

July 2022

# Tracking Girls' Agency:

An Outcome Evaluation of Diverse Programming Amongst  
Community-Driven Organizations in East Africa

An **AMPLIFY Girls** Research Brief Prepared by: Aubryn Allyn Sidle, Brenda Oulo, Ninjel Osano





# TABLE OF CONTENTS

**5** EXECUTIVE SUMMARY

---

**7** INTRODUCTION

---

**8** LITERATURE REVIEW

---

**9** STUDY SETTING

---

**10** DATA AND METHODS

---

**13** RESULTS

---

**20** DISCUSSION

---

**21** LIMITATIONS

---

**21** CONCLUSION

---

**22** REFERENCES

---

**23** APPENDIX

AMPLIFY Girls' mission is to amplify the voices, work, and collective impact of community-driven organizations focused on the power and potential of adolescent girls. We do this by strengthening organizational efficiencies, creating opportunities for collaborative learning, building evidence and engaging in global policy and practice fora. AMPLIFY Girls is a collective of partner organizations, governed by a Board of Directors the majority of whom are elected from amongst the membership.

Suggested Citation: Sidle, A.A., Oulo, B., & Osano, N. 2022. "Tracking Girls' Agency: An Outcome Evaluation of Diverse Programming Amongst Community-Driven Organizations in East Africa" AMPLIFY Girls Research Brief. Nairobi, Kenya.

Aubryn Allyn Sidle (Ph.D.) is a researcher at AMPLIFY Girls and served as Principal-Investigator on this study.

Brenda Oulo (M.Sc., MBChB.), is a researcher at AMPLIFY Girls and served as the Co-Investigator on this study.

Ninjel Osano, (B.S.) is a research assistant at AMPLIFY Girls.

## ACKNOWLEDGEMENTS

This report was possible due to the hard work and partnership of the participating AMPLIFY Girls' partner organizations, who implemented this survey during an incredibly challenging year. These organizations are listed here alphabetically: AkiraChix, Girl Up Initiative Uganda (GUIU), Girls Livelihoods and Mentoring Initiative (GLAMI), Jifundishe, Komera, Learn Work Develop (LWD), Riley Orton Foundation, Sega Girls, Streets Ahead Children's Center Association (SACCA), The Action Foundation (TAF), The Girls Foundation of Tanzania (TGFT) and Wiser Girls. We are also grateful for the feedback and guidance of our colleagues in the AMPLIFY Girls Secretariat, and would like to thank our regional research liaisons, Dr. Alfred Kitawi (Strathmore University), Dr. Eugenia Kafanabo (University of Dar es Salaam), Dr. Wenceslas Nzabwirwa (University of Rwanda), and Dr. John Bosco Isunju (Makerere University), all of whom helped to guide this work in each of our study countries. Last, but not least, we'd like to extend our gratitude to Echidna Giving, whose generous support helped to fund this study.

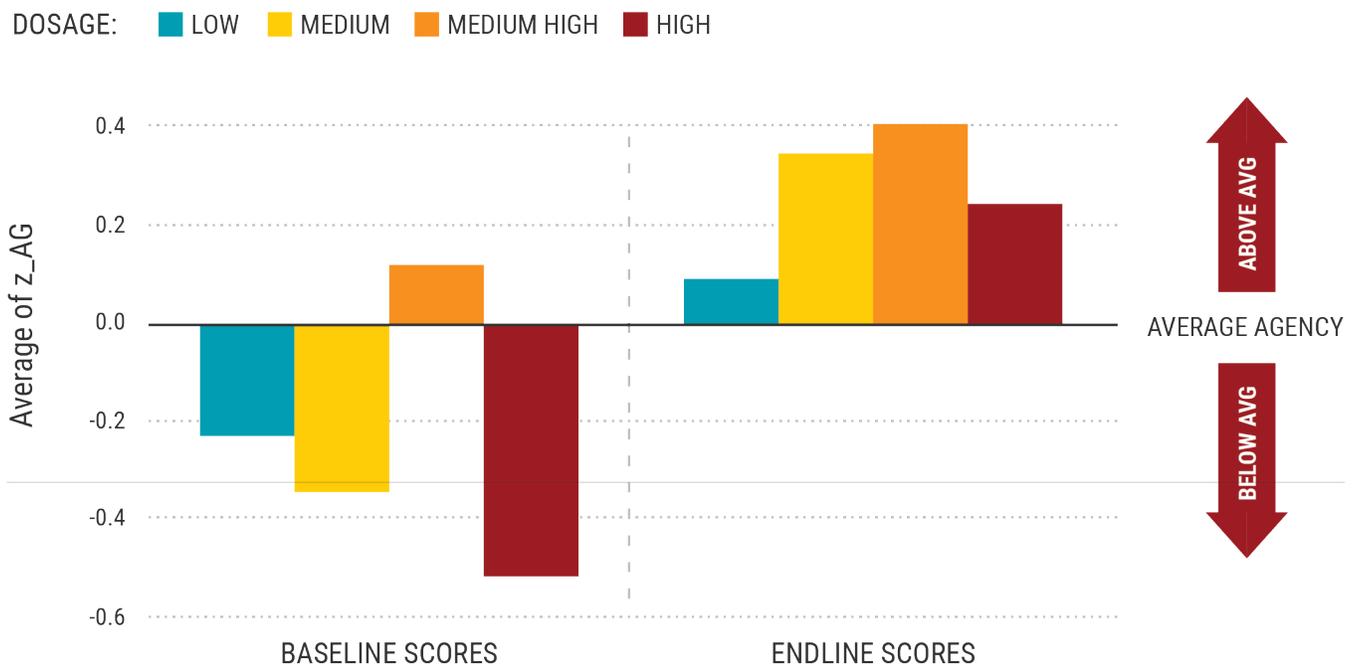
## EXECUTIVE SUMMARY

Community-driven organizations (CDOs) are best positioned to understand the complexity of local context and thus design interventions that will champion gender-transformative change in their communities. In this paper, we offer evidence documenting how the AMPLIFY Girls’ collective of CDOs has impacted such change for girls through an outcome evaluation, tracking these diverse programs’ effects on girls’ agency.

We evaluate agency outcomes in 13 programs and four countries, in a 12-month longitudinal study that utilizes a locally developed and validated psychometric measure of agency (Sidle & Oulo, forthcoming). We use country fixed effects regression modelling to understand the effect of age, program dosage, duration, and baseline agency score on increasing girls’ agency. The study included 899 female beneficiaries, attending programs that were classified into four categories of program dosage (low, medium, medium-high, or high). Our findings show that AMPLIFY Girls’ organizations are significantly increasing girls’ agency scores in every dosage category. On average, organizations participating in the evaluation increase girls’ agency by 6.55% over the course of the evaluation period.

