Pre-Analysis Plan:

Assessing the impact of the HIV prevention tool “Join-In-Circuit on AIDS, Love & Sexuality” in Zambia

Anselm Rink and Felix Elwert*

*Corresponding author: I received valuable feedback from Donald Green, Macartan Humphreys and Monica Grant. Please send correspondences to Anselm Rink at afr2132@columbia.edu.

Contents

11 Background

In Zambia, 13% of the 15 to 49 year old population lives with HIV. The highest num-ber of new HIV infections is among young people (Central Statistical Office (CSO) [Zambia], 2015). To counter the spread of the disease, developmental and governmen-tal actors are increasingly relying on educational behavior change tools. A particularly widely used tool, implemented by the German Development Corporation (henceforth, GIZ), is the so-called "Join-In-Circuit on AIDS, Love Sexuality" (JIC). The JIC uses exercises, pictorial aids and edutainment strategies in an interactive circuit with six sta-tions to teach kids about HIV. The tool aims to improve a) HIV and sexual reproductive health knowledge, b) HIV testing uptake, and c) demand for health services. Previous research (Rink and Wong-Grunwald,¨ 2017) has investigated the direct effect of the JIC on knowledge about STIs as well as self-reported sexual behavior in Zimbabwe, and has found positive effects in both domains.

The present research project evaluates the JIC in Zambia. We randomly assign 170 participating schools to five different JIC treatment arms. The first two arms represent control schools. Here, no JIC will be implemented. The third arm implements the JIC among a random subset of students. The fourth arm implements the JIC among indegree central students. The fifth arm implements the JIC among edge betweeness central students. In each school, the JIC will be implemented in one pre-determined grade. Within each school at least 30 students will be selected. For larger schools, 20 per cent of students in the selected grade are selected.

The design allows us to tackle four distinct questions:

1. For each of the three targeting strategies, what is the overall effect of the JIC on all students?

2. For each of the three targeting strategies, what is the total effect of the JIC on participating students?

3. For each of the three targeting strategies, what is the spillover effect of JIC on non-participating students?

4. What is the most effective targeting strategy to maximize the above effects?

12 Design

2.1 Population

The study's population is the full universe of schools in the Choma and Livingstone provinces of Zambia (a total of 218 schools). The schools were provided to the authors by the Zambian Ministry of Education. 28 schools were randomly selected for a pre-test RCT (randomly assigning 14 schools to treatment, and 14 schools to control). For this reason, the final number of schools analyzed in this study is 204 (excluding already treated schools from the pilot RCT). The schools are split between primary (grade 1 to 7) and secondary schools (grade 8 to 12). In order to not interfere with exam peri-ods, the Ministry of General Education and GIZ decided to implement the JIC among children in grade 6 in primary schools, and grade 11 in secondary schools. Before the intervention, we contacted all 204 eligible schools. During this process, several schools were found to not exist, to be double-counted or to be a different institution altogether (e.g., kindergartens mislabeled as primary schools). The final set of schools was then 170.

The descriptive statistics of the school sample is given in Table 1. On average, the schools have 80 students[1] and 74 percent of the sample are in Choma. 88 percent of schools are considered rural, i.e., they do not lie in a major town. 4 percent of the sample have received a JIC before. The school's funding is as follows. The vast majority (80 percent) are funded by the Government of the Republic of Zambia (GRZ).

[1] We caution that the data behind this variable is highly noisy and based on estimates. For this reason, the data was not used when implementing our block-randomization scheme.

3 percent are funded by the community, 12 percent are privately funded and 5 percent are funded through non-standard grants. Schools' access to governmental and health facilities is captured by three variables. First is the distance to the nearest District Board of Education (DEBS, average of 36 kilometers). Second is the distance to the nearest so-called youth-friendly corner (YFC, average of 20km). Third is the distance to the nearest health facility (HF, average of 4km).

Table 1: Descriptive Statistics of Schools

		N	Mean	SD
School characteristics				
	Students	169	79.0	58.4
	Choma	170	73.5	44.2
	Primary	170	88.2	32.3
	Rural	170	60.0	49.1
	JIC before	170	3.5	18.5
Funding				
	GRZ	157	80.3	39.9
	Community	157	3.2	17.6
	Private	157	12.1	32.7
	Grant	157	4.5	20.7
Location				
	Lat	160	-16.8	2.8
	Lon	160	26.6	0.6
Facilities				
	Nearest DEBS (km)	169	36.4	34.8
	Lat nearest YFC	160	-17.1	0.5
	Lon nearest YFC	160	26.7	0.5
	Nearest YFC (km)	160	19.7	22.2
	Lat nearest HF	160	-17.0	0.6
	Lon nearest HF	160	26.7	0.5
	Nearest HF (km)	160	4.0	3.8

Notes: The Table shows the descriptive statistics of the school sample, showing the overall sample size (N), the average (Mean), and the stan-dard deviation (SD) of the indicated variables.

