

Empowering Adolescent Girls in the Sahel: Evidence from a Multi-Country RCT of the Sahel Women Empowerment and Demographic Dividend Project

Burkina Faso, Chad, Cote d'Ivoire, Mali, Mauritania, Niger

Pre-Analysis Plan

Olivia Bertelli, Claire Boxho, Elise Huillery, Harounan Kazianga, Estelle Koussoubé,
Bastien Michel, Markus Olapade, Léa Rouanet, Mathilde Sage

December 2021

1. Introduction

1.1 Overview and purpose of the study

The Sahel Women Empowerment and Demographic Dividend (SWEDD, P15008) is a regional project aiming to accelerate the demographic transition by addressing both supply- and demand-side constraints to family planning and reproductive and sexual health. To achieve its objective, the project targets adolescent girls and young women mainly between the ages of 10 and 19, and who are vulnerable to early marriage, teenage pregnancy, and early school drop-out. The project is currently implemented by the governments of Benin, Burkina Faso, Cameroon, Chad, Cote D'Ivoire, Guinea, Mali, Mauritania, and Niger.

This pre-analysis plan describes a multi-country impact evaluation. The impact evaluation framework focuses on six of these nine countries (i.e. all but Benin, Cameroon and Guinea) that were part of the first phase of the SWEDD (2015-2020) and where country-level impact evaluations are ongoing. In these countries, nineteen sub-projects have been approved by the SWEDD Regional Steering Committee. All the projects fall into one or more of three windows of eligible interventions. The first window, empowering girls, includes life skills and sexual and reproductive health knowledge projects that build adolescent girls' capacity to lead healthy and productive lives. The second window focuses on improving economic opportunities through support for income-generating activities. The third group of projects falls under the window of keeping girls in school and includes projects to improve girls' enrollment and retention in school.

The multi-country impact evaluation study focuses on evaluating interventions in the two first windows and will generate evidence related to the following research questions: What is the impact of community-based *Safe spaces* on adolescent girls' marriage outcomes and sexual and reproductive health knowledge and behaviors? What about the impact on empowerment indicators such as aspirations and self-esteem? What is the impact of combining community-based *Safe spaces* and livelihood support interventions on involvement in income generating

activities, savings, and earnings? What is the impact of combining *Safe spaces* and livelihood support interventions on health, empowerment, and economic outcomes?

1.2 Description of evaluated interventions

Two interventions are evaluated in the scope of the multi-country impact evaluations in six countries: Burkina Faso, Chad, Côte d'Ivoire, Mali, Mauritania, and Niger. The first intervention includes the participation into *Safe spaces* for adolescent girls where they can gain life skills and receive sexual and reproductive health knowledge. In these *Safe spaces*, girls interact with their peers and mentors, participate in recreational activities and develop life and health skills during a critical period of their development. These spaces act as platforms to deliver services to girls, such as sexual and reproductive (SRH) knowledge, life skills training, literacy and non-formal education, in an informal and accessible location. In Chad, Côte d'Ivoire and Mauritania, *Safe spaces* were implemented both in schools and in communities. In Burkina Faso, Mali and Niger, *Safe spaces* were only implemented in communities. The multi-country analysis presented in this pre-analysis plan will focus on the impacts of *Safe spaces* implemented in local communities and targeting out-of-school girls. The impacts of *Safe spaces* implemented in school will be analyzed in a separate study.

A second intervention that aims to improve the economic empowerment of adolescent girls was evaluated in all countries with the exception of Niger. It includes vocational training and financial support (loan or grant) in all countries, plus entrepreneurship training in all countries but Chad. The purpose of this livelihood support intervention is to equip girls with new technical skills, ease the credit constraint they face and ultimately improve their economic opportunities. This second intervention targeted different age groups per country. Table 1 provides details on the intervention implemented in each country and on the age eligibility.

These two interventions were designed to create the conditions for girls to stay in school longer, marry later and begin having children later.

Table 1: Overview of the interventions evaluated by country

Project initiatives	Burkina Faso	Chad	Côte d'Ivoire	Mali	Mauritania	Niger
1. Life Skills and Sexual and Reproductive Health Knowledge						
Duration of interventions (# of intervention in months)	6	9	6	12	9	7
Age Eligibility	10-24	12-24	10-24	9-19	16-35	10-19
2. Improving Economic Opportunities						
1-Technical/Professional training	Yes	Yes	Yes	Yes	Yes*	No
2-Financial support (grants, loans)	Yes*	Yes	Yes	Yes	Yes*	No
3-Entrepreneurship training	Yes	No	Yes	Yes	Yes*	No
Age Eligibility	15-24	15-24	16-24	15-24	16-35	-

Notes: *Due to delays in the implementation of these components, these would be evaluated using endline data.

The different interventions were implemented between January 2019 and January 2020. The implementation agencies by country are as follows:

- **Burkina Faso:** Ministry of Women, National Solidarity, Family and Humanitarian Action; and Ministry of Health
- **Chad:** Ministry of National Education, Ministry of Women, National Solidarity, Family and Humanitarian Action; and Ministry of Health, Ministry of Plan and Prospective Ministry of youth and sport.
- **Côte d'Ivoire:** Ministry of National Education, Technical Education and Vocational Training; Ministry of Women, Family, and Social Affairs; Ministry of Health and Public Hygiene.
- **Mali:** Ministry of Territory and Population
- **Mauritania:** Ministry of Social Affairs, children and Family, Ministry of National Education and Ministry of sport and youth.
- **Niger:** Ministry of population, Ministry of Women and child Protection Ministry of Primary Education and Alphabetization, Ministry of National Languages and Civic Education, Ministry of Secondary Education.

1.3 Impact evaluation design

The regions in which the impact evaluation is implemented are shown as dark areas in Figure 1 below.