**FIGURE 2**

### CHANGE IN GIRLS’ STANDARDIZED AGENCY SCORES (Z\_AG) BASELINE TO ENDLINE



Furthermore, in country fixed effects regression modelling of girls' agency gains, we found that

- » Girls with the lowest quartile agency scores at baseline had the highest increases in agency. These effects held across all our models, and baseline agency score remained the single biggest predictor of agency outcomes.
- » Medium dosage programs were associated with higher agency gains as compared to high-dosage programs, suggesting that regular, weekly, agency-focused engagement has a bigger impact on girls' agency over course of the year than other, higher intensity delivery models that cover agency as part of a broader curriculum.
- » Each additional month of programming was associated with an increase in girls' agency gains. The effect size of program duration between those that were three months long and 12 months long was similar to the effect size of the medium dosage program structure.
- » Age was a significant predictor of girls' agency outcomes only before accounting for baseline agency score, indicating that the age of participant doesn't matter above and beyond a girl's lack of agency.
- » When not accounting for baseline agency, younger adolescents (ages 10–13) are associated with higher increases in girls' agency compared to young women (age 21–25).

Our evaluation yields a number of important recommendations for programs and policies targeting adolescent girls. To maximize impact on girls' agency, our study suggests that programs should:

- Target the most vulnerable, (those who are likely to have the lowest agency);
- Engage in regular, sustained, and medium-touch program delivery with focused content; and
- Consider program durations longer than three months. (Our analysis indicated that every additional month of programming after three months yielded increases in agency gains);
- Conduct more research that evaluates the effect of programs on girls' beliefs about gender and gender norms, as this appears to be the slowest changing dimension of agency in our study.

## INTRODUCTION

Community-driven organizations (CDOs) are best positioned to understand the complexity of local context and thus design interventions that will champion gender-transformative change in their communities. In 2018, AMPLIFY Girls (AG) was founded by 18 CDOs across four countries in East Africa (Kenya, Uganda, Tanzania, and Rwanda) to quantify and document the work of local organizations and to drive more resources to community practitioners at the forefront of development efforts. AMPLIFY's CDO partnership has since grown to 25 organizations, and current collective work continues to strengthen the role of transnational local networks, thus further defying the stereotype that grassroots movements are too narrowly focused on local issues to contribute meaningfully to global sustainable development goals (SDGs).

In this research brief, we focus on AG's work to document the impact of community-driven organizations' interventions that seek to provide women and girls in East Africa with equal access to education, health care, decent work, and representation in political and economic decision-making processes. At inception, Group Concept Mapping (Trochim and Kane, 2005) was used to reach consensus among AG practitioners on what outcomes constitute success for girls in their communities. Agency—defined as one's ability to effect change in one's life and/or community—was identified as a key outcome for adolescent girls. Adopting girls' agency as a focus of collective evaluation efforts, AG collaborated with Dr. Aubryn Allyn Sidle to create the first locally developed measure of girls' agency—the AMPLIFY Girls Agency survey (Sidle et al., 2019)—which was validated by Dr. Sidle and Brenda Oulo in 2020 (Sidle & Oulo, forthcoming). The AG Agency Measure is currently used for program evaluation among all AMPLIFY Girls' partner organizations.

This research brief provides a detailed description of our process and methodology for conducting an outcome evaluation of AG's diverse programs on girls' agency using the jointly developed and validated quantitative measure of agency. The findings we present herein are informed by some assumptions: First, agency can be learnt and therefore can be evaluated as the outcome of programming intended to improve one's ability to effect change in their own lives and community. Second, agency is a complex and contextualized construct often influenced by the evolving opportunities and constraints of one's environment, but one that can be measured by the four constituent constructs of beliefs and skills; self-beliefs, self-governance skills, leadership skills, and beliefs about the malleability of one's environment (Sidle et al., 2019).

The evaluation findings we present demonstrate that the AMPLIFY Girls Agency Measure is sensitive to change over time and to different interventions and, most importantly, that AMPLIFY Girls' organizations are indeed improving girls' agency through diverse and locally developed programming. Further, we provide insights into the structure of programming that can maximize gains in agency over time. We argue that community-driven organizations are as well positioned as transnational actors to drive progress towards global development agendas.

The main evaluation question guiding this review is: **What is the effect of local programming on girls' agency over time?**

## LITERATURE REVIEW: IMPORTANCE OF AGENCY AND LOCAL ACTORS FOR GIRLS' EDUCATION

Soft skills or social emotional learning competencies are an important part of personal development and, as such, have long been considered critical outcomes of educational programs designed to influence one's academic achievement, health, and economic outlook. Consequently, soft skills are core to the curricula of most girls' life skills education programs, regardless of whether these programs seek to improve participants' education, economic, or health outcomes (Nasheeda et al., 2019; Brush et al., 2022). Increasingly, girls' education practitioners are acknowledging that improving outcomes for girls in low-resource settings requires content related to gender and power, and sexual and reproductive health and rights as core to a girl's successful future (Haberland, 2015). The programs included in this review incorporate life skills education content that covers both soft skills and sexual and reproductive health and rights (SRHR), using various programmatic formats to influence a diverse range of development outcomes for adolescent girls. Despite some sector-wide agreement on soft skills and SRHR content, there are no universally accepted frameworks for evaluating the skills and beliefs learned in life skills programs and only a few that have been locally developed (Mugo, 2020) and/or adapted for specific country settings (Room to Read, 2022).

Agency offers a powerful evaluative framework for documenting the outcomes of life skills education programs (Sidle, 2019). At its core, agency is a multi-dimensional construct—made up of both skills and the beliefs that support a girls' change-making capacity (ibid). Viewed as such, the capacity to exercise agency for women and girls is an important development objective, but evidence on how programs affect agency is profoundly lacking. Qualitative approaches remain a popular methodology for evaluating girls' agency and empowerment, however utilizing a quantitative measure of agency provides organizations with the opportunity to monitor trends over time and make comparisons across groups. As a practitioner-defined and contextualized instrument, the AMPLIFY Girls Agency Measure provides a useful evaluation framework for community-driven organizations and is able to adapt in order to document the impact of various life skills-oriented programs. The AMPLIFY Girls Agency Measure conceptualizes agency as a four-dimensional construct made up of the following indicators: self-beliefs, environmental beliefs, self-governance skills, and leadership skills (Sidle & Oulo, forthcoming). Importantly, our measure is reliant on several structurally-dependent aspects of agency—most notably beliefs about the environment, which are tied to a girls' understanding of gender norms, and how rigidly her opportunities are structured by those norms.

