that the average age of participants was 13 and nearly half of the participants were child domestic workers. Such groups of very young adolescents and domestic workers were frequently thought to be out-of-reach of support programs such as these. However, as a single-sex, safe spaces approach with targeted, house-to-house recruitment, Biruh Tesfa has proven to be highly effective in engaging girls who otherwise lack platforms through which to gain skills and build networks. In addition to creating platforms through groups for disenfranchised girls, building a bridge with medical facilities and higher status women leaders/mentors facilitates' girls access to entitlements, that they otherwise lack access to because of lack of knowledge, confidence or experience. In addition, participation in our program underscores the potential for focussing investments to younger girls aged 10 to 14, around the time of puberty, which is a pivotal time in a girls' life.

Between baseline and endline, participation in non-formal education increased significantly in the experimental group compared to the control group. Further investigation demonstrated that most of the excess participation in non-formal education was attributable to Biruh Tesfa, suggesting that the program was responsible for engaging girls in learning activities who would otherwise not have been able to attend. No significant results were detected in terms of learning outcomes in the population as a whole, as significant improvements in literacy and numeracy were recorded across both groups. This was likely related to an intensive Ministry of Education campaign to reintegrate young people back into the formal schooling system.

Analysis of the sub-group for whom the campaign had not been successful (girls who had still not attended school at endline, despite the campaign) demonstrated significant gains in the aggregate reading and math scores among the experimental group, after controlling for baseline scores, age, migration status and relationship to household head. This finding underscores the impact Biruh Tesfa made on the educational outcomes of the most marginalized group of girls, those who were still not in formal schooling, even after an intensive reintegration campaign by the Ministry of Education.

At endline, girls residing in Biruh Tesfa areas were significantly more likely to have visited a health facility in the previous six months, compared to respondents in the control group. The effect of the program (including provision of a medical voucher) remained after adjusting for baseline health service utilization and other factors. Moreover, among girls who reported using the voucher, 70 percent were first time users of the health system, highlighting that the voucher was instrumental in introducing girls into the formal health sector. This finding is consistent with another evaluation of a maternal health voucher in Kenya, which demonstrated increases in institutional delivery in locations where the voucher was offered (Obare, Warren, Abuya, et. al., 2013).

Biruh Tesfa demonstrates that it is possible to mobilize large numbers of the most marginalized girls through a safe spaces approach, even those in child domestic work. Results indicate that safe spaces platforms can be instrumental in engaging even the most marginalized and isolated girls in both the education and health sectors, while improving learning outcomes.



REFERENCES

Bandiera O, Buehren N, Burgess R, Goldstein M, Gulesci S, Rasul I, Sualiman M. 2012. Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda, enGender Impact: The World Bank's Gender Impact Evaluation Database. Washington DC : World Bank.

Brady M. 2005. Creating Safe Spaces and Building Social Assets For Young Women In The Developing World: A New Role For Sport. Women's Studies Quarterly 2005, vol.33, no.1&2, pp. 44-45.

Edmeades J, Hayes R, Gaynair G. 2014. Improving the lives of married adolescent girls in Amhara, Ethiopia, Washington DC: International Center for Research on Women.

Engebretsen S. 2013. Evaluation of Filles Eveilles (Girls Awakened): A pilot program for migrant adolescent girls in domestic service, New York: Population Council.

Erulkar A. 2014. Building the assets to thrive: Addressing the HIV-related vulnerabilities of adolescent girls in Ethiopia. New York: Population Council.

Erulkar A, Ferede A, Girma W, Ambelu W. 2013. Evaluation of 'Biruh Tesfa' (Bright Future) HIV Prevention Program for Vulnerable Girls in Ethiopia, Vulnerable Child and Youth Studies, 8 (2) 182-192.

Erulkar A, Gebru H, Mekonnen G. 2011. Biruh Tesfa ('Bright Future') Program Provides Domestic Workers, Orphans & Migrants in Urban Ethiopia with Social Support, HIV Education & Skills, Program brief, Addis Ababa: Population Council, March.

Erulkar A, Tamrat T. "Evaluation of reproductive health programs to support married adolescent girls and their husbands in rural Ethiopia," African Journal of Reproductive Health, 18 (2): 68-76.

Levine R, Lloyd C, Greene M, Grown C. 2008. Girls Count: A Global Investment & Action Agenda. Washington, D.C.: Center for Global Development.

Marston M, Beguy D, Kabiru C, Clelland J. 2013. Predictors of sexual debut among young adolescents in Nairobi's informal settlements, International Perspectives on Sexual and Reproductive Health, 39 (1), March.

Obare F, Warren C, Abuya T, Askew I, Bellows B. 2014. Assessing the population-level impact of vouchers on access to health facility delivery for women in Kenya, Social Science and Medicine, 102: 183-189.

Plan International. 2007. Because I am a girl: State of the world's girls 2007, Surrey UK: Plan International.

Safe space network, http://safespacenetwork.tumblr.com/tagged/mod, Accessed April 27, 2014.

Selim, Mona, Nahla Abdel-Tawab, Khaled Elsayed, Asmaa el Badawy, and Heba el Kalaawy. 2013. The Ishraq Program for out-of-school girls: from Pilot to scale-up. Cairo: Population Council.

Temin M, Levine R. 2009. Start with a girl: A new agenda for global health, Washington, D.C.: Center for Global Development

UNICEF. 2013. Ministry of Education in Ethiopia Launches Back to School Awareness Campaign, September, http://www.unicef.org/ethiopia/events_13458.html, Accessed April 27, 2014.

POPULATION COUNCIL
Ideas. Evidence. Impact.

popcouncil.org