Figure 1: Regions included in the SWEDD multi-country impact evaluation

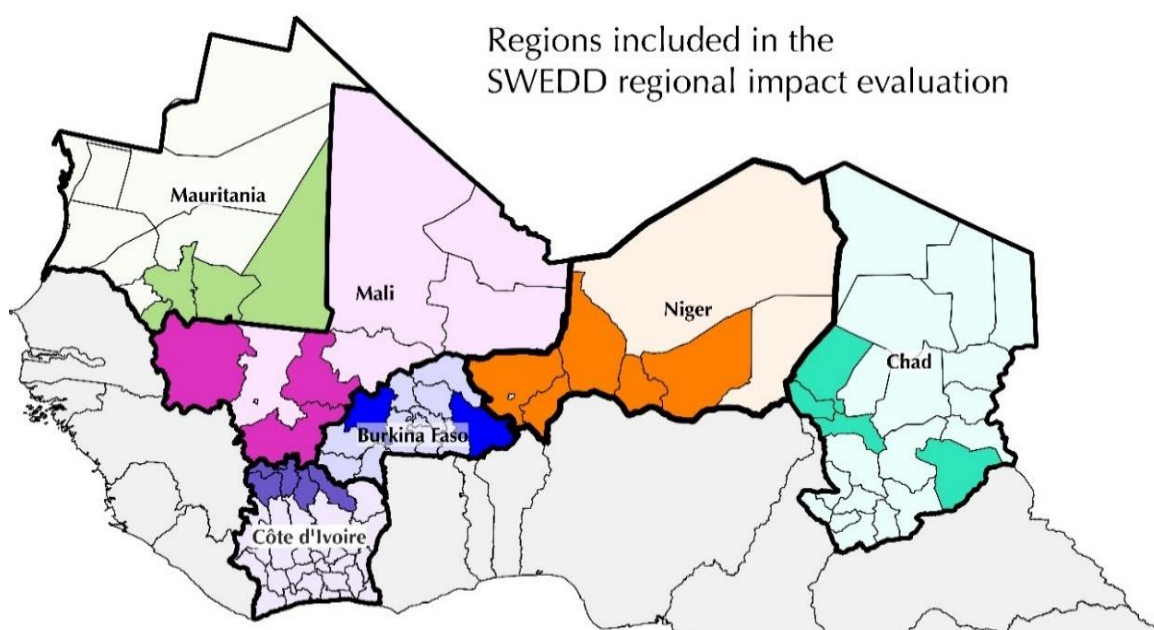


Table 2 describes the sample design for the multi-country impact evaluation and provides information on the units of randomization, the stratification variables and the sample composition by marital status and age of girls.

In each country, clusters were determined using administrative units (rural or urban localities, or villages). These clusters were randomized into control or treatment groups. The randomization was stratified by country-specific variables, including some upper-level administrative units (regions, provinces, etc.), see Table 2 below.

The baseline sample is constituted of out-of-school adolescent girls eligible to the interventions. Priority was given to adolescent girls that were at risk (orphans, single mothers, pregnant single girls, etc.). Additionally, in Cote d'Ivoire and Mauritania, only girls from vulnerable households (as determined by a poverty score) were eligible to the program. The baseline sample size in each country has been determined using power calculations with age of marriage and childbearing as the outcomes of interest.¹

Table 2: Baseline sample by country for the multi-country Impact Evaluation

	Burkina Faso	Côte d'Ivoire	Mali	Mauritania	Niger
Units of randomization	Villages	Localities (rural/urban)	Villages	Localities (rural/urban)	Villages
Number of units	173	168	151	74	198
Stratification variables	11 provinces	% of married girls before 18, Number of girls sampled from the census	6 administrative cercles (second level administrative unit) ; 38 communes	Population of major urban centers	5 regions
Age of girls in IE sample	10-24	12-24	12-24	15-29	10-19
Sample size	8264	2406	3106	2884	3363
Sample size of girls below 15	1261	321	353	0	1312
Sample size of married girls	5793	1456	1872	1670	1538
Other eligibility criteria	None	Poverty index	None	Poverty index	None

Notes: Information on the Impact Evaluation in Chad will be included in a later version.

The multi-country impact evaluation builds on the individual, country-level impact evaluations and the random assignment of clusters into two treatment groups and a control group:

- *Treatment 1: Safe spaces* where adolescent girls receive (SRH) knowledge and life skills training. We will refer to this treatment as *Safe spaces* (T1)
- *Treatment 2: Safe spaces* (T1) plus some livelihood support intervention composed of vocational training, financial support and entrepreneurship training. We will refer to this treatment as *Safe spaces plus IGA* (Income Generation Activities) (T2)
- *Control group*: No interventions.

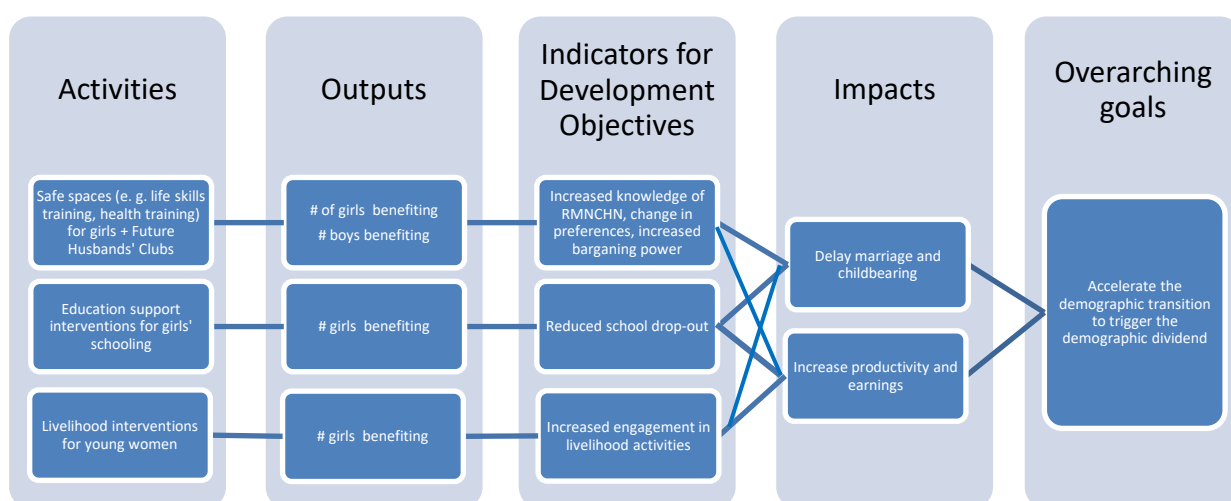
¹ The sample sizes allow us to detect effects like those reported by previous studies (e.g. Bandiera et al., 2017; Adoho et al., 2014), i.e. 0.3SD-0.35SD on our main outcomes of interest.

