

Endline Contribution Analysis

# Young Africa Live

REACHDIGITALHEALTH.ORG

Reach Digital Health

---

**Author**

May 3, 2024

Executive Summary	5
Background and Context	5
Study Design and Findings	5
Contribution Story and Theory of Change Analysis	6
Recommendations	8
Conclusion	9
1) Introduction	10
1.1) Background and context	10
1.2) Project Overview	11
2) Purpose of the report	12
3) Overview of studies and programmatic data sources	14
3.1) Representativeness of each data source	14
Representative Groups	14
Representativity Claims	14
3.2) WhatsApp program monitoring data	15
Backend data description	15
Backend data representativity of target population: Comparison of platform users to national target population	16
Limitations to platform data	17
3.3) WhatsApp Chatbot Pre-Post Study	18
Baseline and Endline Survey Enrolment	18
Endline’s Demographic Representativity of Platform Users	18
State of SRH needs on the platform	19
Endline’s Representativity of young people’s SRH needs (population and platform)	21
Evaluation questions	21
Research hypothesis	22
Estimation strategy	22
Limitations of baseline and endline data	22
3.4) Facebook program monitoring data	24
Facebook page backend data	24
Limitations to platform data	24
3.5) Facebook Survey Methodology	24
Evaluation Design	24
Evaluation questions	24
Research Hypotheses	25

Estimation strategy	25
Limitations of the Facebook Study	25
3.6) Qualitative Study Methodology	26
Data collection process	27
Demographic information	27
Study Limitations	28
4) Contribution Story and Theory of Change Analysis	29
4.1) YAL Theory of Change	29
The theoretical model	29
Instruments used to proxy COM-B constructs	31
4.2) Implementation Fidelity to Theory of Change	31
4.3) YAL Contribution Story	33
Activity 1: Provision of in-depth content on SRH and Mental health topics	33
Output 1: Users see and have access to in-depth content on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships (based on their needs assessment scores).	35
Intermediary Outcome 1.1a: Users read the chatbot content that they are recommended based on their needs assessments (engagement)	36
Intermediary Outcome 1.1b: Users read the chatbot content that they are recommended based on their needs assessments (sentiment)	37
Intermediary Outcome 1.2: Improved knowledge and attitudes regarding contraceptives, sexual health, HIV and STI, sexuality, and healthy relationships.	39
Demonstrable changes in knowledge and attitudes	39
Subjective changes in knowledge and attitudes	43
Outcome 1.1 - Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	44
Demonstrable changes in SRH behaviours	44
Subjective changes in SRH behaviours	46
Outcome 1.2 - Improved uptake of clinical health services related to sexual and reproductive health	48
Activity 2: Provision of a Facebook platform with content and space for peer discussions	51
Output 2.1 - Users see content and peer discussions on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships	52
Output 2.2: Users post their own concerns seeking peer support or input	53
Intermediate Outcome 2.1: Users find posted content or comments relevant, interesting, and useful	55

Outcome 2.1: Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	56
Outcome 2.2: Improved uptake of mental and clinical health services related to sexual and reproductive health	57
Activity 3: Provision of a service finder tool for SRH and mental health services near to users	59
Output 3 - Users see and use service finder tool for SRH and mental health services	60
Intermediate Outcome 3 - Users find the service finder tool helpful	60
Outcome 3 - Increased linkages to healthcare facilities	60
Activity 4: Provision of depression/anxiety screening tool for mental health	63
Output 4 - Users use the depression screening tool	64
Intermediate Outcome 4.1 - <i>Users ask for the LoveLife help-desk to call them back</i>	65
Intermediate Outcome 4.2 - <i>Users directed to mental health content finish the mental health content syllabus</i>	66
Outcome 4.1 - Improved attitudes and behaviors regarding mental health	67
Outcome 4.2 -Short-term mitigation of serious psychological stressors for the youth experiencing depression	69
Impact - Improved SRH, mental health, health empowerment, and health persistence for an increasing number and percentage of adolescents and young people reached with the YAL platform.	70
5) Discussion	72
5.1) Analysis of TOC Validation	72
5.2) Overall Contribution Analysis Limitations	76
6) Recommendations	77
7) Conclusion	78
8) References	80
9) Appendix A - Regression specification	82
10) Appendix B - Survey Instruments	111
10.1) WhatsApp Baseline Survey Questions	111
10.2) WhatsApp Endline Survey Questions	118
10.3) Facebook - B-Wise page Survey Questions	129
10.4) Facebook - Avert Content Survey Questions	134
11) Appendix C - Additional report detail	140
11.1) WhatsApp Chatbot Pre-Post Study	140
Baseline and Endline Survey Enrolment	141

Endline’s Demographic Representativity of Platform Users	141
Endline’s Representativity of young people’s SRH needs (population and platform) 144	
Estimation strategy	145
11.2) Facebook Cross-sectional Study	145
Sample size and methods	145
Data tools	145
Data Collection	145
Data analysis	146
11.3) Qualitative Focus group discussions	146
Data management and analysis	146
Thematic Analysis (TA) was used to examine patterns and organise data into themes. The process undertook the following steps: familiarization of content, generation of initial themes, refining themes, and finally definition of and naming themes. The evaluation team used TA to understand the findings on the effectiveness of YAL’s mobile health intervention, document its impact, and share key recommendations.	146
Data tools	147
Sampling	147
11.4) The YAL TOC and instrument construction	148
The COM-B model of behaviour change	148
Instrument definitions	150
11.5) Activity 1 Supplementary Analysis	151
Output 1 - The COM-B model of behaviour change	151
Intermediary Outcome 1.2 - Improved knowledge and attitudes regarding contraceptives, sexual health, HIV and STI, sexuality, and healthy relationships.	153

## Executive Summary

Mobile health (mHealth) programs present opportunities to enhance the sexual and reproductive health (SRH) and mental health knowledge of adolescents and young people through information dissemination. This document constitutes the endline performance evaluation of the Young Africa Live B-Wise program in South Africa as implemented by Reach Digital Health and the other B-Wise partners including the Elton Johns AIDS Foundation, the South African National Department of Health (NDoH), Avert, and others. The report focuses on assessing users' perspectives regarding young people's access to and utilization of mobile phones to improve SRH and mental health education as well as the associated behaviours. The objective was to determine if the program facilitated access to SRH and mental health education information for young people and identify barriers that needed to be addressed in order to support positive behaviour change. The report presents findings approximately five months after the inception of the YAL program, marking the endline phase.

## Background and Context

South Africa, like other low- and middle-income countries (LMIC), has poor (SRH) indicators among young people. About 10% of the adolescent girls in South Africa are estimated to become pregnant before they reach the age of 20 years old. This is further compounded by an unmet demand for family planning services, highlighting a significant need for reproductive health support and interventions in this demographic (UNICEF 2021). In Africa, many young people lack access to accurate information and high-quality services for their sexual and reproductive health and well-being (UNAIDS, 2020). There are many reasons for this notable deficit of knowledge and access to services which include (but are not limited to) lack of access to youth-friendly services; fear of discrimination and judgement; harmful pre-existing social norms; lack of access to relevant, relatable, and accurate information; and the proliferation of mis- and dis-information through social media channels (Kafwanga et al., 2021).

The use of mobile and wireless technologies (mHealth) has the potential to transform health service delivery globally and support the achievement of the United Nations (UN) Sustainable Development Goals (SDGs) in LMICs. mHealth is increasingly used to deliver health interventions, including SRH interventions for young people. mHealth programs offer opportunities to improve SRH for young people by providing information and support. However, further research is required to inform the development of tailored approaches for this age group.

## Study Design and Findings

The mixed methods evaluation to understand the effectiveness and impact of the YAL platform incorporated three different study approaches and reviewed programmatic monitoring data available through the WhatsApp platform. The three studies included a quantitative baseline and endline comparative analysis for users of the B-Wise WhatsApp service, a one-time survey sent to users of the B-Wise Facebook page, and a mixed-methods qualitative study that interviewed (individually or through focus group

discussions) a sample of YAL users in Gauteng and Kwazulu-Natal. The combination of these different data sources was used to inform a programmatic Contribution Analysis that had three main goals, among others. The first was identifying contributions YAL has made towards SRH and mental health knowledge, attitude, and behaviour changes. The second was understanding the mechanistic relationships through which YAL led to those changes as compared to the program's underlying Theory of Change (ToC). And finally, providing evidence for decision-making to national and international partners and funders as it relates to future phases of the YAL program.

Two key components were considered before crafting the program's contribution story or ToC analysis. First, the research team needed to analyze how representative the data analysed through the endline evaluation were of two central populations: 1) the target population for the program which is South African youth between the ages of 15 and 24 years old with access to a smartphone, and 2) the general user base of the B-Wise platform (in other words all users who had accessed the B-Wise chatbot and completed the onboarding process). Upon analysis, the report finds that the users that partook in the endline survey are statistically different from the broader YAL user population in that there is an over-representation of female users and users in relationships. There are other platform use findings later in the report that also indicate these users engage more significantly than the broader population.

Therefore, the findings in the subsequent sections will report what aspects of the ToC appear to be validated for more highly engaged users that exist within the broader YAL user base.

## Contribution Story and Theory of Change Analysis

**ToC Pathway 1 (Activity 1): Provision of in-depth content on sexual health, mental health, and healthy relationships, through a WhatsApp-based platform, leads to changes in attitudes, knowledge, behaviour, and ultimately, increased uptake of services.**

The YAL program, as evidenced by three studies and programmatic monitoring data, demonstrates correlations between its B-Wise chatbot, which provides sexual and reproductive health (SRH) and mental health content, and improvements in knowledge and attitudes among users. With 100,000 users reached, the program effectively engages its target audience, although there are demographic discrepancies, notably in age, gender, and socioeconomic status. User feedback indicates that users found the content both relevant and useful. The data demonstrates statistically significant improvements in SRH knowledge, body image and consent attitudes, and condom usage (the latter specifically related to the subsample of respondents who are not planning to have a child in the next year) under one of the main regression models used in this paper. However, an additional model was used that included location variables but restricted the sample size significantly. This additional model's analysis demonstrated that some of these findings lose their statistical significance when

analysing the smaller sample for whom location variables were available. For these variables where both regression models are found to be consistent but show mixed statistical significance, more research is needed to confirm the validity of the statistically significant finding. Qualitative data highlights positive behavioural changes, including improved communication and comfort in discussing sensitive topics. However, linking these changes to clinical service uptake is challenging due to survey limitations and tool loss. Nevertheless, self-reported user intentions show an increased likelihood of healthcare facility visits and counselling, supporting the program's potential to generate additional service utilisation for users who actively engage with the platform's content and features.

**TOC Pathway 2 (Activity 2): Provision of a Facebook platform with content and peer discussions on sexual health, mental health, and healthy relationships leads to changes in attitudes, knowledge, behaviour, and ultimately, increased uptake of services.**

The evidence found that the Facebook platform effectively reached a broad audience, boasting over 29,000 Facebook followers and reaching more than 9 million individuals through paid content dissemination. Furthermore, it fostered a large quantity of user engagement, with over 8.9 million post engagements recorded over 13 months, showcasing the potency of social media platforms like Facebook, particularly within YAL's target demographic. User feedback revealed high levels of engagement, with 75% of Facebook group followers accessing the page daily or weekly and 72% of users engaging with paid content, reporting medium to high exposure levels for content shared in the past week. However, while the Facebook component facilitated SRH content viewing, user participation in sharing perspectives and commenting was limited, raising questions about its effectiveness in stimulating sustained peer discussion. Nevertheless, positively perceived content appears to have influenced user knowledge and intentions towards healthier behaviours, with 85% of respondents attributing these self-reported changes to the B-Wise platform, albeit without direct links between exposure to the B-Wise Facebook page and key outcomes. This underscores the need to revisit and refine the program's Theory of Change to understand and foster desired behavioural changes by including a socially focused component like a Facebook page.

**TOC Pathway 3 (Activity 3): Provision of a service finder tool for SRH and mental health services near to users, leads to increased uptake of services.**

Despite its limited duration, the platform came close to reaching its target of 40% awareness among users in need regarding the Service Finder tool's availability. Moreover, the tool demonstrated utility by attracting repeat searches and guiding users to recommended healthcare facilities, indicating its potential to drive user behaviour within the subset of users who were aware of the service and who then used it. However, due to the tool's deactivation, the project could only partially assess its long-term impact on facilitating linkages to healthcare facilities or analyse user experiences regarding recommended service quality. Nonetheless, endline survey data



revealed higher-than-expected user uptake of recommended services, suggesting promising progress. However, these findings are mostly applicable to the more highly engaged YAL user group as compared to the broader YAL user group, and the report shows that users present in the endline did use the tool more than the broader YAL user base. Future phases should incorporate ongoing service finder functionality, increased advocacy on such a tool's availability and purpose to the platform users, and enable critical feedback surveys to better understand service quality from users who do indeed make the jump from the platform to in-person healthcare services.