2.2 Randomization

The 170 schools were randomly assigned to one of five treatment conditions.

1. Control 1 (No intervention) (34 schools)
2. Control 2 (No intervention) (34 schools)[2]
3. Treatment 1 (Random selection) (34 of schools)
4. Treatment 2 (Indegree centrality) (34 of schools)
5. Treatment 3 (Betweenness centrality) (34 of schools)

To leverage pre-treatment covariate data on the schools, we relied on an algorithmic block-randomization procedure proposed by Ryan Moore (2016). Specifically, we blocked on the province (Choma), the school type (Primary), urbanization (Rural) and whether a JIC was implemented before. All variables should, in theory, predict whether the JIC leads to positive outcomes pertaining to STIs. In particular, we would expect schools in Livingstone (an urban center), secondary schools, urbanized schools and those that have not seen a JIC implemented before to have the greatest potential to improve knowledge and self-reported behavior as a result of the JIC. The precise code for the blocking is attached to this application in the Supplementary Information. In Table 2, we report the resulting balance across treatment statuses. The table shows that our block randomiza-tion scheme produced very good balance across the treatment statuses for observable pre-treatment covariates.

[2]We included two random control groups so that one could be dropped in case a possible funding crunch. Indeed, during the baseline, some schools in Control 2 were dropped due to a funding shortage. While we were informed by the survey group that there was no precise selection criterion for dropping the schools, we will nonetheless exclude the schools in select robustness tests.

Table 2: Balance across treatment conditions

	Random		Indegree centrality		Betweenness centrality		Control 1		Control 2	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
School characteristics										
Students	66	50	94	55	78	53	71	52	86	76
Choma	71	46	71	46	74	45	76	43	76	43
Primary	88	33	88	33	88	33	88	33	88	33
Rural	59	50	59	50	59	50	62	49	62	49
JIC before	3	17	3	17	3	17	3	17	6	24
Funding										
GRZ	78	42	82	39	80	41	77	42	83	38
Community	0	0	6	24	3	18	3	18	3	18
Private	16	37	12	33	17	38	13	34	3	18
Grant	6	25	0	0	0	0	6	25	10	30
Location										
Lat	-16	6	-17	1	-17	1	-17	0	-17	0
Lon	27	1	27	1	27	1	27	1	27	1
Facilities										
Nearest DEBS (km)	41	39	37	38	31	30	36	35	37	34
Lat nearest YFC	-17	0	-17	0	-17	0	-17	0	-17	0
Lon nearest YFC	27	1	27	1	27	0	27	0	27	0
Neraest YFC (km)	21	24	19	23	20	20	17	20	22	24
LAT nearest HF	-17	1	-17	1	-17	1	-17	0	-17	1
Lon nearest HF	27	1	27	0	27	0	27	0	27	0
Nearest HF (km)	4	4	4	4	4	4	4	4	3	3

Notes: The Table shows descriptive statistics of five different samples showing the average (Mean) and the standard deviation (SD) of the indicated variables.

13 Surveying

3.1 Baseline

The baseline data collection, conducted by a team of 10 trained enumerators divided into 3 groups, lasted about two months from March 16, 2017 to April 11, 2017 and from May 24, 2017 to July 11, 2017 for a total of 52 days. Data collection had to be interrupted during the month of April due to a month-long school break. The survey instrument is attached to this document. The prime reason for the baseline survey—in addition to collecting outcomes before the intervention takes place—was to map the entire friendship network in a given grade (more below).

Once data collection started, the survey team found additional ineligible schools that were either unreachable by car, did not have the targeted grades or were closed. Overall, the survey team managed to conduct a baseline in 98 out of the 102 schools in the treatment group (96% of the target), 28 out of 34 schools in the control 1 category, and 7 out of 34 in control 2. The final sample, thus, consists of 133 schools for the baseline survey. 37 of the schools are in Livingstone and 96 schools in Choma. 19 schools are secondary schools and 114 are primary schools. All students in either grade 6 (if primary school) or grade 11 (if lower secondary school) present on the day of baseline were interviewed for a total of 8,276 students.