1.4 Theory of change

The design of the SWEDD project is based on the hypothesis that increasing adolescent girls and women social and economic empowerment is essential to: (i) reduce fertility rates, (ii) improve health; (iii) and increase human capital to lower dependency ratios and create a productive adult workforce. Here, “social empowerment” is defined as participation in community-level girls’ programs, with an implicit theory of change that such participation leads to social empowerment. These programs will enable economic empowerment. “Economic empowerment” is defined here as having both the resources to advance economically and the power to make and act on economic decisions. Economic empowerment starts with fair and equal access to markets and institutions, but women also need agency and control over resources to ensure they benefit from economic activities.

Furthermore, the anticipated project’s impacts (delayed marriage and childbearing; increased engagement in activities and earnings) are mutually reinforcing in that marriage/childbearing delays lead to productivity gains by allowing girls the time to finish school and start a livelihood. This facilitates increased engagement in economic activities and allows for higher earnings. As a result, there is an increase in the opportunity cost of having children and hence delays childbearing. Thus, providing a combination of interventions (*Safe spaces* + IGA) is expected to yield more impact than simply providing one intervention. The figure below describes the causal chain between each intervention and its hypothesized impacts.

Figure 2: Results chain for the SWEDD project



2. Methodology

2.1 Specification

(a) Impact of *Safe spaces* and accompanying economic empowerment measures

To estimate the impact of the *Safe spaces*’ intervention and the combination of *Safe spaces* and IGA on our outcomes of interest using all six SWEDD countries, we will estimate different ANCOVA specifications. For outcomes without baseline values, a simple OLS estimation will

be used, considering the same specifications below without the baseline value of these outcomes (y_{i0}). Reported results will include the pooled estimates and the country-specific estimates.

As explained in Section 1.2, the intervention “Improving economic Intervention” focusing on IGA targeted different age ranges per country. Therefore, in order to best estimate the Intention-to-Treat impact of the group T2 *Safe spaces plus IGA*, we will run all specifications described in this section on the entire sample of adolescent girls and on the restricted sample of girls eligible to the IGA intervention. Balance test will verify that the restricted sample is balanced between treatment arms on the main characteristics for the countries where the IGA intervention took place.

Specifications 1 & 2: Short/mid-term impact of *Safe spaces* and *Safe spaces plus IGA*.

$$y_{i1} = \alpha_1 + \beta_1 \text{SafeSpace}_{it=1} + \gamma_1 \text{SafeSpace plus IGA}_{it=1} + \delta_1 y_{i0} + \lambda_1 \pi_i + \varepsilon_{i1} \quad (1)$$

$$y_{i2} = \alpha_2 + \beta_2 \text{SafeSpace}_{it=2} + \gamma_2 \text{SafeSpace plus IGA}_{it=2} + \delta_2 y_{i0} + \lambda_2 \pi_i + \varepsilon_{i2} \quad (2)$$

Where y_{i1} is an outcome variable for individual i at midline ($t=1$), y_{i2} is an outcome variable for individual i at endline ($t=2$), and y_{i0} is the baseline value of the outcome variable considered. SafeSpace_i is the indicator of assignment to treatment arm *Safe spaces* only, i.e. equal to 1 if the individual i lives in a community assigned to receive the *Safe spaces* intervention. $\text{SafeSpace plus IGA}_i$ is the indicator of assignment to the treatment arm “*Safe spaces plus IGA*”, equal to 1 if the individual i lives in a community assigned to receive *Safe spaces* and economic empowerment interventions. Lastly, π_i is the country-specific stratification variable that was used for randomization.

Here, the parameters β_1 and γ_1 will identify the treatment effect of the *Safe spaces* and *Safe spaces plus IGA* interventions respectively on the relevant outcome of interest at midline, while β_2 and γ_2 will identify these same treatment effects at endline. Standard errors will be clustered at the randomization unit level.

Hypothesis Testing:

The main hypotheses we intend to test based on specification (1) and (2) are as follows:

Null hypothesis 1: Neither intervention has an effect on the outcome of interest relative to the control group at midline: $\beta_1 = 0, \gamma_1 = 0$.

Null hypothesis 2: Neither intervention has an effect on the outcome of interest relative to the control group at endline: $\beta_2 = 0, \gamma_2 = 0$.

Null hypothesis 3: The economic empowerment accompanying measures (IGA), offered on top of *Safe spaces*, have no additional effect on the outcome of interest relative to the control group at endline: $\gamma_t = 0$, with $t = \{1, 2\}$ respectively for midline and endline.

In addition, we will estimate the following equation to test the sustainability of the impact and to see if the impact persists in the longer run:

Specification 3: Sustainability of the impact across time

$$y_{it} = \beta_0 + \beta_1 \text{SafeSpace}_i + \beta_2 \text{SafeSpace plus IGA}_i + \beta_3 T_2 + \beta_4 \text{SafeSpace}_i \times T_2 + \beta_5 \text{SafeSpace plus IGA}_i \times T_2 + \beta_6 y_{i0} + \beta_7 \pi_i + \varepsilon_{it} \quad (3)$$

Where y_{it} is an outcome for individual i in period $t = \{1, 2\}$ and y_{i0} is the baseline value of the outcome indicator. T_2 is a time period indicator, equal to 1 if the observation is an endline observation, and ε_{it} is the error term. Here, the parameters β_4 and β_5 will identify the change between midline and endline in the treatment effect of *Safe spaces* and *Safe spaces plus IGA* respectively on the relevant outcome of interest. Standard errors will be clustered at the randomization unit level. For this specification, the data from midline and endline will be pooled.

Hypothesis Testing:

Null hypothesis 1: Neither intervention has a sustainable effect on the outcome of interest relative to the control group at endline: $\beta_4 = 0, \beta_5 = 0$.