**TOC Pathway 4 (Activity 4): Provision of a LoveLife call-back feature, linked to mental health screening, leads to increased uptake of counselling services.**

The WhatsApp chatbot, integrated with a mental health screening tool in the YAL version 2 release, aimed to engage users in mental health assessments and direct at-risk individuals to support services. While 28% of invited users initiated the screening (a percentage that is significantly lower than the targeted 90% of invited users), a high completion rate of 93% was observed (which surpassed the target of 80% of users completing the assessment). This is a positive result to show that the tool itself is engaging once started, but more work can be done to better engage users on starting the self-assessment. Approximately 70% of completed screenings identified users as "at risk," with nearly 9,000 individuals recommended to seek counselling services. Engagement with the callback feature exceeded the target, with 26% of at-risk users utilising it, surpassing the target of 10%. In addition to meeting goals for screening completion and support service referrals, post-analysis via the WhatsApp endline survey revealed a statistically significant decrease in users experiencing issues with depression/anxiety and low social connectedness. Qualitative data suggests that the B-Wise chatbot positively impacted users' understanding of mental health concepts and provided support in navigating mental health challenges.

## Recommendations

Based on the TOC pathway analyses, this report identifies key areas for improvement in future iterations of the YAL platform, as evidenced by conducted studies. Firstly, reactivating the Service Finder Tool or developing a similar offering is crucial to enhance healthcare service utilisation among youth. Understanding reasons for low uptake, such as discomfort or lack of trust, can inform provider selection and improve user experience. Secondly, considering additional participatory focus groups to support the refinement of the YAL program for future phases can enhance representation of minority groups and tailor content to specific demographics, such as under-17s and LGBTQ individuals. Thirdly, ensuring affordability and accessibility, particularly through free modes and offline options, is vital to reach diverse users, especially those in rural areas. Lastly, enhancing the social component, particularly on Facebook, requires moderation for regular thematic content posting and peer-to-peer discussion facilitation to meet users' desires effectively. These improvements are essential for the platform's efficacy and inclusivity in addressing youth healthcare needs.

## Conclusion

The contribution analysis reveals valuable insights into the YAL program's TOC and intervention effectiveness for engaged users. While pathways 1 and 4 of the TOC show relatively strong support, the Facebook component and service linkage through the Service Finder tool, pathways 2 and 3, lack sufficient supportive evidence. Efforts are needed to sustain peer engagement and relevant content dissemination for any proposed social components while restarting the Service Finder feature and improving user knowledge of the feature should be prioritised for increased service utilisation in future phases of the program. The WhatsApp chatbot appears more impactful than the Facebook component, though given which outcomes were significantly improved suggests emphasising the need for enhanced user exposure to critical themes like contraception and STI testing. Additionally, the program design supports user capacity and motivation for SRH/mental health learning, suggesting continued potential for behaviour change among engaged users. To improve engagement across the platform, enhancing user experience, diversifying content approaches, and incentivizing consistent engagement are recommended. Overall, the YAL platform offers vital, accessible support for South African youth, however, with definitive areas for continued improvement, demonstrating the importance of continued investment in digital health interventions to meet youth needs effectively.

## 1) Introduction

Mobile health (mHealth) programs present opportunities to enhance the sexual and reproductive health (SRH) and mental health knowledge of adolescents and young people through information dissemination. This document constitutes the endline performance evaluation of the Young Africa Live B-Wise program in South Africa, focusing on assessing users' perspectives regarding young peoples' access to and utilisation of mobile phones for SRH and mental health education. The objective was to determine if the program facilitated access to SRH and mental health education information for young people and identify barriers that needed to be addressed in order to support positive behaviour change. The report presents findings approximately five months after the inception of the YAL program, marking the endline phase.

### 1.1) Background and context

Young people under the age of 25 years account for 43% of the world's seven billion people (World bank Data 2021). Young people are defined as individuals between the ages of 15–24 years old. This is a crucial period in life during which young people undergo extensive biological, psychological, and sociological changes. It is a crucial time for lifelong health development, and improving health behaviours at this stage of life contributes to the health of future generations. SRH is integral to health and wellbeing during adolescence and beyond. Empirical evidence over the past 20 years has highlighted the challenges faced by adolescents in accessing SRH information and services, including contraception to prevent unplanned pregnancy and mental health.

South Africa, like other low- and middle-income countries (LMIC), has poor (SRH) indicators among young people. About 10% of adolescent girls in South Africa are estimated to become pregnant before reaching the age of 20 years old. This is further compounded by an unmet demand for family planning services, highlighting a significant need for reproductive health support and interventions in this demographic (UNICEF 2021). The 2017 HSRC HIV surveillance report revealed that knowledge levels about HIV have seen a decline. Additionally, while condom use at last sex among 15–24-year-olds is higher than in other age categories, it has also experienced a decline.

In Africa, many young people lack access to accurate information and high-quality services for their sexual and reproductive health and well-being (UNAIDS, 2020). There are many reasons for this notable deficit of knowledge and access to services which include (but are not limited to) lack of access to youth-friendly services; fear of discrimination and judgement; harmful pre-existing social norms; lack of access to relevant, relatable, and accurate information; and the proliferation of mis- and dis-information through social media channels (Kafwanga et al., 2021).

The use of mobile and wireless technologies (mHealth) has the potential to transform health service delivery globally and support the achievement of the United Nations (UN) Sustainable Development Goals (SDGs) in LMICs. mHealth is increasingly used to deliver health interventions, including SRH interventions for young people. mHealth programs

offer opportunities to improve SRH for young people by providing information and support, but further research is required to inform the development of tailored approaches for this age group.

## 1.2) Project Overview

Reach Digital Health, in collaboration with the Elton John AIDS Foundation (EJAF) and Avert, initiated the Young Africa Live (YAL) platform as a response to the deficit in high-quality information for youth. YAL aims to empower young people across Africa, providing them with knowledge and confidence regarding their sexuality, relationships, sexual and reproductive health, and mental well-being. The platform aimed to assist them in making informed choices to enhance both their physiological and psychosocial well-being. It was designed to support:

1. The dissemination and effective consumption of accurate, relevant and relatable information, addressing topics young people care most about which influence their health and wellbeing.
2. Discerning young people's health needs and connecting them with either virtual or physical support services that can effectively address these needs.

Overall, the project aimed to:

1. Understand the motivations and (mis)information that puts young people's sexual health at risk.
2. Create a knowledge base among young people by providing honest, relevant, accessible, and engaging information on the issues that motivate young people to empower them to make informed sexual health decisions and generally maximise their well-being.
3. Provide a forum for young people to safely discuss these issues and learn from one another.
4. Increase demand for, and linkage to, high-quality adolescent-friendly sexual and reproductive health/ family planning/ HIV services and products, including self-care through an array of screening algorithms.
5. Improve self-reported sexual and reproductive health behaviours among young people, through 2, 3, and 4 above.
6. Empower young people to have a voice in their care both in terms of choice and in providing feedback on what they are currently receiving.

YAL provided a blended digital communications and engagement approach. Facebook content and discussions were intended to serve as a platform for more public conversations and discussions among peers, directing users to the private conversation space on WhatsApp if needed. With the WhatsApp service, users could interact anonymously with a tailored chatbot. The WhatsApp chatbot served as a private space for users to ask sensitive questions, seek advice, access sexual health information, and engage with edutainment content. It also facilitated links to appropriate external services when necessary. Additionally, users were referred to external content through [www.B-Wisehealth.com](http://www.B-Wisehealth.com) for extended information as needed. The broad functional architecture of the platform included:

Table 1: Functional architecture of the YAL platform

Function	Description
<b>Browsable content</b>	Browsable and searchable content, including edutainment offerings extending across WhatsApp and Facebook, was created. This content was regularly updated with new material based on what resonated most with the youth
<b>Chatbot</b>	A gamified chatbot with multiple personas was developed to connect with users of different profiles. This guided users through journeys on the platform and answered their questions using natural language processing technology to identify intent and select the best response.
<b>Guided user journeys</b>	Structured decision-trees, in the form of text-based menus, supported informed choices around users' health and well-being. When an offline service was deemed the best outcome, the platform suggested the closest service centre based on geolocation or provided linkage to an appropriate virtual service
<b>Digital screening tools</b>	digital screening tools were envisioned to support young people on a journey of health empowerment. These tools assessed, for instance, clinical eligibility for certain HIV prevention methods such as PrEP, or determined whether a young person needed to be linked to mental health or Intimate Partner Violence (IPV) support
<b>Geomapping</b>	Link users to physical services or link them to a service provider on the WhatsApp platform who can provide live information via text consultation.
<b>Channel crossover</b>	Content was developed on Facebook and Instagram to drive relevant traffic to the privacy of the WhatsApp Chatbot and build the user base. The WhatsApp Chatbot also directed users back to specific Facebook pages and/or other online platforms to engage with peers and to access more detailed information on topics of interest.
<b>Peer-based discussion</b>	Moderators encouraged and monitored peer-to-peer discussions on social media channels, notably on Facebook landing pages where key topics could be explored in greater detail. By leveraging the peer-to-peer functionality of Facebook linked to the secure 1-1 chatbot, the program aimed to tap the need for both privacy and peer recognition in dealing with sensitive issues.
<b>Audience insights</b>	Ongoing anonymized social listening and engagement with youth to help learn what topics young people cared most about, better understand their motivations in engaging in risky behaviours, and discover what content resonated best with them

## 2) Purpose of the report

The mixed methods approach to evaluating the effectiveness and impact of the YAL platform incorporated three different study approaches and the review of programmatic monitoring data available through the WhatsApp platform. The three studies included a

quantitative baseline and endline comparative analysis for users of the B-Wise WhatsApp service, a one-time survey sent to users of the B-Wise Facebook page, and a mixed-methods qualitative study that interviewed (individually or through focus group discussions) a sample of YAL users in Gauteng and Kwazulu-Natal.

The research team designed such a diverse approach to quantify the changes the platform contributed to for YAL users and better understand how those changes happened. The report also aimed to identify any unforeseen benefits or consequences of the YAL approach on users' knowledge, attitudes, and behaviours regarding SRH and mental health. In addition to the three studies, the research team also integrated data from programmatic monitoring data to add additional depth to the review of user engagement and relevance of the platform's content. The purpose of taking a contribution analysis approach to this program evaluation was to achieve the following:

- **Identify Contributions:** First, the report aims to determine the specific contributions that YAL made towards achieving the desired SRH and mental health outcomes. Moreover, the report examines the extent to which the different aspects of the YAL program have influenced the observed changes.
- **Understand Mechanistic Relationships:** Second, the report analyses each step along the YAL Theory of Change. This model captures the assumed mechanistic pathways through which the platform leads to the desired changes. This analysis will identify relevant evidence that either supports or challenges the assumed relationships between various inputs, activities, outputs, and outcomes. This will aid in validating or refining the underlying theory of change by examining whether the observed outcomes align with the expected causal pathways and assumptions.
- **Provide Evidence for Decision-Making:** Third, this report will offer evidence-based insights to stakeholders, including partners within the South African National Department of Health (NDOH), EJAF, and Reach leadership and implementation teams. These insights will focus on the effectiveness and impact of the intervention to inform future decision-making, resource allocation, program design, and improvement efforts.
- **Enhance Accountability and Learning:** Fourth, the report will promote accountability by evaluating the intervention's contributions. Additionally, identify the program's strengths and weaknesses to allow for adaptation and improvement in future implementation phases.