Since some instruments are sensitive in nature (more below), we opted to adminis-ter self-administered surveys on tablet computers supported by ACASI (audio-assisted computerized assisted self-interview software) technology. Specifically, the survey was loaded unto the tablets and each question was accompanied by an audio file that read out loud the question in the desired language (options were Tonga and English). Stu-dents were given headphones for increased privacy. After hearing the question, students were to tap on the relevant answer. Finding the correct answer was facilitated by a) maximizing the number of questions with a vignette or picture or color coded, and by

b) positioning the answers on the left, middle, or right panel of the tablet and have the audio file explain where to find the answer. Two enumerators were present to help and assist in case of confusion or any other problem.

3.2 Network mapping

The network mapping was done as follows. First, enumerators collected attendance lists at every school for the assigned grade. Second, enumerators administered the survey, which included the following question: “Now, I want to know who are the friends that you spend time with in your grade. I am going to read you the names of all students in your grade. For every name, I want you tell me if the student is a friend of yours that you spend time with.” Enumerators then recorded each friend of a given respondent on a separate list.

In a third step, the list was then passed on to the principal investigator (PI). In treatment group 1, a random subset of students was chosen for treatment. In treatment group 2, the PI used the R package keyplayer (An and Liu, 2016), to identify the students with the highest in-degree centrality D_i, defined as the sum of incoming friendship nominations from the rest of the network. Specifically, $D_i = \Sigma_j w_{ji}$, where w_{ji} indicates friendship nominations from j to i. In treatment group 3, the PI used the same package to identify students with the highest edge betweenness centrality. It is defined as: $B^i = \Sigma^{jk} \frac{g^i}{g_{jk}}$, where g^{jk} is the number of shortest paths between students j and k, and g_{jk}^i gives the amount of those paths running through student i. In other words, key-player selects the group of students who, combined as a pair, are directly connected to the highest number of different students compared to any other combination. The code underlying this procedure, including all realized assignments, is available upon request.

3.3 Outcome measurement

The questionnaire, administered during the base- and endline of the project, includes four substantive sections. The first section (A) administers demographic questions (e.g., gender, religion, etc.). The second section (B) administers questions pertaining to knowledge about sexual and reproductive health. This includes family planning (B1), condoms (B2), HIV/Aids (B3), other STIs (B4). The third section (C) administers ques-tions about self-reported behavior. This includes sex (C1), HIV testing (C2), and access of youth-friendly corners (C3). The final section, only asked during the endline, in-cludes questions about the JIC.

3.4 Endline

Endline data collection will happen approximately six months after the JIC implementation and will administer the same questions to the same grades identified at baseline. At this point, the endline is scheduled to take place in October 2017.

14 Empirical Analysis

Our primary outcomes of interest are knowledge and self-reported behavior. To assess the degree to which the JIC improved both outcomes, we will create an overall index that standardizes all our outcome variables. This includes all questions in sections B1, B2, B3, and B4. In section C, this includes all questions, but question C1-8. In a second step, we will split the overall index into a self-reported behavior (section A and B), and a knowledge index (section C). In a third step, we will split the index further into the following topics: family planning knowledge (B1), condom knowledge (B2), HIV/Aids knowledge (B3), STIs knowledge (B4), sex behavior (C1), self-reported testing behavior (C2), and self-reported health access behavior (C3). To account for the networked nature

of the trial, we estimate the overall effects, total effects, spillover effects, and direct effects of the JIC on our main index and on our sub-indexes, separately for the three targeting strategies and across targeting strategies.

4.1 Overall effect

H1: We hypothesize that the JIC improves the main and sub-index scores for students in treated schools.

Separately for each targeting strategy, we estimate the overall effect of the JIC by comparing all students in treated schools to all students in control schools,

$$Y_{ij} = \beta_0 + \beta_1 T_j + \beta_2' X_{ij} + \epsilon_{ij}$$

where Y_{ij} represents the respective index for individual i in school j; X_{ij} marks a vec-tor of salient predictive school- and individual-level pre-treatment covariates[3]. T_j is a dummy variable denoting whether a given school was treated or not. To ease interpreta-tion and in line with Freedman (2008), we will use ordinary least squares. For selected robustness tests, we will also include school-random effects. Given that randomization is at the level of the school, we will cluster standard errors at the school level. Moreover, given that we pre-specify one main hypothesis, we will not adjust uncertainty levels for multiple comparisons.