(b) Heterogeneity of impact by age group and marital status

Existing studies of the impact of *Safe spaces* and other interventions targeting girls to improve sexual and reproductive health outcomes have emphasized the mediating effect of age, with younger girls benefiting more on average (e.g. Amin et al., 2016). One can also expect the participation to the interventions as well as the impacts on the outcomes of interest to vary by marital status. For instance, married girls may benefit less for the interventions as they may have less time to devote to these activities due to domestic responsibilities. On the other hand, because group-based activities have the potential to increase social capital (e.g. Roy et al., 2019), they may particularly benefit married girls who may have less opportunities to interact with their peers due to time constraints. In addition, the curriculum is framed differently for girls under 13 and girls above. Furthermore, recommendations were made to separate the clubs by age group and marital status for the success of the implementation. We will test for the heterogeneity of effects by age group and marital status, by estimating the following ANCOVA specification:

Specification 4: Heterogeneity in impacts by age group and marital status

$$y_{it} = \beta_0 + \beta_1 \text{SafeSpace}_i + \beta_2 \text{SafeSpace plus IGA}_i + \beta_3 X + \beta_4 \text{SafeSpace}_i \times X + \beta_5 \text{SafeSpace plus IGA}_i \times X + \beta_7 y_{i0} + \beta_8 \pi_i + \varepsilon_{it} \quad (4)$$

Where y_{it} is an outcome for individual i in period $t = \{1, 2\}$ and y_{i0} is the baseline value of the outcome indicator. X represents the baseline characteristics of interest, including age and

marital status (married/unmarried). ε_{it} is the error term. Here, the parameters β_4 and β_5 will identify the differences in treatment effects across marital status and across age groups. Standard errors will be clustered at the randomization unit level.

Hypothesis Testing:

Null hypothesis: Neither intervention has a heterogeneous (differential) effect on the outcome of interest relative to the control group across age groups (or respectively marital status groups) at midline (the same test can be conducted at endline): $\beta_4 = 0, \beta_5 = 0$.

The heterogeneity analysis will not be run on the outcomes related the marital status of the adolescent girls. In addition, since there are some differences between baseline and midline questionnaires (for instance, new outcomes added to the midline survey), it will only be possible to estimate an ANCOVA for outcomes that have been collected both at baseline and at one of the two distinct post-intervention surveys (midline, endline).

2.2. Intent-To-Treat (ITT) and Treatment Effect on the Treated (TOT)

From the different specifications above, estimations will be first based on ITT that is, using initial random assignment to treatment arms. The ITT measures the effect of offering the treatments. TOT measures the impact based on those who actually participate in the program. If there were perfect compliance in each treatment arm, then the ITT estimation would yield to same results at the TOT estimation. Since we do not have perfect compliance in the project (some girls assigned to treatment did not participate to the program), we plan to estimate the TOT to measure the impact on participants. We do this by instrumenting actual treatment status (participated versus not participated) by the random assignment to treatment arms dummy. The specification to be used is similar to equations 1 and 2 and will be estimated in the framework of a two-stage least squares method.

2.3. Data

During the life of the impact evaluation, there will be three data collection rounds: (i) baseline survey; (ii) midline survey and (iii) endline survey (Figure 3).

(i) *Baseline survey*: the survey is based on two main questionnaires harmonized across countries: household and adolescent questionnaires. The household questionnaire is designed to collect information on households' living conditions and members². The adolescent questionnaire aims to collect information related to the aspirations and the social behavior of the girl, as well as information on gender, relationships, sexual and reproductive health, education, marriage, economic activities and access to finance.

(ii) *Midline survey*: The purpose of the midline survey is to measure short-term impacts and to assess the performance of the intervention. The survey results will be used to inform

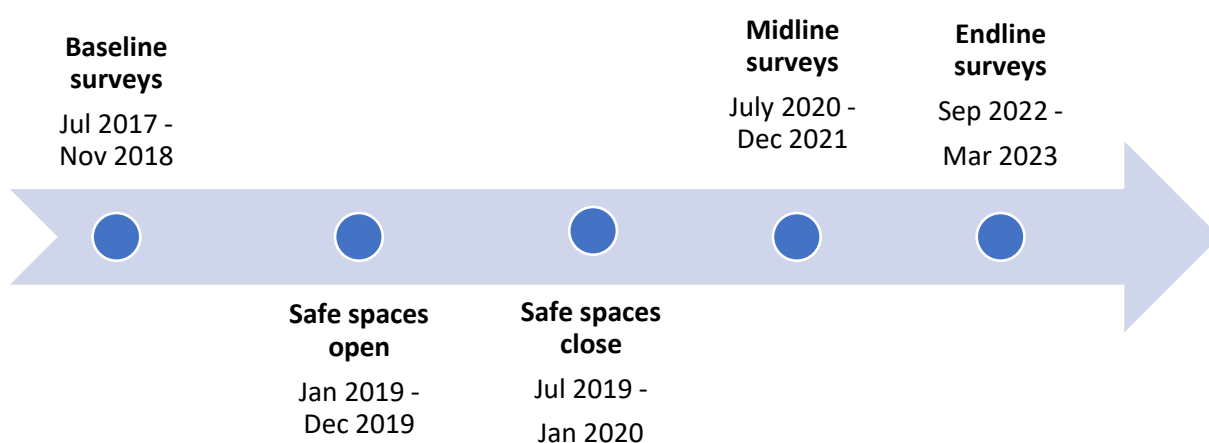
² Note that follow-up surveys will collect data on marriage and fertility for household members under 24 years, as well as the information on gatekeeper(s) aspirations and norms to study within households' spillovers.

stakeholders about effectiveness of project interventions and help to refine project implementation strategies if needed. During this survey, the respondents of the baseline survey will be interviewed again using updated household and adolescent questionnaires.

(iii) *Endline survey*: This will be an additional post-intervention survey used to measure the longer-term impacts of the treatments.

In addition to these survey data, this study will rely on monitoring and evaluation data from the implementing partners. These data will provide valuable information on *Safe spaces'* openings, mentors' and girls' attendance, sessions' duration, content, etc. As much as possible, this administrative data will be matched to each adolescent survey data. The data collection timeline is illustrated on Figure 3.

Figure 3: Timeline of key proposed multi-country impact evaluation activities



2.4. Key outcomes of interest

In order to precisely pin down the different channels and mechanisms sketched in the theory of change, we classify the different outcomes of interest into primary outcomes, mechanisms outcomes and secondary outcomes (see Tables 3, 4 and 5).

2.4.1. Primary outcomes

Primary outcomes are the key outcomes on which treatment impact will be measured. We have grouped them into two main families: (i) childbearing and marriage, and (ii) IGA as depicted in Table 3. Indices construction is detailed below.