## 3) Overview of studies and programmatic data sources

### 3.1) Representativeness of each data source

#### Representative Groups

In evaluating the effectiveness of the YAL program, it is first necessary to identify for which users this report's findings are relevant. This report focuses on three groups of not necessarily overlapping groups;

- A - All South African youth aged 15-24 with access to a smartphone, as the universe of the potential target population.
- B - All users that reach the YAL platform, complete onboarding, and are within the target age group. Comparisons between this group and group A would aim to identify whether Reach has been able to advertise the service to a representative sample of the target demographic and whether there is sufficient interest in joining the platform from those users reached by advertising (separating the platform reach and user interest may not always be possible given the data).
- C - Target users that not only reach the platform but engage with the platform as intended, including subscribing to push-notifications, reading some degree of the prescribed content and completing some needs assessments. Comparisons between groups B and C would aim to identify whether there might be statistically significant differences between these groups that might influence the representativeness of this report's findings for the broader YAL user base . The working assumption is that users who have completed both the baseline and endline surveys potentially represent a higher engagement cohort of the YAL user base. Such differential engagement may be due to the platform being more relevant, interesting, or useful to some groups, or it may reflect some unobservable differences in the populations, such as motivation, consistency, time availability, or other factors. Separating these themes may similarly not always be possible given the data that is available from the three studies or programmatic monitoring data.

#### Representativity Claims

It's important to consider which group the report's data source speaks to as it impacts which groups the report's findings can be extended to. This is crucial for making informed decisions and drawing accurate conclusions for the evaluation sample and, if possible, the broader YAL user base.

- If the report finds that the demographics of the general YAL user base, Group B, are representative of the general youth population in South Africa between the ages of 15-24, Group A, **and** the report demonstrates that the sample users that completed both the baseline and endline survey (e.g. a more engaged set of

users whose data are used for the report’s analyses, or Group C) is representative of the general YAL user base, Group B, then the findings resulting from analysing responses from a more highly engaged subset of users could be reflective of trends for the platform population, Group B, and the generally targeted population, Group A. This would speak to the general value of the platform for the average youth user in South Africa.

- However, if Group C is found to differ from platform users, Group B, either demographically or by default for being more engaged than other users, then findings from Group C may not validly extend to the average young person on the platform, Group B, and likely not to the broader population, Group A. However, whether the platform is associated with impact for those users who engage with the platform (group C) can still be of interest as a finding in itself, indicating that the platform may be impactful for users with similar traits as users found in Group C.

The remainder of this section explains four data sources available for this evaluation and details what representativity claims can be made from each, given the above considerations.

## 3.2) WhatsApp program monitoring data

### Backend data description

From their first message to the WhatsApp chatbot, each cellphone number that engages with the platform has each message and interaction stored at the individual level. From its launch to November 30th, 2023, the WhatsApp chatbot has received messages from 111,658 unique cell phone numbers across its various recruitment strategies<sup>1</sup>. After messaging the line, users are led through a short onboarding process in which users sequentially agree to the platform’s data privacy policy and indicate their age, gender, relationship status, country of residence, income bracket, previous experience with the platform, and complete a series of four questions asking about their external-internal locus of control (IE-4). The last step in the onboarding process is for users to indicate if they are willing to receive daily notifications from the platform, which Reach theorizes is mechanistically necessary for any change in knowledge, attitudes, or behaviours.

Since its launch to November 30th 2023, of the 111,658 unique cell phone numbers contacting the page, 85,588 (71%) complete the full YAL onboarding process, disclosing their demographic information. Although the use of mobile health applications and chatbots has grown exponentially in recent years, definitions of user uptake and engagement is highly variable. Additionally the majority of studies focus on developed country contexts, both making direct comparison to YAL difficult (Torous et al., 2018; Ng et al., 2019). In their review of the real world uptake of digital interventions for

---

<sup>1</sup> This section is concerned with assessing the representativity of the data sources. We return to a discussion of the platform’s performance against the projects stipulated SMART goals, such as 100,000 users, in Section 4.



depression, anxiety and low mood, Fleming et al. (2018) found only one study that reported on an application's registration rate, calculating this at 42%.

### Backend data representativity of target population: Comparison of platform users to national target population

If we assume that drop-off during the onboarding process is constant for all demographic groups<sup>2</sup>This data would represent the population of users that are sufficiently interested enough in the platform to complete registration (Group B as discussed in Section 3.1). Comparing this platform population against the national target population (Group A as discussed in Section 3.1), shows several key similarities and differences. In terms of age, 87% of users in Group B are within the target years of 15-24, with 1% of users being below 15 years old and the remainder being above 24 years old. As such, 13% of users on the platform fall outside of the age range for the national target population. Within those users who are in the target age group (15-24), there are disproportionately fewer users aged 15-17 years old (7%), as opposed to users aged 18-20 years old (44%) and 21-24 years old (49%). This large imbalance is believed to be due to delays in gaining approval from Meta to advertise the platform to underage clients. While uptake among this group has increased since gaining approval, the platform sample, therefore, differs from the national target population in terms of age, being skewed towards users aged 18-24.

Of all registering users who were asked and disclosed their household income, there is a relatively high proportion of users in no-income homes (55%)<sup>3</sup>. The next highest modal response (9%) is users indicating that their household has a total income of R1-R400 per month, and 5% have an income of R401-880 per month. This would imply that 69% of all users supposedly live below South Africa's lower poverty line of R945 per month per capita (Statistics South Africa, 2022). In contrast, SALDRU (2023) estimates that 35% of South Africans live below the poverty line. Statistics South Africa (2020) also indicates that only 21% of households with youth in them have income below R1800 per month. As such, if this data is to be trusted, this would mean that the sample is dramatically skewed toward low-income earners. Given these disproportionate results, it is perhaps more likely that young people are not directly aware of the actual income of their households or may have neglected to consider grants or other piecemeal income when reporting their household income. There are no means of confirming, but it is possible that respondents misunderstood the question and reported their personal income since other studies find that one-third of South African youth in a sexual and reductive health intervention cite earning less than R1600 per month (Closson et al., 2019). Regardless of the cause, we believe that the distribution of answers indicates that it is likely that there is a substantial measurement error in terms of income. However, if such a high proportion of users at least indicate that "no income" is their best guess of household income; this would indicate that most users at least believe

<sup>2</sup> Since it is only through the onboarding process that we gather such demographic information, this assumption cannot be tested.

<sup>3</sup> Unfortunately, no household size data was gathered at registration, as such income for all registration data is captured only at the household level.

their homes to be financially disadvantaged. While this does not represent the broader South African youth population, the platform would appear to be primarily engaging with users from low-income households.

Regarding gender, the platform attracts a representative sample; 49% identify as female, 47% identify as male, and 1.5% identify as non-binary (with 2.5% preferring not to disclose their gender). This shows a representative sample of men and women and an over-representation of non-binary people - estimated nationally at 0.1% (Beyond Zero, 2021). Relationship status is varied, with 54% identifying as being in a relationship, 28% being single, and 18% indicating that "it's complicated". Given that most census and national surveys focus only on youths' marital status rather than romantic relationship status, Reach was unable to find statistics to compare this proportion against.

Finally, since the onboarding process does not capture data on specific health behaviours and the key SRH barriers, we cannot compare how well the general platform represents the SRH needs of the country. The best approximation of this is found in the results of the WhatsApp baseline data, which can be seen below.

Taken together, the data gathered for users accessing the platform, Group B, can be seen to be targeting a relatively even distribution of men and women from largely low-income households, who are mostly between 18 and 24 years old, and the majority of users are romantically involved (either in a relationship or otherwise). For this group of users, all messaging engagement with the platform is then monitored and stored at the individual-level interactions in Reach's data management system. This data is used to calculate any measure of all users' engagement with the platform, such as; content engagement, service finder engagement, and LoveLife engagement, with individual measures discussed in Section 4. The comparison of Group C to Group B will be addressed in section 3.3.

### **Limitations to platform data**

Regarding the B-Wise chatbot data, since Reach managed the chatbot, direct access to all WhatsApp data has been possible. An obvious limitation to this data in terms of its representativity is that it can only gather data for those users who engage with the onboarding questions and in subsequent messaging. As such, it cannot assess which users the platform is reaching but fails to enrol. Additionally, from a user experience perspective, the onboarding process has to be limited to a short number of non-threatening questions, which has informed the few demographic questions gathered. However, this limits the comparability of the platform's reach in terms of broader SRH needs on the platform, requiring the use of the baseline survey as the next best approximation of SRH needs on the platform. One final limitation is that Reach cannot currently review the "status" of a message once sent, whether the message is read or not. As such response to push notifications has had to be inferred through inbound messages to the platform received from users on the same day they receive push messages.

### 3.3) WhatsApp Chatbot Pre-Post Study

#### Baseline and Endline Survey Enrolment

To gather data on i) the type of users engaging with the platform and ii) possible changes in users' barriers, behaviours, and outcomes through engagement with YAL, a simple pre-post study design was adopted on the WhatsApp chatbot service. All users registering with the chatbot between 31st May and 7th June 2023, meeting the study eligibility criteria<sup>4</sup> were invited to participate in a 45-question, voluntary baseline survey. Of the 1999 users invited to participate over this period, 1295 (65%) consented to and completed the baseline survey. All baseline users were then invited to complete the 75-question, voluntary endline survey 5 months later. Between November 23rd and January 16th, 502 (39%) of the 1295<sup>5</sup> baseline completers consented to and completed the endline survey.

Both the baseline and endline survey gathered information on users' registration levels of; i) SRH knowledge, ii) SRH attitudes, iii) psychological capacity, iv) psychological resilience, and v) SRH behaviours, for comparison between the two periods, as well as user demographics and platform review data<sup>6</sup>. Both surveys compensated users for their time. For more detail on the baseline and endline surveys please see Appendix C, Section 11.1.

#### Endline's Demographic Representativity of Platform Users

Before discussing what research questions the pre-post study was designed to enable, it is first important to understand which groups the endline sample is representative of. A statistical analysis comparing the average age, gender, relationship status and total household income of users at registration between those users that complete endline and those that just access the platform, as well as comparing demographics at registration between users that complete endline and those that only complete baseline, is provided in Appendix C, Section 11.1. Based on this analysis we argue that the endline sample is relatively representative of platform users aged 18-24, who have subscribed to outbound messages<sup>7</sup>, for women, and with a slight bias toward users in relationships. And caution should be used in extending these findings to young male and single users, who appear to be disproportionately underrepresented in the endline

---

<sup>4</sup> To be included in the study users needed to; i) be between 18-24 years old, ii) live within South Africa, iii) have never used the B-Wise chatbot before, iv) consent to push messages and v) complete YAL's full registration process. Although the YAL platform also serves the SRH needs of people aged 15-17, it was decided that these users should be excluded from the research scope since consent from an appropriate guardian could not be adequately gathered over WhatsApp.

<sup>5</sup> For the endline survey, the total responses were limited to a maximum of 500 users due to funding and project timeline limitations, so this is not demonstrating a drastic attrition for willingness to participate from baseline to endline

<sup>6</sup> Unfortunately, due to a coding error 10 questions were not exposed to 302 of the 502 endline respondents including 2 questions related to users' location level data which will be relevant to this paper's model specification.

<sup>7</sup> Although, relaxing this condition we find no substantial changes in any of these distributions, indicating that the endline sample would also represent a similar group on non-subscribed users.

sample. Later in section 4, we'll also see that these users self-report higher rates of engagement with both the service finder and the LoveLife callback feature, than the general user base. Combined with the fact that endline users were invited to the final survey over WhatsApp, this may indicate that the endline sample should be considered relatively more engaged than the average user.

### State of SRH needs on the platform

The above analysis captures what selection effects are at play in terms of demographic variables for the full sample. However, endline users could be systematically different in their starting SRH levels. As mentioned above, since the onboarding process only gathers demographic factors it is not possible to compare all platform users against national statistics. In this case, the baseline sample provides the most information of users SRH needs at registration. It is worth noting that analysis in Appendix C, Section 11.1 shows that the baseline response group is largely representative of the overall platform sample in terms of demographics. As such it is at least plausible that these baseline results are representative of the needs of users registering on the platform. Column 1 of Table 2 below presents the proportion of users experiencing various SRH needs at baseline across 5 categories; SRH psychological capability, knowledge, attitudes, behaviours, and psychological persistence. Column 2 looks at these indicators for users who only completed the baseline, and column 3 looks at these same indicators for users who did both the baseline and endline.