Recognizing that girls' and young women's agency is influenced by social structure, we argue that community-driven organizations that operate closest to communities are best positioned to innovate impactful programs that consider the socio-cultural conditions that affect agency (Goldman & Little, 2015). Globally, we are witnessing increasing calls for localization in both the humanitarian and development spaces (Federici et al., 2019; Hickey & Mohan, 2005) as evidence illustrates that social and economic development is dependent on creating and sustaining local organizations (Seyfang & Smith, 2007). In their paper, Gill Seyfang and Adrian Smith suggest that community-level action linked with grassroot innovation is key to sustainable development (ibid) and that female empowerment is a necessary and core concern for development (Konte & Tiriayi, 2019; Lohani & Aburaida, 2017).

The COVID-19 pandemic further accentuated the need for a paradigm shift towards localization, as lockdowns and travel restrictions curtailed the role of large-scale international actors. In many areas of the world, millions of people experienced job losses and the consequent economic impacts on their livelihoods (ILO, 2020). In many developing countries which lack robust government social protection mechanisms, locally-based community actors were able to innovate, and offer the only efforts to mitigate the disproportionate socio-economic impacts of the pandemic, particularly on adolescent girls (Oulo et al., 2021).

In this context, there is an urgent need to evaluate and consider the work of community-level actors as integral to global development efforts. Findings from such evaluations should be a source of learning for all actors, not just local ones, as the focus on global development should avoid downplaying local contributions (Oldekop et al., 2020). In this brief, we contribute evidence on the impact of local actors on girls' agency and make some recommendations about program structure and strategy that best contribute to increasing agency (as measured by a locally-developed measurement tool).

## STUDY SETTING: PROGRAMS, ORGANIZATIONS, AND LOCALE

The programs evaluated in this brief are operated by 12 out of 18 of the 2020 AMPLIFY Girls CDO partners, located in four East African countries, namely Kenya, Tanzania, Rwanda and Uganda<sup>1</sup>. The CDOs' beneficiaries are drawn from diverse contexts ranging from low-income settlement areas in the capital cities of Nairobi (Kenya) and Kampala (Uganda), other catchment areas of smaller cities such as Arusha (Tanzania) and Kisumu (Kenya), to rural agricultural communities in Kayonza and Bugesera districts (Rwanda), Morogoro District and Usa River (Tanzania), and Mhuru Bay (Kenya).



<sup>1</sup>Unfortunately, due to extended COVID lockdowns in Uganda (the longest in the world), very few of our Ugandan partner organizations were able to participate in the study and those that did had only very few beneficiaries who remained in the program during the study period. Two organizations in Kenya operating in very remote settings were also unable to collect endline data.

The adolescent girls who are program beneficiaries of these CDOs are all identified as vulnerable, and face a wide-range of needs and disproportionate risks to their general wellbeing due to the socio-economic, geographic, and social contexts where they reside. The evaluation targeted girls who were enrolled in an AMPLIFY Girls CDO life skills program, and although these programs are incredibly diverse in format and structure, **they can be broadly grouped into three categories of programmatic approach:**

**1**

Girls with the lowest quartile agency scores at baseline had the highest increases in agency. These effects held across all our models, and baseline agency score remained the single biggest predictor of agency outcomes

**2**

Programs that operate low/no-cost private schools for vulnerable girls, and that also provide a host of other wrap-around services in leadership training, economic empowerment, agency, and skills building

**3**

Programs that provide school-based interventions (usually after school or on school breaks) and meet girls' psychosocial, SRH knowledge, and material needs

## DATA AND METHODS: A PROCESS FOR EVALUATING A DIVERSITY OF APPROACHES

We employed a longitudinal study design where participants spent varying lengths of time in CDO programs based on the different programmatic formats, but were assessed at two time points. Program duration varied from three months to 12 months (period ranging from November 2020 to April 2022), and all participants were surveyed upon entry into a program (baseline) and then again after completion of all program components, or at the end of one year (in the case of programs which were longer than 12 months).

To make sense of the diversity of programmatic approaches, we hypothesized that there were likely two aspects of program structure that would most influence agency: program duration and dosage. In our study, duration is simply measured as the number of months a participant spent in the program, while dosage is measured in terms of the approximate number of hours per year a participant spends in life skills-related programming. Interestingly, we found that the four dosage categories corresponded to four distinct programmatic structures, described in Table 1 below.

**TABLE 1**

## DOSAGE CATEGORIES OF PROGRAMS

DOSAGE	HOURS PER YEAR	DESCRIPTION	NUMBER OF PROGRAMS	PARTICIPANTS
LOW	<100	A few weekends a year of intensive programming	1 PROGRAM	50
MEDIUM	100-350	Regular after-school programs meeting for 1-2 hours, 1-2x/week	4 PROGRAMS	219
MEDIUM HIGH	350-1,000	Multiple weeks per year in full-day programming (usually during school breaks)	4 PROGRAMS	256
HIGH	>1,000	Intensive year-round programming (3/4 are in-residence boarding schools)	4 PROGRAMS	369

## STUDY POPULATION

Study participants were adolescent girls and young women aged 10–25 who were recruited and enrolled into one of the 13 AG partner organizations’ programs evaluated by this study and who gave their consent for data collection. AG partners intentionally target program participants who are some of the most vulnerable in the communities that they serve. A few of the dimensions of this vulnerability include: girls who come from low-income households in rural subsistence farming communities, current or former street children, girls living with disabilities, girls residing in high-density urban settlements, teen-mothers, and girls who have limited access to educational opportunities. A brief description of the programs and target populations included in this study can be found in the Appendix in Table 2.

Quantitative data was collected using the AMPLIFY Girls Agency Survey; an instrument developed by community-based practitioners and validated to assess agency as a psychometric construct (Sidle & Oulo, forthcoming). The survey is a 60-item self-report questionnaire which incorporates Likert type responses and was administered in English and two National languages of East Africa—Kiswahili and Kinyarwanda. The AMPLIFY Girls Agency Measure conceptualizes agentic capacity as being comprised of **four key indicators**:

- 1. Self-Beliefs**—a girl’s self-worth combined with her belief in her own abilities.
- 2. Environmental Beliefs**—a girl’s belief in the malleability of her environment combined with her attitudes towards gender and gender norms.
- 3. Leadership Skills**—a girl’s ability to influence others.
- 4. Self-Governance Skills**—a girl’s ability to organize and coordinate her own abilities into action.