In addition, we pre-register the following additional exploratory hypotheses:

- The JIC improves the knowledge index more strongly than the self-reported be-havior index.

[3]Specifically, at the school-level we include: Students, Choma, Primary, Rural, JIC before, GRZ, Private, Nearest DEBS, Nearest YFC, Nearest HF. At the individual-level, we include age (A1), gender (A2), grade (A3), religion (A4), father's primary job (A5), father's literacy (A6 and A7), parents' marital status (A8) and mother tongue (A9).

1• The JIC improves the main index more strongly in urban areas as compared to rural areas.

4.2 Targeting effect

H2: We hypothesize that the overall effect of the JIC is strongest under edge-central targeting, intermediate under indegree targeting, and weakest under random targeting.

We estimate the contribution of targeting to the overall effect with the following equation:

$$Y_{ij} = \beta_0 + \beta_1' S_j + \beta_2' X_{ij} + \epsilon_{ij}$$

where S_j is a vector of indicator variables for random, indegree, and edge central targeted schools, and differences between the elements of β_1 evaluate the differences between targeting strategies.

4.3 Network estimands

Guided by the same hypotheses as above, we also estimate total, spillover, and direct ef-fects of the JIC on the same outcomes, and compare them across targeting strategies. We estimate the total effect of the JIC on targeted students separately for the three targeting strategies analogous to equation (1), except that we now restrict the samples to treated students in treated schools and comparable untreated students in control schools.[4] We estimate the spillover effect of the JIC on non-targeted students separately for the three targeting strategies analogously to equation (1), except that we now restrict the samples

[4] Specifically, under random targeting, we compare randomly targeted students in treated schools to all students in untreated schools; under indegree targeting, we compare treated students in indegree-treated schools to the corresponding share of highest-indegree students in untreated schools; under edge-centrality targeting, we compare treated students in edge-centrality schools to the corresponding share of highest edge-centrality pairs in untreated schools.

to untreated students in treated schools and comparable untreated students in control schools. We estimate the direct effect of the JIC on targeted students under random tar-geting, except that we now restrict the sample to students in random-targeted schools, and estimate

$$Y_{ij} = \beta_0 + +\beta_1 T_{ij} + \beta_2' X_{ij} + \epsilon_{ij} ,$$

where T_{ij} is an indicator equal to 1 if student i in random-targeted school j is treated and 0 otherwise.

The direct effect compares treated students in treatment schools to comparable un-treated students in treated schools. By comparing treated to untreated students in the same school, this effect removes the indirect effect via peers. Note: in our study, how-ever, the direct effect can only be estimated with random targeting, because only with random targeting are treated and untreated students in a given school comparable to each other.

We compare the effectiveness of targeting strategies on the total, spillover, and direct effect at the school level by estimating models analogous to equation (2) without covari-ate controls on the appropriately restricted samples. We note that since the subsamples of treated (or untreated) individuals in the three targeting arms are not comparable, un-adjusted comparisons warrant a causal interpretation only at the group level, and require covariate control (analogous to equation 2) to be interpreted as averages of individual effects.

4.4 Robustness

To ensure that our estimates are robust, we commit to the following additional checks. Specifically, we will re-run the above analyses

1. Excluding all covariates

2. Using randomization inference

15 Final remarks

We commit to:

1• Make all data and code available after the initial publication of any academic paper or 2 years after the experiment is finished, whichever comes first.

2• Answer any questions pertaining to our analysis or to this document.

3• Make any deviations from this PAP explicit in the paper.