Table 3: Primary outcomes

Family of outcomes	Outcome	Sample
Childbearing and marriage	Adolescent ever pregnant [yes=1]	All
	Number of times adolescent has given birth to a child	All
	Age at first child is before 18 years old [yes=1]	18 years old+
	Adolescent ever married [yes=1]	All
	Adolescent's age at first marriage	Married

	Adolescent's age at first marriage is before 18 years old [yes=1]	18 years old+
IGA	Adolescent has been engaged in IGA (last 30 days)* [yes=1]	15+ and/or girls eligible to the IGA intervention
	Total hours worked (last 30 days)*	15+ and/or girls eligible to the IGA intervention
	Amount of money/in kind benefits received from (last 30 days)* (2017 USD PPP)	15+ and/or girls eligible to the IGA intervention
	Adolescent received income from IGA (last 30 days)* [yes=1]	15+ and/or girls eligible to the IGA intervention
	Income Generating Activities index	15+ and/or girls eligible to the IGA intervention

Notes: *We will also look at these outcomes on a reference period of the last 6 months.

The Income Generating Activities z-score is constructed based on the three outcomes in the IGA family (Table 3), following Kling *et al.* (2007) and Buchmann *et al.* (2018). Its construction proceeds as follows. First, each IGA outcome is normalized and “normed”, i.e. the sign is reversed for adverse outcomes if any (this ensures to interpret any increase in the index as higher value). A z-score is then calculated as a simple average of the normed values. Let y_k^* be the normalized value of the k^{th} outcome of the list K outcomes: $y_k^* = \frac{Y_k - \mu_k}{\sigma_k}$, with μ_k and σ_k , respectively the control group mean of the outcome, and its standard deviation. The z-score is given by: $Y = \frac{1}{K} \sum y_k^*$.

2.4.2. Mechanism outcomes

Mechanism outcomes give a sense of relevant channels conducive to the impacts to be measured on primary outcomes. They are grouped into seven main families: (i) knowledge and attitudes in Reproductive, Maternal, Neonatal and Child Health, and Nutrition (RMNCHN); (ii) women empowerment; (iii) gender-based violence; (iv) education; (v) aspirations on marriage and childbearing; (vi) aspirations on education and professional occupation; and (vii) socio-emotional skills. Table 4 provides details on these mechanism outcomes for each family. Indices construction is further explained below.

Table 4: Mechanism outcomes

Family of outcomes	Outcome	Sample
Knowledge and attitudes in RMNCHN	Adolescent knows at least one modern contraception method [yes=1]	All
	Adolescent used condom during the last sexual intercourse [yes=1]	15 years old+ and sexually active
	Adolescent uses any contraceptive method [yes=1]	15 years old+ and sexually active
	Adolescent uses any modern contraceptive method [yes=1]	15 years old+ and sexually active
	Adolescent uses any modern contraceptive method except condom [yes=1]	15 years old+ and sexually active

Women Empowerment	Decision making power index	All
	Decision making power index for couples	Married
Gender-based violence	Adolescent experienced any emotional violence (last 12 months) [yes=1]	15 years old+
	Adolescent experienced any physical violence (last 12 months) [yes=1] *†	15 years old+
	Adolescent experienced any sexual violence (last 12 months) [yes=1] *†	15 years old+
	Adolescent experienced any physical or emotional violence (last 12 months) [yes=1]* †	15 years old+
	Adolescent experienced any violence (physical, sexual or emotional) (last 12 months) [yes=1] *†	15 years old+
	Adolescent experienced spouse control (last 12 months) [yes=1]	15 years old+ and married
	Physical violence frequency and severity index*†	15 years old+
	Sexual violence frequency and severity index *†	15 years old+
	Adolescent finds domestic violence justified [yes=1]	15 years old+
Education	Adolescent's highest level of formal education attained	All
	Adolescent's education level at which is currently enrolled	Enrolled
	Adolescent is currently enrolled at school [yes=1]	All
	Number of months the adolescent was in school over the last three school years	All
Aspirations on marriage and childbearing	Age at which the adolescent would like to have her first child	All
	Age at which the adolescent wants to get married	All
	Adolescent's preferred number of children	All
Aspirations on education and professional occupation	Level of education adolescent wishes to attain	All
	Occupation adolescent wishes to exert	All
	Revenue adolescent wishes to make from desired occupation*	All
Socio-emotional skills	Self-esteem index	All
	General self-efficacy index	All
	Self-awareness index*	All
	Problem solving index*	All
	Perseverance index*	All
	Index on ability to influence others*	All
	Creativity index*	All
	Socio-emotional skills index*	All

Notes: Outcomes noted with * are not available at baseline. † Variables not collected in Mali.

We detail below the construction of indices for i) decision-making power, ii) gender-based violence and iii) socio-emotional skills.

Decision-making power

We account for adolescents' decision-making power by constructing two standardized decision-making indices. One is based on her personal life and the other on her couple's life if she is married. These indices are constructed by summing answer modalities "*strongly disagree*", "*disagree*", "*partly agree*", and "*strongly agree*" coded respectively 1, 2, 3 and 4 to various statements. The statements on the adolescent's personal-level decision-making are as follows: whether she may decide on i) the money she earns; ii) her studies; iii) her own work; iv) her own health care; v) her daily activities; and vi) her movements outside the house. The statements on couple-level decision-making are as follows: whether she may decide on i) her pregnancies; ii) household spending; iii) household major purchases; and iv) visits to her family.

Gender-based Violence (GBV)

The Gender-based Violence (GBV) questions are provided in Appendix A, Tables A1 to A5. Their answer modalities are dummies. Based on each set of dummies, we construct seven indicators: six of them are related to the experienced violence and one indicator reflects the adolescent's attitudes toward domestic violence.

The six indicators on experienced violence account respectively for emotional violence, physical violence, sexual violence, the spouse's control over the adolescent, whether the adolescent experienced any physical or sexual violence, and any violence. These indicators are dummies coded 1 if the adolescent responds "yes" to any of the items in the corresponding table in Appendix A and 0 if there is no occurrence of the considered violence. In case one of the items is missing (eg. the respondent refuses to answer) and all others are 0, the indicator is coded missing, while in case one item is missing and there is at least one item with an affirmative response (1"yes"), the indicator is coded 1 (WHO multi-country study, 2005).

In addition to these indices, we use the frequency in the last 12 months of physical and sexual violence to construct two indices of violence severity: physical violence severity and sexual violence severity. For each violence item in Appendix Tables A3 and A4, the frequency of occurrence is measured on a 4-point Likert scale (0 "Never", "Once", 2 "Sometimes", 3 "Often"). The severity indices are constructed as an average of the frequencies to the considered violence statements and standardized. In addition, we construct an alternative index as a sum of violence frequencies related to the two physical and three sexual violence items in Table A3 and A4 in Appendix A. This second measure will be used as robustness check of the impacts measured on experienced violence severity. In case all of the items pertaining to the construction of an indicator is missing for a given respondent (eg. the respondent refuses to answer), the severity index is coded to 0 for that respondent.