## DATA MANAGEMENT & ANALYSIS

Data was cleaned, dropping duplicates and subsequently coding items using statistical software from Statacorp version 17. Descriptive statistics describing the characteristics of population were produced and tabulated.

Summary statistics were evaluated via a crude ‘collective impact’ agency score as a percentage change over time. Composite agency scores were manually calculated out of ‘100,’ where each survey question was weighted based on the contribution of that item to the overall indicator. Using this method, each individual’s agency scores and related factor scores (self-beliefs, environmental beliefs, self-governance skills, and leadership skills) were generated at baseline and endline. Differences between baseline and endline were calculated and represented as a percent change. The unweighted changes in factor scores and standardized agency baseline to endline were plotted in graphs to understand the trend of each over time.

For inferential analysis, the survey was scored as a four-dimensional model of agency (comprised of the four indicators described above) using structural equation modelling (SEM) to generate composite agency scores while adjusting for measurement error at each time period (baseline and endline). The difference between baseline and endline agency scores was calculated and then standardized, producing our outcome variable denoting a girl’s ‘change in agency’ over time in terms of standard deviation units ( $z_{\text{deltaAG}}$ ).

We then modelled the effects of different characteristics on a girl’s change in agency ( $z_{\text{deltaAG}}$ ), in both univariate and multivariate linear regressions with fixed country effects, after dropping outliers in  $z_{\text{deltaAG}}$ . Essentially this method allowed us to ‘control for’ or hold constant the effects of country on girls’ agency outcomes while exploring the effects of different program and individual characteristics (also known as ‘explanatory variables’) on girls’ agency scores. Explanatory variables included in the estimation were participant’s age (in four categories: 10–13, 14–17, 18–21, 22–25yrs), program duration (number of months), program intensity (dosage—see Table 1), and a variable representing a girl’s baseline agentic capacity in terms of lowest to highest quartile. We represent side by side the univariate effects of each of these variables on the outcome (change in agency scores) and the multivariate effects. Statistical significance was set at the .05 level. We also represent graphically the change in raw agency scores and constituent latent variables across dosage categories from baseline to endline.

---

<sup>2</sup>Weights are the factor loadings of each item as derived from principal factor analysis conducted on the validation dataset. See Sidle et al., 2019.

<sup>3</sup>The specified models were estimated using standard estimation procedures in Stata, allowing the adjustment of the composite scores for ignorable missing data by applying common full information maximum likelihood (FIML) estimation.

## ETHICAL CONSIDERATIONS

Research subjects' protection and ethical approvals were sought from the following research and ethics approval committees: Cornell University's Institutional Review Board in the United States, Strathmore University Institutional Ethics Review Committee (SU-IERC) in Kenya, Makerere University Research and Ethics Committee in Uganda, and the University of Rwanda College of Medicine and Health Sciences Institutional Review Board (CMHS-IRB). Research permits were obtained from relevant government authorities in all four study countries.

## RESULTS

### DATA SUMMARY

Table 3 below represents a simple description of the data demographics. Data represents 13 programs and 12 out of 25 AMPLIFY Girls' organizations for a total of 899 matched (baseline-endline) observations. This analysis is of longitudinal data, meaning that the same participants took the survey at baseline (just entering the programs) and at the endline (in the final weeks of programs). Overall, the ages represented in the data ranged from 10 to 25 with a mean age of 19. Data was collected in both rural and urban areas with some organizations being both rural and urban at the same time with regard to location and beneficiaries served, and the majority of responses (822) were from rural communities. While geographically data was collected in all four countries where AMPLIFY Girls has partner presence, 57% of the respondents were from Rwanda and less than 2% were from Uganda (where schools remained closed and programs suspended due to the COVID-19 pandemic until January 2022).

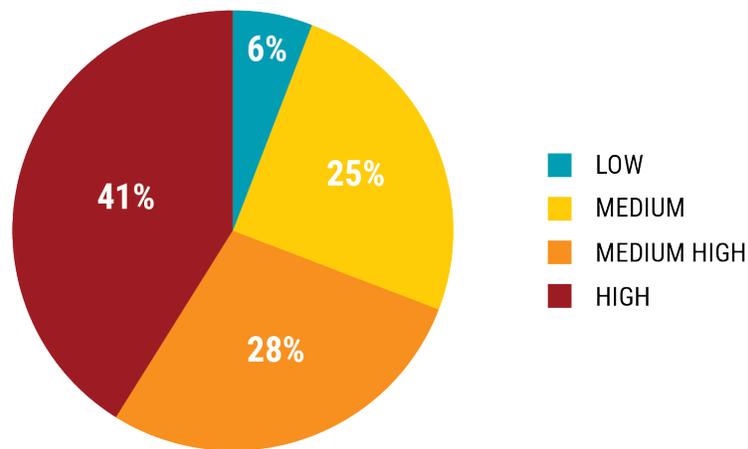
**TABLE 3:**
**DATA SUMMARY**

COUNTRY	TOTAL N BY COUNTRY	PROGRAMS	LOCALE (of participants)	DOSAGE (program intensity)	PROGRAM EVALUATION DURATION	MEAN AGE (range)	TOTAL NUMBER
<b>TOTAL N</b>		<b>13</b>					<b>899</b>
<b>KENYA</b>	<b>215</b>	Codehive	Urban & Rural	Medium High	12 months	21 (17-25)	109
		Mentorship Program	Urban	Low	3 months	14 (10-25)	53
		Smart Girl Program	Urban	High	12 months	12 (11-14)	9
		WISER Houses of Wisdom	Rural	High	12 months	15 (13-17)	47
<b>TANZANIA</b>	<b>159</b>	The Girls Foundation of Tanzania Scholarship	Urban & Rural	Medium High	12 months	14 (12-14)	13
		Sega School	Rural	High	12 months	14 (11-16)	60
		Independent Program	Rural	Medium High	12 months	19 (15-25)	25
		Kisa Scholars	Urban & Rural	Medium	12 months	18 (17-21)	60
<b>RWANDA</b>	<b>505</b>	Sacca: Livelihood	Rural	Medium High	3 months	21 (16-25)	108
		Sacca: Rapid Response	Rural	Medium	6 months	22 (17-25)	80
		Teen Mothers Program	Rural	Medium	12 months	19 (15-22)	66
		Masenge mba Hafi	Rural	High	12 months	21 (16-24)	254
<b>UGANDA</b>	<b>15</b>	Adolescent Girl Program	Urban	Medium	12 months	12 (10-14)	15

In terms of program dosage, four out of 13 programs evaluated represent the medium, medium-high, and high dosage categories, with just (one) program representing low dosage. Accordingly, just 6% of the population were attending low dosage programs, with 25% and 28% attending medium and medium-high dosage programs respectively, and 41% attending high-dosage programs (Figure 1). Program participants were given an endline survey at the end of the program, or at the end of 12 months (whichever came first). Three of the programs evaluated operated for less than 12 months.