Table 2: TOC analysis of attrition from baseline to endline

Variable	All baseline respondents (1)	Did baseline only (2)	Did baseline and endline (3)	Difference (2-3)	P-value (2 vs 3)
<b>Psych Capacity</b>					
Depressed or Anxious at baseline	76%	76%	77%	-1%	
Misusing substances at baseline	22%	23%	20%	3%	
Low social connection at baseline	80%	81%	78%	3%	
<b>Knowledge</b>					
Low knowledge at baseline (Less than 1.5 correct on 3 SRH knowledge questions)	16%	12%	23%	-11%	***
<b>Attitude</b>					
Poor body image at baseline	33%	31%	37%	-6%	**
Poor sex positivity at baseline	17%	17%	16%	1%	
Poor gender at baseline	38%	41%	34%	7%	***
Poor consent valuation at baseline	25%	26%	22%	3%	
<b>Behavioural</b>					
Used contraception at last sex at baseline	56%	53%	59%	-6%	*

Used condom at last sex at baseline	47%	48%	46%	2%	
Less than 2 sexual partners last month at baseline	91%	90%	93%	-3%	**
Having ever tested for an STI at baseline	78%	76%	80%	-4%	*
Poor self perceived healthcare at baseline	48%	47%	50%	-3%	
Persistence					
Low locus of control at baseline	41%	44%	36%	8%	***
Low self-esteem at baseline	22%	21%	23%	-2%	
Total users	1,295	793	502		

A general description of this population at baseline shows that the average baseline respondent is someone with high levels of depression/anxiety, is not abusing substances, has high social disconnection, with high knowledge of SRH information, good body image, sex positivity, and relatively good gender and consent attitudes. It shows that generally there is inconsistent use of condoms within this population, though it appears the users don't have multiple partners and have generally been tested for STIs previously. In addition a majority of these users have good self esteem and a relative internal locus of control. This description is somewhat more resilient than what might be assumed for individuals coming from vulnerable households with no income, which further supports the assumption that there was measurement error in the reporting of household income.

Column 1 of Table 2 above, shows meaningful evidence of a need for intervention among the sample. The results for psychological capability are mixed with 76%, 22% and 80% of users identified as "at-risk" of depression/anxiety, misusing substances, and low social connection. In comparison to national statistics, Craig et al. (2022) find that 26% of South African adults score as depressed/anxious on the PHQ-9, indicating that the sample may experience more depressive symptoms than the national population.

Of all baseline respondents, 16% do not appear to have sufficient SRH knowledge<sup>8</sup>. Given that SRH knowledge is a constructed index of questions, direct comparison to other statistics is difficult. Indeed, there is a high degree of variability in measures of SRH knowledge in the literature. The nationally representative South African Demographic House Survey (2016) finds that youth's knowledge of at least one valid form of contraceptive was "near universal". However, depth of knowledge appears variable, with many young people not knowing how to use a particular contraceptive despite knowing the method (Pleaner et al., 2022). As such, a knowledge rate of ~16%

<sup>8</sup> Note that this result differs slightly from the baseline report result of 12%. This is due to one knowledge question being left out for a sub-sample of users at the endline. In order to make the samples comparable all users' baseline scores were adjusted to only include questions that they were also asked at the endline.

appears plausible, given this study's SRH questions which are a mix of simple identification and some understanding of STIs.

In terms of attitudes regarding sexual relationships, column 1 shows substantial room for improvement, with 33%, 17%, 38%, and 25% of users identified as having poor attitudes about their body image, sex positivity, gender equality in relationships, and valuation of consent in sex respectively.

In terms of SRH outcomes, only 56% and 47% of sexually active baseline respondents used any form of contraception or condoms when last having sex. This is somewhat lower than the 60% of sexually active youth who report using contraceptives (SADHS, 2016) and 59% of sexually active South African youth who self-report using a condom at the last sexual encounter (Simbayi et al. 2019), indicating that baseline respondents are at least as poor in terms of condom usage as the national target group.

Most users have ever tested for STIs/HIV (78%) and have had 1 or fewer sexual partners in the last month (91%), however, 48% of users still identified themselves as having relatively poor sexual and reproductive healthcare behaviours. Lastly, while only 22% of the sample had low self-esteem at baseline, 41% of users had a predominantly external locus of control.

#### **Endline's Representativity of young people's SRH needs (population and platform)**

As mentioned previously, of the 1,295 users that took the baseline survey, 502 went on to complete the incentivized endline survey. This represents a 61% attrition from baseline to endline, but it should be noted that this is due to budgetary constraints within the project to reduce the sample size while still trying to be sufficiently powered. Section 11.1 provides an evaluation of how well this endline sample represents the initial baseline sample. It concludes that those users completing the endline survey differ from those just completing the baseline in terms of being predominantly female and somewhat more likely to be in a relationship when registering. The endline sample also has lower initial SRH knowledge, worse body images, better initial gender attitudes and a larger proportion of users have an internal locus of control. That said, the endline sample is relatively representative of the psychological capacity, of the SRH attitudes, and the majority of SRH behaviours of the baseline sample, however, only for those users similar to the endline sample in the ways just mentioned. As such, an extrapolation from these users onto either the average youth or the average subscribed target user should therefore be seen with caution. While the endline sample is therefore not representative of all users, the sample is still a highly valuable group to assess, in order to determine what kinds of changes in SRH outcomes are possible on the platform, even if only for these kinds of users.

#### **Evaluation questions**

The evaluation objectives included assessing changes in SRH knowledge, attitudes and behaviours, along with understanding users' experiences of the platform and its perceived impact. As such, the pre-post evaluation was designed to answer the following research questions:

1. Establish endline levels of SRH knowledge, attitudes, perceptions, and behaviours of adolescents and young people aged 18-24 in South Africa for users who completed the WhatsApp Chatbot journey and those who dropped off.
2. Compare changes in SRH knowledge, attitudes, perceptions, and behaviours between baseline and endline for users who completed and dropped off the chatbot.
3. To establish whether YAL's target groups are adequately retained on the WhatsApp chat bot after enrolling in the intervention.
4. To gauge users' experiences of the WhatsApp chatbot as an intervention and its various features.

**Research hypothesis**

Following the YAL TOC, we would expect:

1. Engagement with the WhatsApp chat bot will be significantly associated with decreases in barriers to sexual and reproductive health, namely;
  - a. low SRH knowledge, poor attitudes regarding one's body image, sex positivity, gender equality in relationships and consent in sex, depression and/or anxiety, low interpersonal connectedness and high substance reliance.
2. Engagement with the WhatsApp chat bot will be associated with an increase in either of the identified SRH persistence measures (internal locus of control and/or self-esteem) for 18-24 year olds in South Africa.
3. Users will report positive experiences and regular use of the B-Wise WhatsApp chatbot's features, as well as substantial impacts due to the platform from their subjective perspective.

**Estimation strategy**

As in any pre-post study, a change in variables from baseline to endline is the primary means of inferring impact in this study. In such cases, a simple McNemar test or paired t-test is appropriate for estimating changes in proportions or means between the periods. However, to account for the possibility that the nature of trends varies between groups or to account for exogenous variables that could affect an outcome of interest and also change over time, this study employed the use of paired subjects, mixed model linear regressions. For a justification of the choice of this estimation strategy against other statistical methods, please see Appendix C, Section 11.1.

**Limitations of baseline and endline data**

A first limitation to note is that all results from the McNemar tests or mixed-model regressions aim to identify the effect of time on the outcomes of interest, using this as a proxy for the possible effect of the platform. Where other exogenous and unobserved variables may also change across time this may therefore incorrectly identify the effect of the program. Unfortunately, this is a natural constraint of pre-post studies. While we are able to control for some exogenous changes, such as changes in income, there are a number of changes that are likely unobserved. The coefficient on time therefore serves

as our best proxy of the possible effect of the program on each of these outcomes. As mentioned above, there is reason to be sceptical of users' income responses, baseline household size has had to be inferred by endline numbers, and the only other time variant demographic variable captured by the survey is relationship status, which is later argued to be at least partially endogenous to the program. As such the ability to partial out time-variant changes exogenous to the model is limited. Additionally, users' location-level variables were, unfortunately, not captured for over 50% of the sample. As such this report's preferred model specification does not include location-level variables. Annex A provides a motivation for this report's preferred model, and it demonstrates that the inclusion of location variables (for the sub-set of respondents with location data), does not meaningfully improve the model's goodness of fit.

In order to detail one strength and limitation of the pre-post study, we must first outline the natural functioning of the B-Wise chatbot. All users that sign up to the WhatsApp chatbot are led through a short onboarding survey, and those that opt into regular notifications then receive a sequence of assessments spread out over their first few weeks on the service. These assessments include the PHQ-4 depression/anxiety screening, assessment of users SRH knowledge, attitudes, and behaviours, their locus of control on the IE-4, and many more. These assessments are used by the platform to determine users' prescribed content buckets and the relative priority of topics. These questions are then automatically asked again once users have received all of their prescribed push-message content.

A strength of the study is that the recruitment strategy and execution of the pre-post study were designed in order to mirror this user journey and all pre-post assessments that are already asked on the YAL platform were now asked in the baseline and endline survey for those users willing to participate in the study (with users that complete the baseline or endline survey then not receiving those same in-built platform assessments). While the baseline and endline surveys are therefore relatively representative of the type of survey engagement that users would naturally be asked on the platform, this does also mean that the baseline and endline surveys only capture responses from users willing/able to answer long surveys in one sitting. Additionally, both the baseline and endline surveys are associated with relatively large financial airtime compensations which may lead to a consideration that those incentives have inserted bias regarding which users answer the baseline and endline survey. There is no evidence in the data, however, to support this claim given that income is statistically indistinguishable for platform users, baseline users, and those users that go on to complete endline. Nonetheless, given the concerns regarding the accuracy of the income variable this should be viewed with some caution,

There were also a few unfortunate errors in the execution of the surveys. Each of these has already been stipulated explicitly above. These errors either meant that some questions were not asked of all participants, reducing the sample size for some analysis, or were left out of the baseline and had to be inferred from the endline (such as household size at baseline being inferred from the household size at endline).



## 3.4) Facebook program monitoring data

### Facebook page backend data

Engagement data related to the B-Wise Facebook page was gathered using the analytics and reporting tools provided by Facebook. These tools offered insights into the performance of the B-Wise Facebook platform, including the total number of page visitors, impressions per post, and the number of users clicking on Facebook ads. Additionally, engagement metrics such as likes, comments, and shares were tracked to assess user interaction with the content posted on the page. However, given Facebook's privacy policy, all data is only accessible in aggregated form and no individual-level data is provided by any of the Facebook tracking tools. Additionally, a limitation of this data includes the possibility of double counting in these metrics, which may inflate the reported engagement figures. This may occur when a single user interacts with a post multiple times, thus artificially inflating the reported engagement figures

### Limitations to platform data

For the Facebook data, that all user data is only accessible in aggregate form has meant that linking behaviours across the platforms is not possible. As such linkages between platforms cannot be directly observed and are rather self-reported. Additionally, management of the Facebook arm of the project fell under the responsibility of a partner organisation, tasked with providing the required data for the Facebook indicators. Accessing this data posed a challenge as it was not readily available. Specifically, comprehensive metrics detailing the performance of the Facebook page from the program's inception to the present were unavailable and consequently not provided in this report.

## 3.5) Facebook Survey Methodology

### Evaluation Design

In January 2024, the research team invited over 200 WhatsApp chatbot users aged 18-24 to complete a once-off survey to reflect on their experiences with the complementary B-Wise Facebook page, including the frequency of their engagement, their perception of the content, peer engagement and perceived effects the platform had for them. A total of 178 usable surveys were received and all respondents were compensated for their time. For a more detailed description of the Facebook Cross-sectional survey, please refer to Appendix C Section 11.1.

### Evaluation questions

1. To establish whether the Facebook arm of the B-Wise platform adequately reached a representative sample of South Africa's youth aged 18-24.
2. To gauge users' perceptions of whether the B-Wise Facebook page adequately fulfilled its intended purposes as set out in the B-Wise theory of change.
3. To identify possible correlations between users' reported engagement with the B-Wise Facebook page and important Sexual and Reproductive Health Behaviours.
4. To gauge users' engagement with both the B-Wise Facebook page and the Young Africa Live WhatsApp chatbot.
5. Establish a link between Facebook respondents and WhatsApp chatbot engagement rates.

### **Research Hypotheses**

1. The majority of users report that discussions on the B-Wise Facebook page were relevant, interesting, and useful for their sexual health needs.
2. Users who report greater degrees of engagement on the B-Wise Facebook page (total duration and frequency) on average reported healthier sexual and reproductive behaviours.
3. Most users on the B-Wise Facebook page registered with the B-Wise WhatsApp Chatbot.