Finally, in order to account for violence within couple (Intimate Partner Violence-IPV), we will construct all the above violence indicators for adolescents who had a partner at baseline.

Socio-emotional skills (SES)

The socio-emotional skills (SES) statements are displayed in Appendix B, Tables B1 to B5. The answers modalities to each SES statement are “*strongly agree*”, “*partly agree*”, “*disagree*”, “*strongly disagree*” coded respectively 1, 2, 3 and 4. Based on these answers, the following indices are constructed: General Self-efficacy, Self-esteem and Creativity.

The General Self-efficacy index is constructed by summing the answers to 9 statements (ps1-ps5 in Appendix Table B1; ps17-ps19, in Appendix Table B2; and ps31 in Appendix Table B4), following Schwarzer and Jerusalem (2010) and Chen *et al.* (2001).

The Self-esteem index is constructed by summing the answers to statements as suggested by Rosenberg (2015). Since most of these statements are not available at midline, we rely on the following statements: i) I am an original thinker; ii) I know the skills I have and others do not; iii) People like to follow my ideas; iv) I am good at reading social situations to present myself well.

Creativity index is constructed by summing answers to the following statements: i) I am able to come up with new and different ideas; ii) I like to think about new ways of doing things; iii) I’m finding new ways to do things; and iv) I am an original thinker.

Four additional SES indices account for skills such as Self-awareness, Problem solving, Perseverance and Ability to influence others. Each of these indices are constructed by summing the answers to their respective underlying statements as described in Appendix B.

Finally, the socio-emotional skills index in Table 4 is a summary index constructed by summing the answers to all the socio-emotional skills statements, as described in Appendix B. Socio-emotional skills indices will be standardized, based on the same methodology described for standardized IGA index.

2.4.3. Secondary outcomes

Secondary outcomes are grouped in the following families: (i) RMNCHN, (ii) IGA, (iii) finance, (iv) spillovers on the adolescent girls’ parents, (v) spillovers on the children and young adults living in the adolescent’s baseline household, (vi) gender attitudes, (vii) mental wellbeing and life satisfaction, and (viii) nutrition. Table 5 explains each of these outcomes. We provide further details on the construction of indices below.

Table 5: Secondary outcomes

Family of outcomes	Outcome	Sample
Knowledge and attitude in RMNCHN	Adolescent has ever heard of HIV/AIDS [yes=1]	All
	Adolescent agrees that using a condom during sex intercourse can reduce HIV/AIDS risk [yes=1]	All

	Index of opinion on maternal health care	All
	Index of attitude towards maternal health care	Adolescents who gave birth at least once
IGA	Adolescent worked in a non-agricultural IGA sector (last 6 months) [yes=1]	All
	Adolescent total number of IGA activities	All
	Number of IGA the adolescent girl initiated (last 6 months)	All
	Adolescent's standardized time spent on domestic chores	All
Finance	Adolescent borrowed money in the last 12 months [yes=1]	All
	Amount of money borrowed in the last 12 months (PPP)	All
	Adolescent saved money in the last 12 months [yes=1]	All
	Amount of money saved in the last 12 months (PPP)	All
Spillovers on the parents[†]	Minimum age at which household head wishes the girls living in the adolescent's baseline household get married	All
	Gap between the minimum age at which the household head wishes the adolescent girls and adolescent boys who live in the baseline household to marry	All
	Proportion of girls living in the adolescent's baseline household for which the household head wishes a high-status job (army, government or private sector) as future employment	All
	Gap between the proportion of girls and proportion boys who live in the baseline household for which the household head wishes a high-status job (army, government, or private sector) as future employment	All
	Gender attitude index of the household head (1-44)	All
	Gender attitude index of the adolescent's mother (1-44)	All
Spillovers on the children and young adults living in the adolescent's baseline household[†]	Median number of children that the women aged 8-24 years old and living in the adolescent's baseline household gave birth to since 2019 (excludes the sampled adolescent)	All
	Minimum age at first marriage of girls who were unmarried at baseline in the household (excludes the sampled adolescent)	All
	Gender equality index 1	Ever married
	Gender equality index 2	Ever married

Adolescent's gender attitudes	Ideal age for a woman to marry	All
	Ideal age for a man to marry	All
	Gap between the ideal age to marry for women and men	All
	Ideal level of education for her (future) daughter(s)	All
	Gap between the ideal level of education for her (future) daughter(s) and her (future) son(s)	All
	Adolescent thinks it is desirable for her (future) daughter(s) to have a professional occupation once adult [yes=1]*	All
	Difference between whether the adolescents thinks desirable for her (future) daughter(s) to have a professional occupation once adult compare to her (future) son(s)*	All
Mental well-being and life satisfaction	Mental health index*	All
	Overall adolescent feels happy or very happy [yes=1]	All
Nutrition	Food diversity: number of different foods adolescent consumes*	All
	Household food security index	All

Notes: Variables noted with * are not available at baseline. † Variables not collected in Niger and partially collected on a sub-sample in Cote d'Ivoire.

On RMNCHN indicators, we construct two indices to account for both opinion and attitude towards maternal health care. The opinion index is constructed by summing 4 dummies (1 or 0) indicating whether the adolescent finds important (i) to go to prenatal care visit when pregnant, (ii) to deliver in a hospital, (iii) to go to postnatal care visit after giving birth, and (iv) considers family planning important. The index on attitudes towards maternal care is constructed by summing three dummies on whether: (i) the adolescent went to prenatal care visit when pregnant of her last child, (ii) the adolescent gave birth to her last child at the hospital, and (iii) the adolescent went to postnatal care visit when pregnant of her last child.

The index on the adolescent's time spent on domestic chores is constructed by summing the time spent in collecting wood, fetching water, cooking, and cleaning. We then compute a z-score based on the sum obtained.

The indices on attitudes towards gender of the household head and of the adolescent's mother are constructed by summing answers to 11 statements listed in Appendix Table C. The answer modalities are “*strongly disagree*”, “*disagree*”, “*partly agree*”, and “*strongly agree*” respectively coded 1, 2, 3 and 4. Items are reversed if needed to indicate gender equitable attitudes.