**FIGURE 1**

## PARTICIPANT PERCENTAGE BY DOSAGE

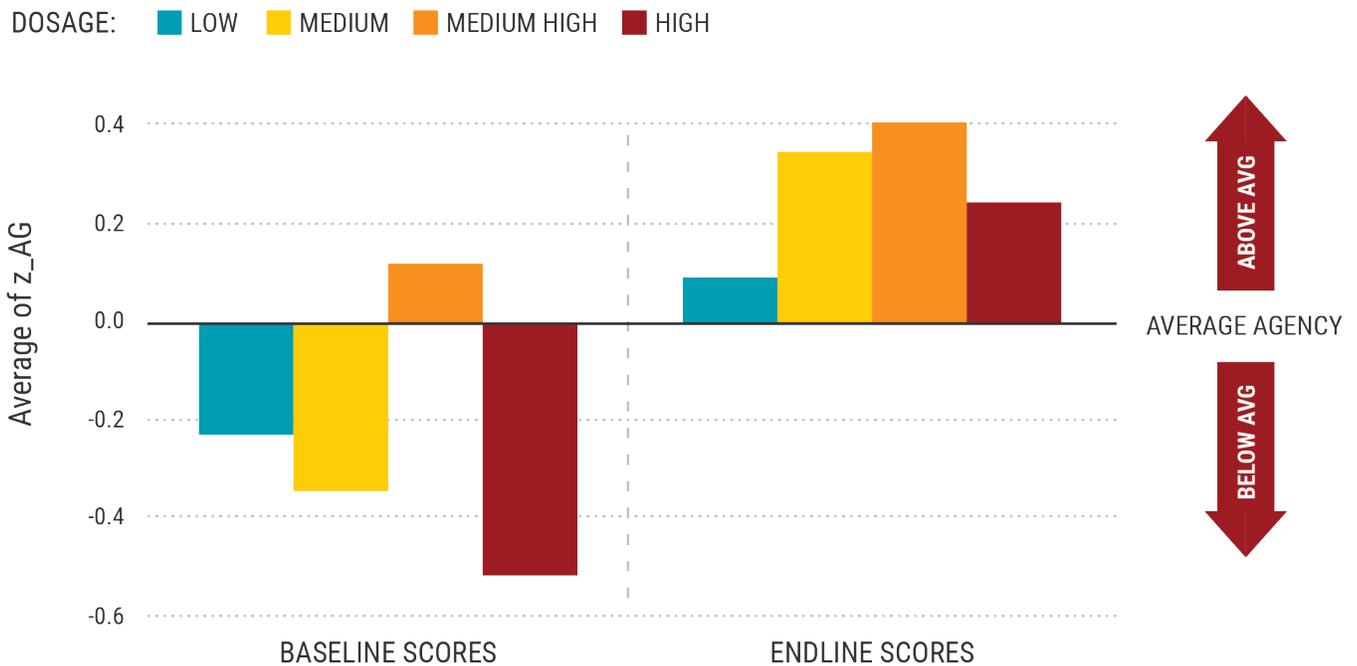


## TRENDS IN AGENCY BASELINE TO ENDLINE

In Figure 1, we show standardized baseline and endline agency scores by dosage category where zero '0' represents the average agency score across the study population. Thus a negative agency score indicates 'below average' and a positive score as 'above average.' Scores were aggregated for all participants by dosage category.

**FIGURE 2**

### CHANGE IN GIRLS' STANDARDIZED AGENCY SCORES (Z\_AG) BASELINE TO ENDLINE



A visual examination of Figure 2 shows clearly that girls' agency scores increased across all dosage categories. Participants attending high dosage programs started with the lowest agency scores, followed by medium dosage participants and low dosage participants. Medium-high dosage participants started with above average agency but still showed increases in their average agency scores at endline. We also performed a difference of paired means test between baseline and endline groups in each dosage category to determine whether these gains were statistically significant or due to random chance, and found that the differences were statistically significant in every dosage category ( $P < .05$ ).

Next we considered the baseline to endline trends of the four agency indicators: self-beliefs, environmental beliefs, self-governance skills, and leadership skills, to see if there were differences in how these four factors contributed to differences in agency scores overall. Table 4 shows the differences in summary statistics of the raw scores from baseline to endline, and the percentage change in agency scores. The last column indicates whether these differences were statistically significant at the five percent level of significance ( $p < .05$ ). We found a statistically significant increase of 6.55% in average agency scores across the study period. Additionally, all four agency indicators, showed statistically significant positive change over time, meaning they were not due to random chance.

**TABLE 4**

### RAW SCORE DIFFERENCES BASELINE TO ENDLINE

	BASELINE N=899	ENDLINE N=899	CHANGE BASELINE TO ENDLINE	P-VALUE
<b>SELF-BELIEFS</b>	24.32/30	26.36/30	<b>+2.04</b>	0.000
<b>ENVIRONMENTAL BELIEFS</b>	6.72/10	7.52/10	<b>+0.81</b>	0.000
<b>SELF-GOVERNANCE SKILLS</b>	22.70/30	24.59/30	<b>+1.89</b>	0.000
<b>LEADERSHIP SKILLS</b>	23.49/30	25.31/30	<b>+1.8</b>	0.000
<b>AGENCY</b>	77.23/100	83.78/100	<b>+6.55%</b>	0.000