16 Supplementary Information

6.1 Blocking code

```
rm(list = ls())
setwd("")

data <- read.csv("population~cleaned.csv")
data$id <-seq(1:length(data[,1]))
set.seed(52273)
library(blockTools)


data$groups = 1

data$nearest~yfc<-ifelse(is.na(data$geo~lat~yfc), 1, 0)


out<-block(data,
           groups = "groups",
           n.tr = 5,
           id.vars = c("id"),
           block.vars = c("choma_vs_livingston", "primary_vs_secondary", "rural vs urban", "jic implemented before"), #we only block on non
                  #"dist_debs_km", "dist_yfc_km", "dist_hf_km"),
                  #"grz", "community", "private", "grant", "geo_lat", "geo_lon",
                  #"male_pupil", "female_pupil","pre_test"),
           algorithm = "optGreedy",
           distance = "mahalanobis",
           level.two=FALSE,
           verbose=TRUE)


assg<-assignment(out, seed = 5212)
outCSV(out)
```

```
treatments<- read.csv("Group1.csv")
treatment1<- cbind(treatments[2], "random")
colnames(treatment1)<- c("Unit", "Treatment")
treatment2<- cbind(treatments[3], "change_maker")
colnames(treatment2)<- c("Unit", "Treatment")
treatment3<- cbind(treatments[4], "change_maker_friends")
colnames(treatment3)<- c("Unit", "Treatment")
treatment4<- cbind(treatments[5], "control1")
colnames(treatment4)<- c("Unit", "Treatment")
treatment5<- cbind(treatments[6], "control2")
colnames(treatment5)<- c("Unit", "Treatment")
treatment <- rbind(treatment1, treatment2, treatment3, treatment4, treatment5)
treatment  <- treatment[order(treatment$Unit),]
treatment <- as.matrix(treatment[2])
data <- data[order(data$id),]
data$treatment <- treatment
table(data$treatment)


data$treatment_change_maker <- c(ifelse(data$treatment=="changé_maker",1,0))
data$treatment_change_maker_friends <- c(ifelse(data$treatment=="change_maker_friends",1,0))
data$treatment_random <- c(ifelse(data$treatment=="random",1,0))
data$treatment_control1  <- c(ifelse(data$treatment=="control1",1,0))
data$treatment_control2  <- c(ifelse(data$treatment=="control2",1,0))

#write.csv(data, "randomization.csv")
```

6.2 Survey instrument

Section X. Logistics and Network Measurement

Questions to be asked and/or filled in by enumerator.

X1	*Fill in the name of the school*	Name of school _____
X2	*Fill in the ID of the school*	School ID _____
X3	*Ask student*: What is your first name? *Confirm with student that name was spelled correctly*	Student first name _____
X4	*Ask student*: What is your last name? *Confirm with student that name was spelled correctly*	Student last name _____
X5	Enter ID of student from enrollment roster	Student ID _____
X6 (baseline AND endline!)	*Ask student*: Now, I want to know who are the friends that you spend time with in your grade. I am going to read you the names of all students in your grade. For every name, I want you tell me if the student is a friend of yours that you spend time with. [*Prompt: Slowly read through enrollment list and prompt for answer at every name (pause, look, "is this a friend that you spend time with?"). Whenever student labels individual friend, put down number on the right.*] Protocol for AIR: At every school, obtain enrollment protocol from school officials. The protocol needs to be numbered such that every student has a unique ID. Each enumerator then gets one sheet. This requires making copies (by hand or electronically). AIR should write up a protocol on how this is undertaken.	ID 1 _____ ID 2 _____ ID 3 _____ ID 4 _____ ID 5 _____ ID 6 _____ ID 7 _____ ID 8 _____ ID 9 _____ ID 10 _____ ID 11 _____ ID 12 _____ ID 13 _____ ID 14 _____ .....

Section A. Demographic Characteristics

First, we have some general questions about your background.

A1	What day, month and year were you born?	Day Month Year
A2	What is your gender?	Male Female
A3	What grade are you in?	Number _____
A4	What is your religion?	Islam Catholic Apostolic Pentecostal

		Anglican Christian, other Other Don't know
A5	What is your father's primary job?	Farmer ...
A6	Does your father have a hard time reading (e.g., a newspaper or a letter)?	Yes No
A7	Can your father write in English?	Yes No
A8	Are your parents married to each other?	Yes No
A9	What is your native language?	Nyanja Lozi Tonga Kaonde English Luvale Lunda Other Don't know

Section B: Knowledge about sexual and reproductive health

Section B1: Family planning / contraceptives

Now, we have some questions about contraception. By that we mean ways in which one can avoid getting pregnant.