Similarly, we construct two indices reflecting the adolescent's attitudes towards gender equality, based on the 7 statements listed in Appendix D. Having two indices will allow to test the robustness of the results to the different outcome constructions. The answer modalities to the statements are: 1 “male”, 2 “female”, or 3 “both, collectively”. A first gender equality index (IXgenequality_1) is constructed by summing dummies equal to 1 whenever the adolescent responds 3 “both, collectively” to any of the 7 statements. This index value increases by 1 whenever the answer to a statement is “both, collectively”, such that the

minimum value of the index is 0 and its maximum is 7. The second gender equality index accounts for the extent to which spouses are conforming to traditional gender role beliefs in the household. This index is constructed based on two sets of dummies. A first set of 3 dummies define traditional male roles in the household and are respectively coded 0 if the adolescent responds “male” to statements 1, 2 or 7, and 1 otherwise. The second set of 4 dummies defines traditional women roles in the household. These dummies are respectively coded 0 if the adolescent responds “female” to statement 3, 4, 5 or 6, and coded 1 otherwise. The index is then constructed by summing the 7 dummies.

A mental health index is constructed by summing answers to statements listed in Appendix E. The possible answers are “*not at all*”, “*not more than usual*”, “*a bit more than usual*” and “*more than usual*” coded respectively 1, 2, 3 and 4. Items are reversed if needed to indicate positive feelings.

Finally, we will also use nutrition indicators and construct two indices: a food security index and a food diversity index. The food diversity index is constructed by summing dummies indicating whether different foods (vegetables, main staples, milk and dairy, meat fish, fruits, etc.) are consumed by the adolescent the previous day. The food security index is the sum of two dummies: (i) in the last 6 months, you faced a situation where you didn’t have enough food to feed the household [yes=1]; and (ii) in the last 7 days, a member of your household skipped a meal because your household did not have enough food [yes=1].

All indices will be standardized based on the methodology discussed for IGA index standardization.

3. Other methodological considerations

3.1 Multiple hypothesis testing

The main objective is to control for Type I error rate (false positive or false discovery) by adjusting p-values. In the framework of this study, we are going to account for false discovery rate by correcting the p-value within the families of outcomes (as described in Tables 3-5 above) and report the corresponding “sharpened q-values” using Anderson’s (2008) routine.

P-values will be adjusted for the number of distinct outcomes of the same family on which impacts are estimated separately. However, in order to improve power and limit false discoveries we will rely as much as possible on summary indices.

3.2 Attrition

Despite significant effort invested to track the maximum of adolescent girls in follow-up surveys, we expect attrition to happen. After follow-up surveys, we will test whether there is a differential attrition across the different treatment arms. The test will be based on a regression of attrition on the treatment dummies as follows:

$$Attrition_{i1} = \alpha_1 + \beta_1 SafeSpace_{it=1} + \gamma_1 SafeSpace\ plus\ IGA_{it=1} + \varepsilon_{i1}$$

Where *Attrition* is a dummy equal to 1 if the adolescent girl was not found.

Null hypothesis tested: There is no statistical differential attrition across treatment arms: $\beta_1 = 0$; $\gamma_1 = 0$.

If attrition is not problematic, that is, non-rejection of the previous null hypothesis at 5 % significance level, then the different estimations considered in the framework of this study will be robust to attrition. In case attrition is problematic and in particular correlated with the baseline values of our primary outcomes, we will reweight the sample using Lee's (2009) bounding method and inverse probability weighting. This will allow us to bracket out treatment effects.

3.3 Outliers

Outliers often arise in the distribution of monetary values such as revenue and savings. Our approach will be to winsorize at the appropriate percentile based on the distribution of the outcomes. We will also rely on an inverse hyperbolic sine transformation (IHST)³ to approximate a log specification without dropping the zeroes, as is common in the literature with wealth and income data (see Callen et al. 2019, Dupas, 2018).

³ $\log(y_i + (y_i^2 + 1)^{1/2})$.

References

- Adoho, F., S. Chakravarty, D. T. Jr. Korkoyah, M. Lundberg, and A. Tasneem (2014). "The Impact of an Adolescent Girls Employment Program: The EPAG Project in Liberia". *Policy Research Working Paper* no. WPS 6832.
- Amin, Sajeda, Johana Ahmed, Jyotirmoy Saha, Md. Irfan Hossain, and Eashita Farzana Haque (2016). "Delaying Child Marriage through Community-Based Skills-Development Programs for Girls: Results from a Randomized Controlled Study in Rural Bangladesh." New York and Dhaka, Bangladesh: Population Council.
- Anderson, M. L. (2008). Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American Statistical Association*, 103(484):1481- 1495.
- Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, and Munshi Sulaiman. (2020). "Women's Empowerment in Action: Evidence from a Randomized Control Trial in Africa." *American Economic Journal: Applied Economics*, 12 (1): 210-59.
- Buchmann, N., Field, E., Glennerster, R., Nazneen, S., Pimkina, S., & Sen, I. (2017). Power vs money: Alternative approaches to reducing child marriage in Bangladesh, a randomized control trial. *Unpublished Manuscript*.
- Callen, M., De Mel, S., McIntosh, C., & Woodruff, C. (2019). What are the headwaters of formal savings? Experimental evidence from Sri Lanka. *The Review of Economic Studies*, 86(6), 2491-2529.
- Chen, Gilad, Stanley M. Gully, and Dov Eden (2001). "Validation of a new general self-efficacy scale." *Organizational research methods* 4, no. 1: 62-83.
- Dupas, P., Karlan, D., Robinson, J., & Ubfal, D. (2018). Banking the Unbanked? Evidence from three countries. *American Economic Journal: Applied Economics*, 10(2), 257-97.
- García-Moreno, C., Jansen, H.A., Ellsberg, M., Heise, L. and Watts, C. (2005). "WHO multi-country study on women's health and domestic violence against women". World Health Organization.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.
- Molina-Millan, T. and Macours, K. (2017). Attrition in Randomized Control Trials: Using tracking information to correct bias, IZA Discussion Paper, 10711.
- Roy, S., M. Hidrobo, J. Hoddinott, and A. Ahmed (2019). Transfers, behavior change communication, and intimate partner violence: Postprogram evidence from rural Bangladesh. *Review of Economics and Statistics*, 101(5), 865-877.
- Rosenberg, Morris. (2015). *Society and the adolescent self-image*. Princeton University Press.
- Schwarzer, Ralf, and Matthias Jerusalem. (2010): "The general self-efficacy scale (GSE)." *Anxiety, Stress, and Coping* 12, no. 1 329-345.
- WHO (2015), *The Multi-country Study on Women's Health and Domestic Violence against Women*.