**FIGURE 3**

### TRENDS IN AGENCY INDICATORS BASELINE TO ENDLINE

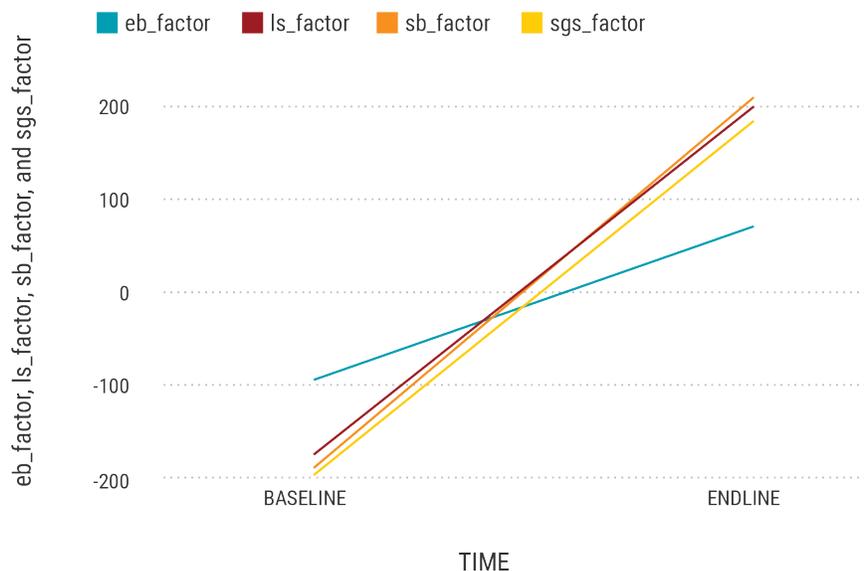


Figure 3 shows the change in factor scores from baseline to endline for all four agency indicators, and agency overall. Leadership skills, self-governance skills, and self-beliefs lines show similar increases over time to agency overall. However, when considered here as a stand-alone construct the environmental beliefs line is notably more constant, indicating that the pace of change of this indicator as an unweighted factor is much slower than the other dimensions of agency.

## EFFECTS OF AGE, DOSAGE, DURATION, AND BASELINE AGENCY ON GIRLS' AGENCY

To understand the effects of program and individual characteristics on girls' agency scores, we constructed linear regression fixed effects models, which hold constant the variable effect of country on girls' agency while considering the contribution of other variables. Table 5 shows the regression results of both univariate and multivariate analysis. In column 1 we show the univariate (singular) effects of age, dosage, program duration, and baseline agency score on changes in agency (z\_deltaAG). In columns 2 and 3 we show how the effects of each of these variables change when considered in combination—first without baseline agency score (column 2) and second with baseline agency score (column 3). Only those coefficients with statistically significant p-values ( $p < .05$ ) should be interpreted as having a meaningful effect on girls' agency. In the case of our continuous variable (duration), the coefficient can be interpreted as the effect of a one-month increase on girls' agency (represented in terms of standard deviation units). In the case of categorical variables (age, program dosage, and baseline agency score), the coefficient can be interpreted as the effect of each specific category, as compared to the reference category, on girls' agency outcomes.

**TABLE 5**

### UNIVARIATE AND MULTIVARIATE EFFECTS ON GIRLS' AGENCY

		COLUMN 1: UNIVARIATE FIXED EFFECTS		MODEL 2: MULTIVARIATE FIXED EFFECTS		MODEL 3: MULTIVARIATE FIXED EFFECTS	
	CATEGORY	COEFFICIENT	PVALUE	COEFFICIENT	PVALUE	COEFFICIENT	PVALUE
OUTCOME z_deltaAG				R <sup>2</sup> = 0.15		R <sup>2</sup> = 0.45	
AGE (years)	10-13	0.181	0.201	0.303*	0.054	0.044	0.726
	14-17	0.159	0.141	0.202	0.072	0.063	0.486
	18-21	0.121	0.111	-0.011	0.883	0.050	0.404
	22-25	Reference Category		Reference Category		Reference Category	
PROGRAM DOSAGE	LOW	-0.148	0.339	1.107**	0.000	0.080	0.726
	MEDIUM	0.084	0.297	0.578**	0.000	0.295**	0.486
	MEDIUM-HIGH	-0.376**	0.000	0.101	0.258	0.040	0.404
	HIGH	Reference Category		Reference Category		Reference Category	
PROGRAM DURATION (months)	N/A	0.056**	0.000	0.123**	0.000	0.028**	0.026
BASELINE AGENCY SCORE (quartiles)	1ST	1.697**	0.000			1.644**	0.000
	2ND	0.792**	0.000			0.793**	0.000
	3RD	0.481**	0.000			0.486**	0.000
	4TH	Reference Category				Reference Category	
N=894   *p<.05   **p<.01							

## UNIVARIATE EFFECTS: LOWER BASELINE SCORES ASSOCIATED WITH HIGHEST GAINS

Above and beyond the effect of country, column 1 suggests that there are no significant effects of age on girls' agency scores when considered in a univariate model. Similarly, there are no significant effects of either low or medium dosage program models when compared to high-dosage programs. Interestingly, medium-high dosage programs are associated with a .376 standard deviation decrease in overall agency score when considered alone. Each month increase in program duration is associated with a positive and significant effect on girls' agency, while baseline agency scores are also associated with a significant and positive effect as baseline scores decrease. In other words, girls whose baseline agency scores were in the lowest quartile had the largest increases in agency, as compared to girls who scored in the highest quartile. This is consistent with our baseline-endline group comparisons (Figure 1) which suggested that medium-high dosage participants (who also started with the highest average agency) had the lowest gains.

## MULTIVARIATE EFFECTS: MEDIUM DOSAGE AND LONGER DURATION YIELD BIGGEST IMPACT

We considered two country fixed effects models (model 2 and model 3) side by side in Table 5, that show the multivariate effects of explanatory variables when omitting and accounting for baseline agency score. Model 2 accounts for approximately 15% of the variation in girls' change in agency scores as indicated by the  $R^2$  statistic ( $R^2=.15$ ), while the model in column 3 represents approximately 45% of the variation in the dependent variable ( $R^2=.45$ ), suggesting that model 3 offers a more comprehensive explanation of variability in girls' agency outcomes contributed by the effects of specific variables.

Overall, our preferred model (model 3) showed that program duration, baseline agency, and dosage all affect gains in agency, although there were no significant effects of age. However, model 2 shows that compared to the highest age group (age 22–25 years), girls in lower age groups (those below the age of 18) are associated with higher increases in agency scores than those above the age of 18. When baseline scores are added in model 3, the effects of age on agency gains disappears, suggesting that a girls' baseline agency score is a more important predictor of agency gain than a certain period of adolescence or young adulthood.