### **Estimation strategy**

To understand how different factors interact, we conducted a multiple logistic regression analysis on a number of key binary variables of interest. The models included all relevant variables like sex, household status, social class, HIV status, and previous exposure or participation in SRH discussion. The logistic regression provided estimates of associations between these variables and self-reported SRH knowledge, attitudes, and behaviour, helping us explore what influences self-assessments. We looked at each variable's contribution in explaining the variance in self-reported health, showing coefficients, standard errors, and significance levels.

### **Limitations of the Facebook Study**

The findings of this evaluation should be considered given certain limitations. Firstly, the evaluation lacked a baseline measure or control group, and the cross-sectional study design prevented us from establishing causal relationships regarding program effects. The study relied on participants' recall of campaign exposure, which could introduce biases, such as reverse causality. In other words, individuals who reported higher exposure to the campaign might already hold strong opinions about SRH issues. Consequently, their reported exposure to the campaign might be influenced by their pre-existing attitudes rather than the intervention itself. To address this, we attempted to minimise the effect by incorporating three levels of exposure to assess the frequency of use-response relationships rather than solely comparing exposed and not exposed groups.

Secondly, the evaluation relied on retrospective reports of potentially sensitive information, introducing the possibility of recall and reporting biases. Some information,

such as details about the last sexual encounter and participation in program activities or receipt of services, pertaining to past events could be influenced by these biases. Individuals may struggle to accurately remember specific details about their last sexual encounter, leading to inaccuracies in their reports. Similarly, when asked about their participation in program activities or receipt of services, individuals may selectively recall events or overestimate their level of engagement due to memory limitations. Further, the evaluation is susceptible to social desirability bias, as participants might feel compelled to respond to questions about attitudes, behaviour, and the applicability and usefulness of B-wise in a socially acceptable manner, possibly influenced by B-wise content. While this bias is inherent in self-reported outcome measures, the fact that users complete surveys independently may help alleviate this concern compared to in-person or phone surveys conducted by enumerators. Furthermore, participants were guaranteed anonymity and confidentiality, encouraging an environment conducive to providing unbiased responses. Additionally, by emphasising that there were no right or wrong answers, participants were reassured and encouraged to respond truthfully, potentially enhancing their sense of psychological safety during the study (Bendoly, 2014).

Another limitation relates to the recruitment strategy. Delays in survey approval by the National Department of Health (NDOH) necessitated a change in the recruitment method. Initially, the plan was to recruit participants by posting a link on the Facebook page visible to all members and page users. However, this approach had to be changed due to the requirement for NDOH approval to post on the Facebook page. Consequently, the recruitment strategy was amended to use the WhatsApp database to reach potential participants. This change may have introduced a potential for selection bias, as the demographics and behaviours of individuals reached via WhatsApp may differ from those targeted through the original Facebook recruitment strategy

### 3.6) Qualitative Study Methodology

Although quantitative research provides statistical data that measures and describes a causal relationship or lack thereof between variables of interest, it does not give an in-depth understanding of those variables, nor does it give an understanding of the participant's individual experiences and social reality. Hence, Reach sought a qualitative inquiry to add in-depth insights into the participants's experiences of the platform to this evaluation. Conducted through focus groups and individual interviews, this qualitative component sought to examine whether and how the intervention worked for YAL WhatsApp and Facebook users, identifying barriers and facilitators to using the platform and validating causal mechanisms hypothesised in the theory of change. Thus, allowing users to identify what they believe to be the primary drivers of changes in their knowledge, attitudes, and behaviour (or, in the alternate case, why changes may have not occurred).

## Data collection process

An open-ended semi-structured interview guide was used to collect data with AYPs (participants). Nine Focus Group Discussions (FGDs) were conducted with youth based in KwaZulu Natal and Gauteng Province. The qualitative research team conducted two face-to-face FGDs in Gauteng Province and facilitated the other seven discussions via WhatsApp. The consultants used WhatsApp to conduct discussions with AYPs who (a) preferred online interactions, (b) could not make it to the face-to-face sessions due to school examinations or being at work, (c) who had relocated or (d) who feared meeting the researchers without prior communication from B-wise. A total of nine Individual Interviews (IDI) were conducted with AYPs. Six interviews were conducted face-to-face in KZN, while the other three were conducted virtually in Gauteng.

A major limitation of both the WhatsApp and Facebook quantitative surveys is that they were not able to include users younger than 18, since parental permission could not be obtained. As such reaching under-age users was a key goal of the qualitative component. Unfortunately, the research team was only able to recruit and interview three minors. Given the qualitative study was the only opportunity to solicit feedback from this group, the report will try to highlight useful feedback from this demographic group, however, this indicates that findings for users below 18 are significantly constrained in this report and more work could be done in the future to engage with under-18 users to enrich these findings. feedback received from the adolescent respondents is not necessarily reflective of broader learnings for the 18-year-old and older population of YAL users.

## Demographic information

Among the participants who participated in the evaluation, were 34 young women and 19 young men, with a total of 53 participants. Table 3 below summarises the demographic information of the research participants. Eight out of these 53 identified as from the LGBTQI+ community. From the LGBTQI+ community, most identified as either Bisexual or Gay. A total of five were in school, 17 were out of school (either employed or in tertiary), and 31 were not in school, not employed, or in training (NEETs). This study had participants predominantly from Gauteng Province. In Gauteng, youth were based in the City of Johannesburg, the City of Tshwane, the city of Ekurhuleni, and the city of West Rand. The participants from KZN were based in Ugu, uMgungundlovu and eThekweni municipalities.

Table 3: Summary of participants' demographic information

Young men	Young women	Heterosexual	Bisexual/ Gay	In school	Employed or Tertiary	NEET
19	34	45	8	5	17	31

## Study Limitations

Like most studies, the current study's design is subject to several limitations. Thus, the results must be interpreted cautiously, and the enumerated limitations should be considered. These limitations, however, present themselves as an opportunity and key reflections for engaging with young people for future research:

- **AYP who changed their numbers:** Some provided numbers were no longer active. Determining how long these numbers had been out of service was difficult.
- **Relocations:** AYP had relocated from the areas they registered when they first used the Chatbot. This was why some participants could not participate in the face-to-face FGDs, even though they were still active on the Chatbot.
- **From chatbot to face-to-face:** AYP who pulled out of the sessions after agreeing to participate in the discussions. Some had agreed to the sessions but later feared that the researchers might not have been legitimate, therefore pulling out at the last minute. To address this concern once it was spotted, the platform pushed reassuring messages to the intended users verifying the identity and authenticity of the researchers to bolster their confidence in the study. Thus, users who did ultimately participate in the focus group discussions were less wary of the researchers involved.
- **Network challenges and load shedding (power cuts) schedules:** During the FGDs, there were times when the discussions would be interrupted by load-shedding, which affected the flow of the conversations. Coordinating the discussions was also challenging because of the different load-shedding schedules.
- **Minors:** Some of the minors refused to give contact details of their caregivers for the researchers to obtain consent for participating in the discussions. Participants expressed that they did not want their families to know that they were on the Chatbot in case they asked to see what kind of information they were consuming. This was a limitation of the study.
- **Time Constraints:** The tool for data collection was not piloted due to time constraints. This meant that any challenges or invalid questions that did not capture the information they were meant to measure were experienced during implementation and thus only identified during the analysis.

## 4) Contribution Story and Theory of Change Analysis

### 4.1) YAL Theory of Change

A main focus of implementing this contribution analysis is to examine the veracity of the originally proposed Theory of Change (ToC) and to understand what assumptions held true, which assumptions did not, evaluate the outcomes that the intervention produced, and understand what deviations occurred during implementation that could have impacted program effectiveness or the underlying theory within the model. To improve the sexual, reproductive, and mental health of youth in South Africa, YAL's program was designed based on the COM-B behavioural change model (see Appendix C, Section 11.4 for more detail). This model posits that interventions that impact individuals' capability, opportunity, and motivation can lead to improved behaviours. Capability refers to an individual's knowledge, skills, and ability to engage in the behaviour. Opportunity refers to factors that enable individuals to execute a specific behaviour. Motivation refers to an individual's disposition to want to do the behaviour instead of treating it as a taxing necessity (West and Michie, 2020).

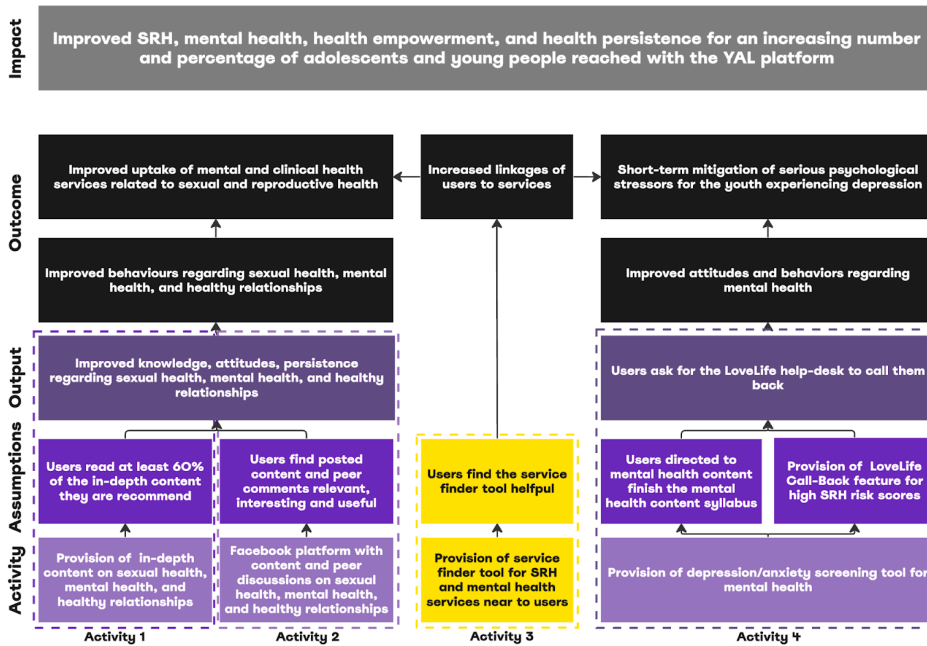
#### The theoretical model

The program focused on four central interventions:

1. Provision of in-depth content on contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships via engaging users through a chatbot.
2. Provision of a Facebook platform with content and space for peer discussions on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships.
3. Provision of a service finder tool for SRH and mental health services near to users.
4. Provision of depression/anxiety screening tool for mental health. All users subscribed to push messages were invited to complete the screening tool, with all users identified as "at-risk" being directed to the LoveLife Call-Back feature. This LoveLife Call-Back feature was also available through the platform's main menu.

These four activities, along with their associated assumptions, outputs, outcomes, and impacts are represented visually in a simplified version of the YAL ToC in Figure 1 below. For the more comprehensive ToC, please access the following [Miro board](#).

Figure 1: Simplified YAL Theory of Change



In summary, the theory of change posits that if the platform provides users with in-depth content on SRH and mental health themes, with the opportunity to engage with additional content and peer discussion through an associated Facebook page, while supporting users to find relevant services near them, and offering a screening tool to assist users with assessing their mental health, then users should demonstrate improved knowledge, attitudes, and persistence regarding sexual health, mental health, and healthy relationships, and, ultimately, impact user uptake on mental and clinical health services related to sexual and reproductive health.

**Instruments used to proxy COM-B constructs**

Reach then identified a number of indices to serve as the best available proxy to measure each of these components. SRH knowledge is captured by one index of 3 questions on safe sex practices. SRH attitudes is comprised of 4 sub-indices, measuring personal body image, sex positivity, beliefs about gender equality in sex and users' valuation of consent in sexual relationships. Three indices independently monitored users' psychological capabilities; depression/anxiety as measured by the PHQ-4, substance misuse and degree of perceived social connectedness. The team were interested in 5 SRH behaviours; condomisation at last sex, contraception at last sex, having 1 or fewer monthly sexual partners, having ever tested for an STI, and users' self reported perception of their quality of self health care. Finally, two measures of persistence were monitored; users' locus of control as measured on the IE-4 and their self-esteem as measured on the Rosenberg SE-10. For more detail on the exact questions comprising the indices and how they align with the COM-B model, see Appendix C, Section 11.4.

The sections that follow will first investigate how YAL implementation reflected this ToC and any areas where there may have been deviations and then will analyse the evidence produced through the three studies and programmatic monitoring data to evaluate where the theory's causal assumptions held true and which assumptions, if any, need to be refined and improved for future implementation.

## 4.2) Implementation Fidelity to Theory of Change

In general, the YAL program implementation followed the framework proposed in the ToC; however, a few external events led to some deviations across certain activities.