B1-1	Have you heard of family planning before?	Yes No		
	People have different opinions about cotraception. Do you think the following statements are true, false or do you not know?			
		True	False	Don't Know
B1-2	There are free contraceptives available in Zambia.	1	2	3
B1-3	Only married women are allowed to use contraceptives.	1	2	3

B1-4	Men have the right to to determine when a woman should become pregnant.	1	2	3
B1-5	Women are very unlikely to get pregnant if the man correctly uses a condom.	1	2	3
B1-6	Women are very unlikely to get pregnant if the woman washes her vagina after having sexual intercourse.	1	2	3
B1-7	A woman is very unlikely to get pregnant the very first time she has sexual intercourse.	1	2	3
B1-8	It is recommended that all girls visit the health facility for yearly check-ups after they become sexually active.	1	2	3
	Next, we have some questions about particular family planning methods.			
B1-9	Do you know what an intrauterine device (IUD) is?		Yes No → B1-12	
	For the following two questions, please decide whether you think each statement is true or false or indicate that you don't know.			
		True	False	Don't Know
B1-10	IUD carries a risk of infertility	1	2	3
B1-11	IUD is a short-term method	1	2	3
B1-12	Do you know what an injectible contraceptive is?		Yes No → B1-15	
	For the following two questions, please decide whether you think each statement is true or false or indicate that you don't know.			
		True	False	Don't Know
B1-13	Injectables only last one month	1	2	3
B1-14	If you stop taking injectables, you still cannot get pregnant for at least a few months	1	2	3
B1-15	Do you know what birth control pills are?		Yes No → B2	

	For the following two questions, please decide whether you think each statement is true or false or indicate that you don't know.			
		True	False	Don't Know
B1-16	Women should take birth control pills every day	1	2	3
B1-17	If women miss just one or two birth control pills, it does not matter	1	2	3

Section B2: Condoms

Now, we are going to ask you a few questions about condoms.

	People have different opinions about condoms. Please tell me if you think each statement about condoms is true or false or whether you don't know.			
		True	False	Don't Know
B2-1	Condoms prevent women from getting pregnant.	1	2	3
B2-2	Condoms can be used more than once.	1	2	3
B2-3	Condoms prevent men and women from getting HIV/AIDS.	1	2	3
B2-4	Condoms prevent men and women from getting other sexual transmitted infections.	1	2	3
B2-5	Condoms are suitable only for steady relationships.	1	2	3
B2-6	If a girl wants to use a condom, that means she does not trust her partner.	1	2	3
B2-7	Condoms cannot help against sexually transmitted infections.	1	2	3

Section B3: HIV and AIDS

Now, we are going to ask you a few questions about HIV/AIDS.

B3-1	Have you ever heard of HIV/AIDs?	Yes No → B4		
	Please tell me whether you think the statement about HIV/AIDS is true, or false, or whether you don't know.			
		True	False	Don't Know

B3-2	You can see whether a person has HIV/AIDS.	1	2	3
B3-3	It is possible to cure AIDS.	1	2	3
B3-4	Sleeping with a virgin can cure AIDS.	1	2	3
B3-5	A mother can transmit HIV to her child when she is pregnant.	1	2	3
B3-6	You can get AIDS by sharing food with a person who has AIDS.	1	2	3
B3-7	You can get AIDS because of witchcraft.	1	2	3
B3-8	You cannot get HIV the first time you have sex.	1	2	3
B3-9	You can get HIV through kissing.	1	2	3
B3-10	A healthy person cannot get AIDS.	1	2	3
B3-11	AIDS can be transmitted by mosquitos.	1	2	3
B3-12	Everybody is going to get AIDS at some time in his or her life.	1	2	3

Section B4: Sexually transmitted infections (STIs)

Next, we want to ask you about sexually tranmitted diseases other than HIV/AIDS.

B4-1	Have you ever heard about sexually transmitted diseases other than HIV/AIDS?	Yes No → C		
	Please tell me whether you think the statement about sexually transmitted diseases is true, or false, or whether you don't know.			
		True	False	Don't know
B4-2	In Zambia, only sex workers have sexually transmitted diseases.	1	2	3
B4-3	You can always see if somebody has a sexually transmitted disease.	1	2	3
B4-4	Male circumcision reduces the risk of getting sexually transmitted diseases.	1	2	3
B4-5	Sexual transmitted diseases can cause infertility	1	2	3
B4-6	Sexually transmitted disease can never be cured.	1	2	3

	Please tell me whether you think it is true or false if a person can get tested for sexually transmitted diseases at the following places.			
		True	False	Don't Know
B4-7	Public clinic	1	2	3
B4-8	Private clinic	1	2	3
B4-9	In church	1	2	3
B4-10	Some NGO clinics	1	2	3
B4-11	Drug shop	1	2	3
B4-12	Gas station	1	2	3
B4-13	At the witch doctor	1	2	3

Section C: Self-reported behavior

Section C1: Sex

Now, we are going to ask you direct questions regarding your sexual experiences. If a question makes you feel uncomfortable, you can skip the question.