Similarly, examining the effects of program dosage on girls' agency outcomes between model 2 and model 3, the significant positive effects of low dosage programs compared to the high dosage category become null once baseline agency score is added to the model. Importantly, medium dosage programs continue to have significant ( $p=0.001$ ) and substantial (.295 SD) positive impact on girls' agency scores above and beyond the effect of baseline agency, age, and program duration, when compared to the high dosage category. There is evidence that while the effect of program duration decreases after accounting for baseline agency, each month increase in program duration significantly increases the gains in agency scores. Finally, the effect size of girls' baseline agency scores remains virtually identical in univariate (column 1) and multivariate models, showing substantial significant effects on girls' agency the lower the starting score.

## DISCUSSION

Our findings have important implications for program design and participant recruitment. In our study, the largest gains in agency were associated with girls who scored in the lowest quartile of agency at baseline. This indicates that programs will have a bigger impact if they target the most vulnerable populations, or those populations most likely to have low agency. While such a conclusion might be intuitive for practitioners, NGOs have long been criticized for targeting participants based on ease of access or other aspects of expediency, rather than based solely on need or merit. The effects of baseline agency score also suggest something important for program practice about the malleability of agency outcomes and the potential for change amongst those lowest scoring participants. To provide more detailed guidance on targeting low-agency populations, more research is needed to understand how agency scores vary with other characteristics of vulnerability such as socioeconomic status, minority ethnic/linguistic group, etc.

Second, the diversity of the 13 participating programs' length and structure allowed for a unique opportunity to investigate the relationship between program intensity (dosage) and duration on girls' agency gains. Most notable from our analysis is the implication that higher dosage is not necessarily associated with higher impact. In fact, there seems to be a 'sweet spot' of program dosage where sustained weekly engagement for short time periods, with more focused life skills content—as represented by medium dosage programs—has a higher effect on girls' agency gains than daily intensive engagement.

Third, the lagging trend line in environmental beliefs (Figure 2) both underscores the difficulty of improving girls' beliefs about the world around them and the potential impact of success. In our measure of agency, the concept of environmental beliefs is characterized by a girl's beliefs about gender norms and gender roles in her community and future, and beliefs about how rigidly her opportunities will be structured by these norms. Gender norms are understandably some of the most entrenched social constraints that girls face in building and exercising their agentic capacity. Our study indicates that more research is needed to ascertain what programs might be most effective at improving girls' views on gender and gendered norms, but that this might be a key area for greater impact on girls' agency.

Finally, our analysis finds that longer duration of programs is associated with higher agency gains, and that these effects were above and beyond the effects of program dosage and baseline agency, meaning that length of program has an impact on girls regardless of where they start. Each additional month of programming above and beyond three months was associated with a .028 standard deviation (sd) increase in agency gains; this translates to .25 sd at 9 months and .335 sd at 12 months—which rivals the effect of program dosage. For practitioners targeting agency outcomes, this offers practical guidance that indicates that, much like the teaching and learning of academic subjects, improving girls' agency is a sustained under-taking related to the building of complex skills and beliefs that take time and attention.

## LIMITATIONS

Our findings should be viewed in light of a few limitations. First, although our design is longitudinal, it is non-experimental, meaning that causal inferences cannot be made from the associations presented in our analysis. Relatedly, our data only captures a 12-month period, and thus we are unable to draw conclusions about programs that are longer than 12 months. It could be that programs which are longer than 12 months continue to improve agency scores at a similar rate, or there could be threshold effects. Programs which are actually multiple years in duration but which were assessed at the 12 month mark in our study could actually achieve higher impact if carried through to the end. At least eight programs in the study operate for two to four years.

Second, although the diversity of programs in our study allowed us to investigate questions related to program structure, this diversity also was a limiting factor in our ability to draw more specific conclusions about program characteristics that impact agency. Dosage categories are still quite broad, and represent a wide range of time-engagement. Our study population was also not evenly distributed across dosage categories—with only one program identified as ‘low dosage.’ Similarly, the study population was not evenly distributed across countries, with only 15 observations from Uganda, and over half the sample coming from Rwanda.

Finally, our assessment of agency is limited by the instrument which we used to measure agency gains. As mentioned, the strengths of the AMPLIFY Agency Measure include that it was locally developed and adapted for East Africa by practitioners; however, it is a self-reported, Likert scale survey and as such likely suffers from acquiescence bias, social desirability bias, and other biases related to these two characteristics of measurement instruments.

## CONCLUSION AND RECOMMENDATIONS: GUIDANCE FOR PRACTITIONERS

Despite these limitations, we believe our study offers some important guidance for practitioners and policy makers designing or investing in programs that impact girls’ agency. For maximum impact on agency, our study suggest that programs should:

- Target the most vulnerable (those likely to have the lowest agency);
- Engage in regular, sustained, focused, and medium-touch program delivery; and
- Consider program durations longer than three months. (Our analysis indicated that every additional month of programming after three months yielded increases in agency gains).
- Conduct more research that evaluates the effect of programs on girls’ beliefs about gender and gender norms, as this appears to be the slowest changing dimension of agency in our study.

We also note that these recommendations are a starting point for identifying best practices for supporting girls’ agency and we recommend further investigation into the specific practices and structures which support agency in the short and long-term. Our analysis has highlighted several specific areas of need in this regard, including studies which are longer than 12 months in duration and consider the contextual factors and processes which are responsible for affecting agency gains.

In conclusion, this evaluation is an example of how to document 'localized' impact. Life Skills programs are incredibly diverse in scope and content, and we argue that this diversity is likely important for serving the needs of unique communities of girls living in diverse settings with diverse opportunities and constraints in their lives. Rather than pass over community-driven organizations which are often best positioned to serve these young women, our methods and evaluation processes should include these important actors in all of evidence, practice, and policy making. As agency is a complex and contextualized construct often influenced by the evolving opportunities and constraints, so too are effective approaches to addressing girls' agency. Overall, we argue that CDOs are as well positioned as transnational actors to drive progress towards global development agendas such as gender equity.