A critical deviation from the proposed model was the loss of the Service Finder capabilities of the YAL platform. Through Service Finder, youth seeking access to public and private care services (either prompted or by their own volition) should have been able to use this feature to be linked to professional physical or virtual health services based on their needs and a specified location. These services should include clinics and health facilities, contraceptive provision and family planning, PrEP and PEP provision, safe spaces and care for victims of abuse, and educational services for learning new skills related to SRH. In addition, youth were meant to be able to rate their experience of using these services. This data was meant to be fed back into the chatbot's design to help improve the quality of recommended services. Unfortunately, the feature had to be deactivated in June 2023 following a contract termination between the NDOH and their partner WitsRHI. The feature has remained deactivated since, though the partner organisation SoulCity has indicated their intention to reactivate the feature in the future.

A second deviation to be considered was the Facebook component of the YAL ToC. As of June 2023, there was a decline in posted content on the page and a lack of thematic content posted at any time in the post-June 2023 period which means users were not being exposed to additional SRH content via this component of the B-Wise platform. It



was intended to be a space for additional user engagement, where youth could also engage in peer discussions and ask specific questions to the moderators of the Facebook page. This was significantly absent in the period post-June 2023 and thus within the Facebook study survey, respondents were asked to recall their engagement from over seven months ago. In addition, the program faced challenges in obtaining timely approval from the Department of Health (DoH) for new ad content, leading the team to reuse previously posted top-performing ad posts or those that did not receive optimal exposure. This recycling of content might have diminished interest among some youth, affecting the effectiveness of the Facebook component. The ToC highlighted the importance of providing a platform for peer discussions to enhance youth knowledge, attitudes, and the adoption of healthier sexual and SRH behaviours. However, the decrease in content posting on the Facebook page and the reuse of ad content could have affected the ToC pathway.

Finally, there was a slight deviation within Activity 4. The original plan was to have a digital screening tool as part of the platform's broad functional architecture. This tool would help young people in their journey towards health empowerment by assessing things like eligibility for HIV prevention methods like PrEP, or determining if a young person needs support for mental health or Intimate Partner Violence (IPV). Initially, WitsRHI was supposed to fund this as part of the YAL program, but the program had to change course and focus on the LoveLife call centre due to the partnership with WitsRHI not coming to fruition.

The next section of this report summarises the contributions of the YAL program to its intended outputs, outcomes, and impact measures.

## 4.3) YAL Contribution Story

### Activity 1: Provision of in-depth content on SRH and Mental health topics

Level	Description	Indicators
Output 1	YAL users see and have access to in-depth content on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships (based on their needs assessment scores)	Number of registered WA users Number/ Percentage of users that link between Facebook and WhatsApp channels
Intermediary Outcome 1.1	Users read the chatbot content that they are recommended based on their needs assessments	LF-29a: Average of The percentage of days individual users send messages to the line, relative to the days they have been receiving push-messages i.e. opt-in/subscribed users. (for all users active 2 days post-registration) LF-29b: Average of The percentage of days individual users that are screened as depressed/anxious send messages to the line, relative to the days they have been receiving push messages. i.e. opt-in/subscribed users. (for all users active 2 days post-registration)
Intermediary Outcome 1.2	Improved knowledge, attitudes and persistence-barriers regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	Percentage of users who initially scored at risk on SRH knowledge at baseline that then show sufficient knowledge at endline  Percentage of users who initially scored at risk for the following sub-set of SRH literacy, body image, sexual positivity, gender attitudes and consent assessments (see column G), that then show a reduction in each barriers score at endline  Percentage of users who initially scored at risk for the following sub-set of connectedness, depressions/anxiety, substance misuse assessments (see column G) that then show a reduction in each barriers score at endline

Outcome 1.1	Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	<p>Percentage of users that show an improved total score on the self-perceived healthcare from baseline to endline</p> <p>Improvement in the average number of healthy behaviours reported by WhatsApp users from baseline to endline</p>
Outcome 1.2	Improved uptake of mental and clinical health services related to sexual and reproductive health	<p>Percentage of users that indicate they are more likely to visit a clinic or other health facility for their sexual needs since being on B-Wise</p> <p>Percentage of users that indicate they are more likely to visit a counsellor clinic or other health facility for their sexual needs since being on B-Wise</p>

To understand if the YAL platform effectively provided in-depth content to users, the research team analysed data from the relevant results framework indicators shown above as well as specific questions that were included in the comparative analysis of the baseline and endline survey data and analysis of findings from both the Facebook survey and qualitative study.

**Output 1: Users see and have access to in-depth content on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships (based on their needs assessment scores).**

Launched on the 27th of October 2022, the B-Wise chatbot allows users to browse and search over 474 content pieces which are advertised to users through a series of 174 push messages. The content pieces include all of the above-mentioned themes, as well as guidance on seeking clinical assistance, enjoying sex, investigating gender, and understanding one's body. All content was developed in collaboration with Avert and approved by the South African National Department of Health (NDOH).

To expose a sufficient number of youth users to the B-Wise content, the platform aimed to recruit 100,000 users to the WhatsApp chatbot (SMART Goal 1). From its launch to November 30th 2023, the WhatsApp chatbot has received messages from 111,658 unique cell phone numbers, with 85,588 completing the YAL onboarding process in its entirety. In terms of whether the platform managed to recruit its intended age group, it's worth noting that not all users who access the platform disclose their age (30% do not reach this question). As detailed in Section 3.2, of the 78,160 users that disclose their age, 68,040 (87%) are 15-24 years old. Assuming that non-response in registration is independent of age this would imply that 96,807 of the 111,658 unique numbers belong to youth in the target demographic. While therefore very nearly achieving the target goal, Section 3.2 has already noted that of these users, youth aged 15-17 make up a relatively small percent of the target group reached. This large imbalance is related to delays in gaining Meta approval for advertising to minors. As such attaining approval has already seen improvements in the representation of minors on the platform and Reach will need to monitor that these efforts lead to more representative proportions in the future.

Beyond having users reach the platform, Reach's logical framework also aimed to have 25% of arriving users subscribe to regular push messages. In this regard, the platform exceeded its goal, with 38,825 users (36% of all unique numbers) subscribing to regular messaging. With these data in mind, the platform recruited very close to the number of target users, with a higher percentage of those users registering to receive the daily messages which would drive users to the in-depth content on the targeted themes. Finally, Reach had aimed to have 25% of users link between the WhatsApp chatbot and the Facebook channel (SMART Goal 2). Due to Meta's privacy policies, individual level Facebook usage data is not available to the research team. As a next best available data

point, endline users were asked about their knowledge and use of the Facebook page. For detail on these responses, refer Section 11.5 in Appendix, which finds that 67% of the endline respondents cite having visited the Facebook page and that 51% of those users indicate visiting the Facebook page monthly or weekly. As such the aim of having traffic from the WhatsApp chatbot reaching the Facebook page appears to be supported, at least for the type of users represented in these groups.

### **Intermediary Outcome 1.1a: Users read the chatbot content that they are recommended based on their needs assessments (engagement)**

At the start of the project, the research team set out to measure user engagement with content by analysing the percentage of users that read prescribed content, with a target of 25% of users reading 60% that is sent to them, based on their needs (SMART Goal 3). Unfortunately, it was discovered that it is impossible to identify which content pieces users have been prescribed based on their barrier assessments. This is because the list of content users receive for each assessment outcome has gone through many iterations as content has been added to the platform without a record of these historic mappings being kept. As such, the data structure would be unable to identify which content pieces users should have seen based on their assessment results over the intervention period.

This limitation was shared with the EJAF team in September 2023, and the research team and EJAF agreed to a suitable proxy measure for this intermediary outcome which was SMART goal 29a: *On average, users send messages to the line on 15% of the days that they receive push-messages (for all opted in users active 2 days post-registration)*. This works as a relatively strong proxy of SMART goal 3 and a measure of engagement overall since users are sent daily content concerning the themes identified in the needs assessment. These messages only invite users to engage with the content. To gain access to the educational content, users must reply to a push message with the name of the content piece they would like to see. As such, the percentage of days that users request access to their prescribed content should closely mirror the percentage of the prescribed content that users have engaged with.

Across the launch of the program until the end of November 2023, we find that, on average, users send messages on 15% of the days they receive push messages, achieving the SMART goal. Interestingly, this statistic shows a slight downward trend throughout the program's run time, with users responding to push messages on 22% of the days they receive push messages between launch and Feb 2023, 20% between March and May 2023, 15% between June and August 2023, and 13% between September and November 2023. This downward trend appears to be a symptom of users becoming less engaged over time. Indeed, when considering engagement relative to enrolment date, Table 4 below shows a similar trend, with users' average response rate at 26% within the first two weeks of registration, dropping to 15% by the second week, and then hovering around 10% for the rest of the program.

Table 4: Change in engagement rates by days post registration

Days post reg	Percentage on LF-29a	Days post reg	Percentage on LF-29a
2-13	26%	56-69	9.7%
14-27	14.5%	70-83	8.6%
28-41	12%	84-97	9.1%
42-55	10.5%	98-111	11.8%

This pattern of high initial drop-off rates that then quickly flattens is seen as a “ubiquitous phenomenon” for most digital interventions (Druce, Dixon and McBeth, 2019), with interventions characterised by high enrolment and high initial attrition. The lack of common definitions for attrition rates across both studies and meta-analyses makes comparing YAL’s attrition rate against other interventions difficult (Amagai et al., 2022; Druce, Dixon and McBeth, 2019). In their analysis of 8 large-scale, m-health studies focusing on mental health, Pratap et al. (2020) find that the median participant retained in each study varied widely, from between 6 days to just 2 days. Pooling the data from all 8 studies, they find the median duration of engagement pre-dropout to be just 5.5 days.

**Intermediary Outcome 1.1b: Users read the chatbot content that they are recommended based on their needs assessments (sentiment)**

As identified in the YAL theory of change, an avenue that m-health interventions should prioritise in order to keep users engaged on the platform and for them to internalise content, is to ensure that users feel that the content is relevant to their SRH needs, is interesting, and is useful in their own lives. SMART goals 4 and 5 speak to this: *65% of users find the WhatsApp chatbot content relevant, interesting, and useful, and, on a 5-point Likert scale the content averages a score of 3.5 for each of those categories.* Qualitative questions to this effect were included in the endline survey and are reported below.

Table 5: Users’ feedback on the platform and features

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>WhatsApp content</i>				
Content related to your sexual needs	Likert: 5	Very related (46%)	Related well (27%)	493
Content was interesting when it related	Likert: 5	Very interesting (57%)	Quite interesting (25%)	498

Content was useful for managing your SRH and relationship needs

Likert: 5

Extremely useful (63%)

Quite useful (25%)

501

Table 5 above shows that 73% of endline respondents thought the chatbot content was either very or well related to their sexual needs (46% and 27% respectively). When asked about how interesting the relevant content was, 82% of respondents indicated that they found the content either very or quite interesting (57% and 25% respectively). Finally, 88% of respondents indicated that the content was either very or quite useful for managing their sexual and reproductive health and relationship needs (57% and 25% respectively). Treating each Likert answer as a numeric response on a continuous scale, with 1 as the lowest possible score and 5 as the highest possible score, the average scores for relevance, interest, and usefulness of the content are 4, 4.4, and 4.5 respectively. Translating these scores into average sentiments should be viewed with some caution, since the distance between each interval on the scale is not necessarily constant, however, assuming it was, this would indicate that on average, users find the content related well to their needs, was quite-to-very interesting and was quite-to-extremely useful for their needs. Overall, this shows strong support for the claim that the content was relevant, useful, and interesting, at least for the type of users represented in the endline survey (more highly engaged users).

In addition to the quantitative findings from the WhatsApp endline analysis, the qualitative study also asked about the relevance and relatedness of content during the FGDs or IDIs. The qualitative analysis found that participants said comprehensive information on relationships, including decoding what type of relationships the participants are in, and information about sex are some of the topics and content relevant to them as young people. One of the participants elaborated further on this and said that on the platform, you can talk about relationship questions that would be difficult to discuss with parents.

**"That's not easy for us to ask our parents about relationships, but with the bot, we can ask questions regarding the kind of relationship that you're having and how the relationship is going, and the bot can tell you if ever you guys are just dating or you are just dating for fun"**

**[Female participant, focus group discussion]**

Focus group participants found the content on the platform relevant to their needs, particularly regarding mental health, sexual health education, intimate relationships, and gender identity.