C1-1	In the last six months, did you visit a health clinic to talk about issues relating to sex or family planning?	Yes No → C1-3
C1-2	When was that?	1 month ago 2 or 3 months ago 3-6 months ago 6-12 months ago More than 12 months ago
C1-3	In the past 6 months, have you been to the health facility as a patient?	Yes No
C1-4	(Only males) Are you circumcised?	Yes No → C1-6
C1-5	When were you circumcised?	1 month ago 2 or 3 months ago 3-6 months ago 6-12 months ago More than 12 months ago
C1-6	Have you ever had sexual intercourse?	Yes → C1-8 No → C1-7

C1-7	What is the main reason that you have not had sex? → After, move to section C2.	No opportunity yet No sex before marriage Afraid of getting pregnant Afraid of getting diseases I want to be abstinent Don't know
C1-8	What was the main reason you have had sex?	Married Love Fun/curiosity Friend influenced me I received money in return I wanted to get pregnant I was forced Don't know
C1-9	How many sexual partners have you had in the last 12 months?	0 1 3-5 More than 5
C1-10	Do you have a steady sexual partner?	Yes No
C1-11	The last time you had sex, did you or your partner use a condom?	Yes No → C1-13.
C1-12	Do you use condoms every time you have sex?	Yes No
C1-13	Are you currently using any family planning method or contraception?	Yes No

Section C2: HIV testing ***Now, we are going to ask you questions about HIV testing.***		
C2-1	Have you ever been tested for HIV?	Yes No → C3
C2-2	In the last six months, have you been tested for HIV?	Yes No → C3

C2-3	How many months ago was that?	1 month ago 2 or 3 months ago 3-6 months ago 6-12 months ago
C2-4	Do you know your test result?	Yes No

Section C3: Youth-friendly corners ***Now, we have two questions about youth-friendly corners.***		
C3-1	Have you ever heard of a youth friendly corner at the health facility?	Yes No → D
C3-2	In the past 12 months, have you visited a youth friendly corner at the health facility?	Yes No

Section D: JIC (endline only) ***Now, we have a few questions about the Join in Circuit (JIC).***		
D1	Are you currently wearing a wristband from the JIC?	Yes No
D2	Have you participated in a Join-In Circuit run in the last six months?	Yes → D5 No
D3	Have you heard about the Join-In-Circuit?	Yes No
D4	Did any of your friends talk to you about the JIC or things they learned at the JIC? → After question, skip to D12	Yes No
D5	On a scale from 1 to 5, how interesting did you find the JIC? 1 means you didn't find it interesting at all and 5 means you found it very interesting.	Number _____
D6	After the JIC, have you gone to visit a health facility because of what you learned during the JIC?	Yes No
D7	Did the JIC make it easier to talk about sex with your friends or partner?	Yes No
D8	Did the JIC made you uncomfortable or embarassed at times?	Yes No

D9	On a scale from 1 to 5, how much new information did you learn at the JIC? 1 means you didn't learn anything new and 5 means that you learned a lot of new information.	Number _____
D10	How many of your friends have you talked to about things that you learned at the JIC?	Number _____
D11	On a scale from 1 to 5, how much did you enjoy the JIC? 1 means you did not enjoy the JIC at all and 5 that you enjoyed the JIC very much.	Number _____
D12	We are done! Thank you very much for your participation in this survey!	

End of Data Collection Tool

References

An, Weihua and Yu-Hsin Liu. 2016. keyplayer: Locating Key Players in Social Net-works. R package version 1.0.3.
URL: https://CRAN.R-project.org/package=keyplayer

Central Statistical Office (CSO) [Zambia]. 2015. “Zambia Demographic and Health Survey 2013-14.”.
URL: https://www.dhsprogram.com/pubs/pdf/FR304/FR304.pdf

Freedman, David A. 2008. “On regression adjustments in experiments with several treatments.” The Annals of Applied Statistics 2(1):176–196.

Moore, Ryan T. 2016. blockTools: Blocking, Assignment, and Diagnosing Interference in Randomized Experiments. R package version 0 6-3.

Rink, Anselm and Ramona Wong-Grunwald¨. 2017. “How effective are HIV behaviour change interventions? Experimental evidence from Zimbabwe.” Journal of Develop-ment Effectiveness 120(2):1–28.