Appendix

A. Statements on gender-based violence

Table A1: Domestic violence

1	Is it justified for a husband/partner to hit/beat his wife if she burns the food?
2	Is it justified for a husband/partner to hit/beat his wife if she argues with him?
3	Is it justified for a husband/partner to hit/beat his wife if she goes out without telling him?
4	Is it justified for a husband/partner to hit/beat his wife if she neglects the children?
5	Is it justified for a husband/partner to hit/beat his wife if she refuses to have sex with him?
6	Is it justified for a husband/partner to hit/beat his wife if she talks to him about using protection against HIV?

Table A2: Emotional violence

7	Has anyone said or done something to humiliate you in front of other people in the last 12 months?
8	Has anyone threatened to hurt or harm you or anyone close to you in the last 12 months?
9	Has anyone insulted or belittled you in the last 12 months?

Table A3: Physical violence

10*	Has anyone hurt you physically, punched, shoved, slapped, hit you in the last 12 months?
11*	Has anyone kicked you, dragged you to the ground, tried to strangle you, burned you, threatened you with a knife or pulled a gun on you in the last 12 months?

Notes: Items noted with * are not available at baseline.

Table A4: Sexual violence

12*	Has anyone ever physically forced you to have sex when you didn't want to in the last 12 months?
13*	Has anyone forced you to do other sexual practices that you didn't want to do in the last 12 months?
14*	Has anyone touched you in a sexual way, I mean your breasts, your buttocks or your private parts, without your permission, touching without permission includes pinching, grabbing, rubbing or fondling against your will, either directly or through your clothing in the last 12 months?

Notes: Items noted with * are not available at baseline.

Table A5: Spouse control over the adolescent

15	In the last 12 months, did your husband/partner has been jealous/angry if you talk to other men?
16	In the last 12 months, did your husband/partner has accused you of being unfaithful?
17	In the last 12 months, did your husband/partner has prevented you from seeing your female friends?
18	In the last 12 months, did your husband/partner has tried to limit your contact with your family?
19	In the last 12 months, did your husband/partner has insisted on knowing where you are at all times?
20	In the last 12 months, did your husband/partner has not trusted you with money?
21	In the last 12 months, did your husband/partner has prevented you from working?

B. Statements on Socio-Emotional skills

Table B1: Problem-solving skills

1	When I am faced/confronted with a problem, I can usually find several solutions.
2	If I am in trouble, I can usually think of a solution
3	I solve most problems if I put the necessary effort
4	Thanks to my skillful and creative thinking, I know how to handle unforeseen situations
5	I can always solve difficult problems if I try hard enough
6	When I have a decision to make, I take the time to try to predict the positive and negative consequences of each possible option before I act
7*	When I have a problem to solve, one of the first things I do is get as many facts about the problem as possible
8*	When I am trying to solve a problem, I think of as many options as possible until I can no longer come up with any more ideas
9*	I am able to come up with new and different ideas
10*	I like to think of new ways of doing things
11*	I come up with new ways to do things
12*	I am an original thinker
13*	I plan tasks carefully

Notes: Items noted with * are not available at baseline.

Table B2: Perseverance skills

14	I finish what I begin
15*	Setbacks do not discourage me
16*	I am diligent
17*	If someone is against me, I can find ways to get what I want
18	It is easy for me to stick to my aims and achieve my goals
19	I am confident that I could deal conveniently with unexpected events.

Notes: Items noted with * are not available at baseline.

Table B3: Self-awareness skills

20	My behavior often puzzles me (reverse negatively coded items)
21*	I understand my own behaviors
22*	I am aware of my thoughts
23*	I monitor my thinking to ensure it is accurate
24*	I analyze my behavior after I make mistakes.
25	I know the skills I have that other people do not have
26*	I assess my strengths and weaknesses in new situations
27	I critique my own abilities
28*	I review how I am thinking when I make a mistake
29*	I have a clear sense of who I am
30*	I understand other people's thoughts, feelings and actions better than my own (reverse negatively coded items)

Notes: Items noted with * are not available at baseline.

Table B4: Emotional regulation skills

31	I can remain calm when I am facing difficulties because I can rely on my abilities to cope.
----	---

Table B5: Ability to influence

32*	I can communicate my ideas in a way that will convince people to agree with me
33*	People like to follow my ideas
34*	When I ask for help, I receive it
35	I am good at getting people to help me when I need it
36	I am good at examining social situations to determine how to present myself well
37*	I observe social situations carefully before deciding how to present an idea to others
38	I am able to adjust my behavior to make a good impression

Notes: Items noted with * are not available at baseline.

C. Statements on attitudes towards gender (household head and adolescent's mother)

Table C: Attitudes towards gender

1	Woman's most important role is to take care of her home and cook
2	Washing clothes, and bathing and feeding the children are mother's responsibilities
3	There are times when a woman deserves to be beaten
4	It is a woman's responsibility to avoid getting pregnant
5	A man should have the final word about decisions in his home
6	A woman should tolerate violence in order to keep her family together
7	I would feel insulted if my spouse/partner asks to use a condom
8	A man and a woman should decide together what kind of contraceptives to use
9	If someone insults me, I will defend my reputation, by force if I have to
10	To be a man, you need not to show weaknesses.
11	An unexcised woman is not faithful to her husband

D. Gender equality statements

Table D: Gender Equality Statement

1	Who should have the highest level of education in the family?
2	Who should earn money for the family?
3	Who should be responsible for washing, cleaning and cooking?
4	If there is no water pump or tap, who should go fetch water?
5	Who should be in charge of feeding and washing the kids?
6	Who should be in charge of caring for sick people?
7	Who should help the children with their homework?

E. Mental health statements

Table E: Mental health

1	Have you lost self-confidence?
2	Have you considered yourself as a useless person?
3	Have you been reasonably happy, overall?
4	Did you feel that you had a useful role in life?
5	Did you feel that you could not overcome difficulties?
6	Have you been able to enjoy your normal daily activities?
7	Have you been able to cope with your own problems?
8	Have you been feeling unhappy and depressed?