## REFERENCES

- Brush, K. E., Jones, S. M., Bailey, R., Nelson, B., Raisch, N., & Meland, E. (2022). Social and Emotional Learning: From Conceptualization to Practical Application in a Global Context. In *Life Skills Education for Youth* (pp. 43–71). Springer.
- Federici, F. M., Gerber, B. J., O'Brien, S., & Cadwell, P. (2019). *The International Humanitarian Sector and Language Translation in Crisis Situations. Assessment of Current Practices and Future Needs*. London; Dublin; Phoenix, AZ: INTERACT The International Network on Crisis Translation.
- Goldman, M. J., & Little, J. S. (2015). Innovative grassroots NGOs and the complex processes of women's empowerment: An empirical investigation from Northern Tanzania. *World Development*, 66, 762–777. <https://www.sciencedirect.com/science/article/abs/pii/S0305750X14002708>
- Haberland, N. (2015). The Case for Addressing Gender & Power. *International Perspectives on Sexual and Reproductive Health*, 41(1), 31–42. <https://doi.org/10.1363/4103115>
- Hickey, S., & Mohan, G. (2005). Relocating Participation within a Radical Politics of Development. *Development and Change*, 36(2), 237–262. <https://doi.org/10.1111/j.0012-155X.2005.00410.x>
- International Labour Organization. (2020). *ILO Monitor: COVID-19 and the World of Work*. Updated estimates and analysis. Int Labour Organ.
- Konte, M., & Tirivayi, N. (Eds.). (2019). *Women and Sustainable Human Development: Empowering Women in Africa*. Springer.
- Lohani, M., & Aburaida, L. (2017). Women empowerment: A key to sustainable development. *The Social ION*, 6(2), 26–29.
- Mugo, J. (2020). Welcome ALiVE–RELI's Firstborn Initiative (Assessment of Life Skills and Values in East Africa). Regional Education Learning Initiative. <https://reliafrica.org/welcome-alive-reli-s-firstborn-initiative/>
- Nasheeda, A., Abdullah, H. B., Krauss, S. E., & Ahmed, N. B. (2019). A narrative systematic review of life skills education: Effectiveness, research gaps and priorities. *International Journal of Adolescence and Youth*, 24(3), 362–379.
- Oldekop, J. A., Horner, R., Hulme, D., Adhikari, R., Agarwal, B., Alford, M., ... & Zhang, Y. F. (2020). COVID-19 and the case for global development. *World development*, 134, 105044.
- Oulo B., Sidle A.A., Butler M. (2021). *Stemming Pandemic-Related Losses in Girls' Education- Promising Practices from the AMPLIFY Collective 2021*.
- Room to Read (2022). "Building the Adolescent Life Skills Assessment for Girls (ALSA for Girls): Synthesis Report on Lessons Learned." <https://roomtoread.org/alsa-synthesis-report>
- Seyfang, G., & Smith, A. (2007). Grassroots innovations for sustainable development: Towards a new research and policy agenda. *Environmental politics*, 16(4), 584–603.

## REFERENCES

- Sidle, A.A. & Oulo, B. (Forthcoming). Assessment of a Practitioner-Derived Framework for Measuring Girls' Agency in East Africa. *Comparative Education Review*.
- Sidle, A. A. (2019). Action on Agency: A Theoretical Framework for Defining and Operationalizing Agency in Girls' Life Skills Programs. *Gendered Perspectives in Development Working Paper Series, Paper #313*(University of Michigan).
- Sidle, A. A., Ahigika, D., Aszed, W., Bashuna, S., Belle, S., Bideri, D., Bohart, E., Butler, M., Fortunate, K. H., Fowler, Z., Kamau, L., Littman, J., Makumayenzi, V., Mariki, E., Mirisho, A., Musumba, P., Omare, M., Omondi, D., Oulo, B., Oyugi, D., Sarakikya, E. & Wolf, K. (2019). Measuring Girls' Agency in East Africa—Co-Creating Contextually Specific Tools for Evaluation (AMPLIFY White Paper). AMPLIFY Girls.
- Sidle, A.A. & Oulo, B. (Forthcoming) "Assessment of a Practitioner-Derived Measure of Adolescent Girls' Agency in East Africa."
- Trochim, W., & Kane, M. (2005). Concept Mapping: An introduction to structured conceptualization in health care. *International Journal for Quality in Health Care*, 17(3), 187–191.

## APPENDIX

**TABLE 2**

PROGRAM NAME	ORGANIZATION	TARGET POPULATION	DESCRIPTION
Mentorship Program	The Action Foundation of Kenya (TAF)	Girls living with physical disabilities, residing in high-density urban settlements in Nairobi.	A three-month Life skills and sexual and reproductive health and rights (SRHR) education.
Kisa Scholars	Girls Livelihood and Mentoring Initiative (GLAMI)	Low-income girls attending A-level secondary school who previously attended government O-Level schools.	A two-year leadership, empowerment, and social change training program for A-level girls.
Adolescent Girls Program	Girl Up Initiative of Uganda (GUIU)	Low-income urban girls attending government secondary schools.	A one-year school-based girl empowerment program covering life skills and SRHR content.
Teen Mothers Program	Komera	Rural girls and teen mothers, who come largely from subsistence farming backgrounds.	A three-year curriculum in Life skills, family planning, nutrition, wellness and SRH.
Rapid Response	Streets Ahead Children's Center Association (SACCA)	Young women who were current or formerly street children.	A three-month, community-based public awareness program focused on sanitation, COVID-19 Hygiene and SRH.
Livelihood	SACCA	Children and young adults who were current or formerly street children.	A six-month vocational training courses in various trades including hair dressing, construction, masonry.
Independent Study Program	Jifundishe	Young adults who did not finish secondary school and are unable to return to formal school settings.	A two-year academic program to teach secondary curriculum and support students in leadership and life skills.
Codehive	AkiraChix	Low income, vulnerable girls and young women who do not have the opportunity attend university.	A one-year training program in tech leadership, coding, professional development, and agency.
Scholarship Program	The Girls Foundation of Tanzania (TGFT)	Low-income, vulnerable, but highly motivated/talented adolescent girls, seeking to attend secondary school.	A six-year scholarship, leadership, and empowerment program for adolescent girls.

## APPENDIX

**TABLE 2 CONTINUED**

PROGRAM NAME	ORGANIZATION	TARGET POPULATION	DESCRIPTION
WISER Houses of Wisdom	WISER Girls' School	Rural and low-income girls living in the community of Mhuru Bay, Kenya.	A four-year boarding school and holistic empowerment, life skills, and SRH curriculum.
Sega	Sega School	Low-income but academically talented girls coming from throughout Tanzania.	A three-year boarding school and holistic leadership, life skills, and SRH curriculum.
Smart Girl Program	Riley Orton Foundation	Low-income girls living in urban settlements outside Kisumu.	A two-year school-based SRH, GBV, and life skills program.
Masenge Mba Hafi	Learn Work Develop (LWD)	Adolescent girls living in the rural village in Kayonza District	A two-year holistic life skills program based on three main pillars: family promotion, anti SGBV, and child protection.