**"We live in communities where once a young person says they are feeling depressed or anxious, they are told they like attention. Many of us do not know how to talk about mental health. We do not even know what mental health is."**

**[Male Participant, focus group discussion]**

Moreover, they reported high levels of engagement with the platform, accessing information on various topics and seeking support when needed. Respondents shared that the WhatsApp chatbot being available throughout the day made it easy to use and

convenient for the participants when they had questions, especially those in school or working. Participants shared said the chatbot is a place where you can ask many kinds of questions and is more accurate and precise than their peers when answering questions,

**“...when I got there (on the Chatbot), I saw that I could ask whatever question I wanted or had. I could ask it, especially the questions I could not ask the people around me. I feared judgement so much, but on this platform, I can also ask about those topics I would be shy about. I learnt a lot.”**

**[Female participant, focus group discussion]**

**“I believe the difference is that people are sometimes misinformed or wouldn’t go into detail. But with the App, you get information even more than what you expect”**

**[Female participant, focus group discussion]**

Another participant stated that the platform should also include teenagers below the age of fifteen years because that is the phase during which many adolescents start to be curious about different topics like sex and sexuality and want to explore their bodies. It was reported that adolescents start at the age of twelve to be curious about their bodies and want to experiment sexually. Some participants added that some teenagers also start exploring intergenerational relationships and/or transactional sex with older partners (Blessers and sugar daddies/mamas); hence, the app would be beneficial to them.

## **Intermediary Outcome 1.2: Improved knowledge and attitudes regarding contraceptives, sexual health, HIV and STI, sexuality, and healthy relationships.**

### **Demonstrable changes in knowledge and attitudes**

Section 3.3 of this report found significant SRH needs among the sample of users completing the baseline survey, with particular room for improvement in SRH attitudes, behaviours, persistence, and psychological capabilities in particular. Section 11.5 in Appendix C details that between baseline and endline, average income increased substantially for those users that completed the endline survey. As such, rather than simply compare respondents’ baseline and endline scores with a McNemar test, it is more appropriate to use a paired subjects, mixed model regression in order to isolate the unique effect of the change in time on users’ knowledge and attitudes. Columns 1 and 2 of Table 6a below present the results of 2 mixed-model specifications (Model 1 and Model 4 – see Appendix A for an explanation of all five models that were considered) for each knowledge and attitudinal barrier separately. For ease of interpretation, only the coefficient of time (moving from baseline to endline) is reported, with column 1 reporting the coefficient and p-value on time for a simple regression of time (as a dummy variable) on each outcome of interest. Column 2 then reports the research teams’ preferred



model specification, which includes all baseline demographic variables<sup>9</sup> (except for location level variables), as well as an interaction effect of income per capita and time. Note that no interaction of changes in relationship status and time were accounted for due to the potential endogeneity and bias this would introduce into the regression. For an explanation motivating Model 4's goodness of fit relative to other specifications, please see Appendix 9A. To be noted, the research team did consider an additional model, Model 5, which was similar to Model 4 but where geographic location, province and urban indicators specifically, were used as an invariant variable in the regression equation. However, that additional model had a significantly lower sample size due to the inconsistency of asking users their location across survey instances, and as noted in Appendix A, had some consistency issues when performing goodness of fit tests for certain outcomes. Thus, throughout the report, the research team will refer to the preferred model as the main point for discussion, but in cases where this additional model was found to be consistent, sufficiently powered, and in disagreement with the main model's finding, we will point out that inconsistency to demonstrate where some results need further exploration. A summary of which models were consistent on which outcomes, as well as where statistically significant results appeared and whether the regressions were sufficiently powered to detect those results, is found in Table A11 in Appendix A.

For ease of interpretation, the coefficients on time have been colour-coded, with green, red, or no colour, representing a statistically significant "beneficial change", a statistically significant "detrimental change", or a non-statistically significant change over time, respectively.

Table 6a: Coefficient on time for all barriers of interest from mixed model regressions

Controls included in model specification	Mixed model regressions			
	1	4	4	4
Time invariant controls	-	X	X	X
Urban and province dummies	-	-	-	-
Interaction of time and income	-	X	X	X
	<u>Coefficient on time dummy</u>		<u>ICC</u>	<u>Power</u>
<i>Knowledge</i>				
Prop. With low SRH knowledge	-0.066***	-0.075**	0.42	0.813
<i>Attitudes</i>				
Prop. With poor body image	-0.145***	-0.130***	0.37	0.999
Prop. With poor sex positivity	-0.011	-0.017	0.29	0.076
Prop. With poor gender attitudes	-0.028	-0.011	0.39	0.126
Prop. With poor consent attitudes	-0.082***	-0.080**	0.21	0.925

<sup>9</sup> The full list of time invariant variables are; age, gender, HIV status, incidence of weekly hunger (only available at endline), exposure to other SRH content pre YAL, baseline relationship status and log of household income per capita.

Sample size (range across outcomes)	502	424	424	424
-------------------------------------	-----	-----	-----	-----

Column 1 presents the results from the univariate mixed-model regression of time on each outcome of interest, including no additional controls. It is equivalent to a straight McNemar test as reported in Table 6a and mirrors those findings. The results of the fully specified model in column 2 are largely similar to the coefficients in column 1 of the univariate regression of time on each of the outcomes. Column 2 implies that once accounting for all possible time variant and invariant confounders across the period, the proportion of users with poor SRH knowledge decreases by 8 percentage points significant at the 5 percent level (exceeding the target of a 10% decrease from baseline, ie. 2.3 percentage points). The proportion of users with poor body images or consent attitudes is found to decrease by 13 and 8 percentage points, respectively, significant at the 1 percent level and 5 percent level (exceeding the target of a 10% decrease from baseline, ie. 3.7 and 2.2 percentage points respectively). The coefficients on time for the percentage of users with poor sex positivity or gender attitudes are not significant in either the straight McNemar, nor fully specified regressions.

Finally, columns 3 and 4 report the intra-class correlation coefficient (ICC) attained when running the fully specified mixed model and the post-hoc power calculations associated with each outcome given the observed change in the proportion of outcomes from baseline to endline and the associated ICC. This shows that all 3 of the outcomes that see significant changes are powered at or above 80%. As such, this would imply that the above changes serve as reliable estimates of the change in each barrier of interest over the course of the program, at least for platform users similar to those represented in the endline. Note that the table also indicates that the sample is underpowered to statistically distinguish the small observed changes in sex positivity and gender attitudes from random chance. However, this study is only interested in validating claims of changes in proportions of 10 percent or larger.

The measures for SRH knowledge, individual’s body image and attitudes regarding consent in relationships are each indices made up of 3, 2 and 2 questions respectively. As such, table 6b below presents the decomposition of these questions using the fully specified model as motivated above.

Table 6b: Coefficients on time for each item on the barrier indices decomposed

	No.	Proportion at baseline	Time dummy	ICC	Power
<i>Knowledge</i>					
Prop. who do <u>not</u> believe that using condoms every time reduces risk of STIs	424	5%	-0.034*	0.32	0.136

Prop. who do <u>not</u> believe that having only one sexual partner reduces the risk of STIs	165 <sup>10</sup>	20%	-0.061	0.31	0.085
Prop. who use a relatively ineffective method of contraception	392	29%	-0.100***	0.38	0.904
<i>Body Image</i>					
Prop. who do not “feel good” about themselves	424	40%	-0.105***	0.37	0.986
Prop. who do not “feel good” about their bodies	424	45%	-0.079**	0.37	0.984
<i>Consent</i>					
Prop. who agree that Robert has the right to force Samantha to have sex in the vignette	424	4%	0.034	0.13	0.105
Prop. who are not comfortable saying no when they don’t want to have sex	392	28%	-0.103***	0.24	0.967

It shows that the platform is associated with a 10 percentage point ( $p < 0.01$ ) reduction in the number of respondents who report using a relatively ineffective form of contraception (from 29% at baseline). Post-hoc calculations show that this change is also powered at 90%. This finding is relevant in South Africa, which has a high rate of unintended pregnancies. In their nationally representative survey of South African women, Chersich et al. (2017) find that two-thirds had unintended pregnancies in the past 5 years. They also found that half of women aged 15-19 and 20-25 who had become pregnant during the period had specifically not wanted to have a child when becoming pregnant. Therefore, it is meaningful to see significant substitution away from relatively ineffective contraceptives (such as the rhythm or pull-out methods) towards effective contraceptive methods (such as condomisation, injectables, and orals).

While there is, therefore, an increase in knowledge of effective contraception, there is no change in knowledge about STI prevention. There is no observable change in the percent of respondents that do not know that having an exclusive sexual partner or always using condoms reduces the likelihood of contracting an STI. However, both of these sub-indices are substantially under-powered, making it improper to draw conclusions as to whether knowledge of these factors changed over time or not. Additionally, the lack of an effect of time on the condom knowledge may be due to a floor effect, where 95% of respondents at baseline already knew that condoms decrease the likelihood of contracting STIs.

Similarly, the change in attitudes regarding consent in sex appears to be driven by one of the indices, with a 10 percentage point reduction ( $p < 0.01$ ) in the number of users who are uncomfortable “saying no” when they do not want to have sex (from 28% at baseline, powered above 90 percent). Unfortunately, we were not able to find statistics

<sup>10</sup> This question was only included in the second release of the endline survey, and as such is missing for most endline respondents.

in the literature regarding the prevalence of this sentiment among South African youth more broadly. There is no observable change in the proportion of users that agree that a man has the right to force sex as described in the vignette, however, only 4% of people agreed with this statement at baseline - presenting another possible ceiling effect, in addition to the statistic that is found being underpowered.

The significant change in the body image index appears to have been driven by both of its component questions. The proportion of respondents that do not "feel good" about themselves or their bodies decreased by 11 percentage points ( $p < 0.01$ ) and 8 percentage points ( $p < 0.05$ ) respectively (from a baseline of 40% and 45% respectively). Both results are also powered at the 90 percent level. In their review of the literature, Nolen & Panisch (2022) find that positive body images are associated with contraceptive use, condom use, STI testing, preventative sexual behaviours and seeking medical attention. Additionally, they note that the literature has broadly recognised a strong relationship between body image, the prevalence of eating disorders, exercise and mental health. As such improvements in each of the body image sub-indices is a valuable outcome.

When comparing these results from the Model 4 regression to those that resulted from using the Model 5 regression when it was found to be consistent, there is an additional level of uncertainty that is brought to the surface. For the overall knowledge index, the knowledge subcomponent on condom usage and effective contraceptive methods, Model 5 found results that were in the same direction, but lower magnitude. In addition, these Model 5 estimates were not statistically significant, and all were not sufficiently powered to find results of those magnitudes. Therefore, there is agreement across models that there was a general trend towards a reduction in these outcomes, but with the increased specification but lower sample size, the Model 5 regression did not find statistically significant changes. The research team thus believes that while there is indication that the knowledge and attitude changes are trending in the desired direction given Model 4 findings, further studies that are sufficiently powered and further specified would need to be done to validate these findings. In addition, Model 5 also consistently estimates the effects on poor consent attitudes. This model similarly shows a result that is in the same direction (a reduction) and magnitude (7.5 percentage points), but it is not statistically significant, and it is insufficiently powered. Similar to the knowledge index, there is a consistent indication that users demonstrate a reduction in poor consent attitudes, further studies would also benefit the ability for the research to conclude definitively on this outcome.

### **Subjective changes in knowledge and attitudes**

While all of the analysis thus far has focussed on observable changes in users' knowledge and attitudes, the endline survey, as well as the qualitative study, also gathered data on users' subjective reflections on how the platform may have affected their knowledge and attitudes. Due to restrictions in survey length, each question focused on a particular knowledge or attitude topic; condomisation and attitudes around sexual relationships in general. Table 7 reports the modal and second most frequent response to a number of self-reflective questions asked in the endline survey.

Table 7: Qualitative impact of platform from endline users' perspective

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Knowledge</i>				
Since YAL, change in your knowledge of condoms' importance	Likert: 5	Improved a lot (52%)	Improved a bit (23%)	489
Source that most influenced total knowledge about using condoms	Categorical: 8	YAL WA/FB (53%)	Other than YAL (47%)	375
<i>Attitudes</i>				
Source that most influenced attitudes around sexual relationships	Categorical: 7	YAL WA/FB (43%)	Friends or partner (14%)	502

A relative majority of users believe that in the past 5 months their knowledge of the importance of condoms has “improved a lot” (52%), with another quarter believing that their knowledge “improved a bit” (23%). In fact, only 2 percent of respondents believed that their knowledge had worsened, either by a little or by a lot. As such the qualitative data supports the findings of the knowledge tests, showing increases in self-perceived knowledge outcomes. Additionally, of those 375 users who report an improvement in their knowledge, 53% identify YAL as the source that has most influenced their knowledge about the importance of condom usage. Note that the question does not ask about the source of the change in their knowledge but rather the main source of their overall condom use knowledge. That after only 5 months, YAL serves as the primary source of many endline users' knowledge about the importance of condoms is a relatively powerful finding.

From the qualitative study, when looking at participant feedback on how the platform improved their knowledge on various topics, participants stated that they gained new knowledge and understanding about mental health issues, sexual health, and contraception methods. Other participants said that the topic of sexual health covered on the platform gave them more information about sex, demystified some of the knowledge they had, and even learned more about prevention methods. For example, one of the younger female participants shared,

**“I was curious about contraceptives but was scared to ask the people around me. I was worried they might judge me because of my age; hence, I used the App to get the needed information. I have not started engaging in sexual activities but wanted the information for when I am ready to.”**

**(Female, 16 Years old)**

While this is not a generalizable finding given the unique differences of users under 18 years old and above, it demonstrates that the B-Wise chatbot can be a place for

younger users to ask questions that might feel too difficult to ask of their peers, family, or health care workers. Another participant indicated that she learned a lot about sex and relationships, and it helped her to do an introspection on the things she was doing in her relationship. She stated that the platform helped her to live with a positive outlook on life and improved communication and sex between her and her partner. Some of the female participants indicated that they felt they were now able to negotiate for safer sex in their relationships and have a better understanding of what consent for sexual intercourse looks like for them.

## Outcome 1.1 - Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships

### Demonstrable changes in SRH behaviours

An underpinning of the YAL ToC is that improvements in knowledge and attitudes will also lead to improved SRH behaviours. The first target related to this desired outcome is SMART goal 7a, or, a 10% increase in the proportion of users performing each of the healthy SRH behaviours (separately) and a 10% decrease in the proportion of users with low self-perceived healthcare from baseline to endline<sup>11</sup>. Similarly to the changes in knowledge and attitudes above, Columns 1 and 2 of Table 8 present the coefficient on time for the simple and fully specified mixed model of time on each behavioural outcome of interest, with column 3 and 4 reporting the post-hoc ICC and associated power for each outcome of interest under model 2.

Table 8: Coefficient on period for all behaviours from mixed model regressions

	Mixed model regressions			
Controls included in model specification	1	2	2	2
Time invariant controls	-	X	X	X
Urban and province dummies	-	-	-	-
Interaction of time and income	-	X	X	X
<i>Behaviours</i>	<u>Coefficient on time period</u>		<u>ICC</u>	<u>Post-hoc Power</u>
Used condom	0.091***	0.080*	0.39	0.891
Used contraception	0.047*	0.032	0.29	0.397
Ever tested STI	0.049**	0.038	0.39	0.671

<sup>11</sup> Note, this wording is slightly different from that in the logical framework which stipulates a 10% increase in total healthy behaviours. Given the shared interest between both EJAF and Reach to understand changes in the individual behaviours of interest we have taken the liberty of analysing each separately.

1 or fewer partners	-0.008	-0.043*	0.22	0.081
<i>Self-perceived healthcare</i>				
Poor self-perceived healthcare	-0.068**	-0.109**	0.26	0.528
Sample size (range across outcomes)	467-502	386-424	386-424	386-424
<i>Sub-sample analysis</i>				
Used condom (given no plan for children in the next year)	0.102***	0.106**	0.37	0.771
Used contraception (given no plan for children in the next year)	0.052	0.052	0.26	0.418
Sample size (range across outcomes)	341	283-286	283-286	283-286

Once moving to the preferred form of the model, column 2 shows that only 2 outcomes are statistically significant at the 5 percent level, once controlling for all time-invariant variables (other than location) and including an interaction term between time and income per capita. The results indicated that the proportion of respondents with low self-perceived healthcare decreased by 11 percentage points, or 22% relative to the baseline proportion ( $p < 0.05$ ). However, column 4 shows that this finding only has a power score of 53%, as such while this indicates significant improvements for the sample, a more highly powered sample is needed to draw larger inferences to other users like the endline respondents.

The result for all users' condomisation is adequately powered (at 89%), but is only statistically distinguishable from zero at the 10 percent level. Therefore, there is less compelling evidence of changes in condomisation for all endline respondents. While this full sample is interesting, the analysis of contraception and condomisation should rather be restricted to those users that are not interested in having a child in the near future, as other users may be purposefully avoiding contraceptive methods. For the sub-sample of users that indicated that they specifically did not plan on having a child within the next year, column 4 shows that the coefficient on condom usage remains large and significant under the fully specified model. It would indicate that, all else being equal, the proportion of respondents using a condom at the last sexual encounter, increases by 10.6 percentage points ( $p < 0.05$ ) and exceeds the target of a 4 percentage point change, for endline users not looking to have a child in the next year. Additionally, this outcome is powered to detect a change as large as was observed 77% of the time, close to the rule of thumb marker of 80%. As such we conclude that this change is substantial, significant, and adequately powered, giving us sufficient evidence to believe that similar changes would have been seen for other users similar to the endline sample and who do not plan on having a child in the next year, all else being equal.

Given the high rates of unintended pregnancies, discussed in the knowledge section, it is unfortunate that there is no observed change in the proportion of users using contraceptive methods in the sub-sample of users uninterested in having children. That there is no increase in contraception, despite a concomitant increase

in condomisation, may support the finding from the knowledge section of a substitution towards effective contraceptive methods, rather than an overall increase in the use of contraception for the endline population. Given that 23% of South African women who become pregnant were using some form of contraceptive method when they became pregnant (Chersich et al., 2017), this is a meaningful problem for this population.

When comparing these results from the Model 4 regression to those that resulted from using the Model 5 regression when it was found to be consistent, there are some differences worth considering. For the outcomes on contraception, both for the total endline group and for the sub-group not intending to have children in the near future, Model 5 shows reductions in the use of contraception (2.2 and .1 percentage points less respectively), however neither of those results are statistically significant or sufficiently powered. For all other outcomes discussed in this section, Model 5 estimates were found to be inconsistent, thus we rely on the Model 4 estimates.

### Subjective changes in SRH behaviours

In addition to estimating changes in users' self-reported behaviours from baseline to endline in a pre-post manner, the endline survey and qualitative study were also interested in gathering users' feedback on their own subjective perception of the impact of the platform on their SRH behaviours. Table 9 reports the modal and second most frequent response to a number of self-reflective questions asked in the endline survey.

Table 9: Impact on behaviours from endline users' perspective

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Behaviour: Condomisation</i>				
Do you plan to use condoms more consistently than 5 months ago?	Likert: 5	A lot more (67%)	A little more (21%)	356 <sup>12</sup>
Do you plan to test for STIs more consistently than 5 months ago?	Likert: 5	A lot more (64%)	A little more (23%)	492
Source that has most influenced plans to use condoms or test for STIs (given increase in plans)	Categorical: 8	YAL WA/FB (47%)	Other than YAL (53%)	353 <sup>13</sup>

Analysing user feedback, 67% and 21% of users indicated that they plan to use a condom "a lot more" and "a little more" consistently than they did 5 months ago when registering with YAL<sup>14</sup>. Therefore, Table 9 would indicate a substantial increase in users'

<sup>12</sup> Once excluding users that indicate they either might (64) or are planning to have a child (64) this year.

<sup>13</sup> Restricting users to just those that indicated an increase in their intention to use condoms or test for STIs

<sup>14</sup> Once excluding users that indicated that they might or will have a child within the next year.



intention to regularly use condoms, for users similar to those completing endline and not planning on having children. These increased intentions appear to then be translating into increased action, as evidenced for both the full and sub-sample of respondents in Table 8 above. Additionally, 64% and 23% of all endline respondents also indicate that they plan to test for STIs and HIV “a lot more” or “a little more” than they did when registering with the platform 5 months ago. When paired with the findings from the above table that “ever testing for an STI” hadn't significantly changed from baseline to endline, this self-reported intention to test improving may reflect a recent change in intention (perhaps meaning the user hadn't yet had time since the change in intention), or that even with increased intention, systemic barriers exist for these users that don't allow intention to translate into action.

Finally, the endline survey also asked respondents to identify the primary source that influenced their “plans to use condoms or test for STIs/HIV”. Note again that this question is ambiguous and unfortunately does not specifically ask about the source of users’ change in condomisation/testing plans, but rather could be interpreted as the most important source affecting their views overall. Restricting this question to just those 353 users that indicated an increased intention to condomise/test, 47% of all respondents identified the B-Wise WhatsApp chatbot or Facebook page as the primary influence on their behavioural intentions. As before, the ambiguous interpretation of the question makes it difficult to conclude on which proportion of users are identifying the platform as the primary influence of their behaviours, and what proportion are identifying the platform was primarily responsible for this increase in intention.

Interestingly, several participants in the qualitative study shared that they preferred the B-Wise chatbot to traditional healthcare services for several reasons. For example, participants reported still struggling to get information about sensitive topics from their local healthcare facilities. Specifically, young women reported being judged for wanting information on family planning. Finally, most participants across gender lines and age groups identified the lack of a safe space and fear of judgement as primary motivators for using the App. For example, one female participant stated,

**“I would not feel that much comfortable talking to a nurse or face-to-face, so the bot is much easier because you just type in the message and it responds”**

**[Female participant, focus group discussion]**

**“.... they usually judge you. You can ask questions about sex with the Chatbot, but you can never touch some of these subjects with people.... because people are way too judgemental”**

**[Female participant, focus group discussion]**

In addition, the qualitative study also surfaced other positive changes in SRH behaviours such as increased comfort and confidence when discussing sensitive topics such as mental health, sexuality, and relationships, improved communication within their relationships about sexual health issues or in communicating their needs and

boundaries, and improved perceptions regarding contraception such as stating they'd be more proactive in seeking contraceptive services after engaging with the platform.

## Outcome 1.2 - Improved uptake of clinical health services related to sexual and reproductive health

Based on the YAL TOC, the logical framework included SMART goal 7b: *2.5% of WhatsApp users indicate they are more likely to visit a clinic or other health facility for their sexual needs since being on B-Wise*. Thus, for this section, the baseline and endline survey asked users several subjective and behavioural questions about their SRH, mental health needs, and health-seeking behaviours. Table 10 reports the modal and second most frequent response to these self-reflective questions.

Table 10: Impact on clinic linkages from endline users' perspective

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Behaviour: Linkages to healthcare facilities</i>				
Needed healthcare facility for SRH or MH	Binary	Yes (64%)	No (36%)	494
Number of visits to a healthcare facility, given was in need for SRH/MH	Continuous	Mean = 3.3	-	138 <sup>15</sup>
YAL has changed how likely you are to visit a health facility, for SRH/MH needs	Likert: 5	A lot more likely (46%)	A little more likely (36%)	199 <sup>16</sup>

Of the endline sample, 64% of users experienced at least one sexual or mental health concern that they felt required visiting a healthcare facility. On average, those users in need of healthcare assistance visited a healthcare facility 3.3 times while registered on B-Wise. The modal response was 2 visits across the period (21%), and only 13 respondents (10%) did not visit any healthcare facility despite being in need. Of those respondents who did seek care, 80% indicated that they were seen and treated, while 13% were seen but received no diagnosis from their visit, and 5% were not seen despite visiting a facility<sup>17</sup>. This would indicate that, given the need for a healthcare facility, YAL users similar to the endline sample visit a clinic and receive treatment through doing so. Unfortunately, the baseline survey for the WhatsApp pre-post study did not include

<sup>15</sup> Although 316 users indicate needing to visit a sexual or mental healthcare facility across the period, this follow-up question about the use of healthcare facilities was only received by 138 in-need users taking the survey in its second release.

<sup>16</sup> This question was only received by those users in the second release of the endline survey.

<sup>17</sup> A table of this data was not generated but can be upon request.

questions about users' need for SRH/MH services and use of healthcare facilities<sup>18</sup>. As such, this report cannot directly estimate the endline sample's change in health-seeking behaviour. However, Table 10 shows that 82% of the endline respondents indicated that YAL had increased their likelihood of visiting a healthcare facility (with 46% of these saying they are a lot more likely to visit a facility). It is important to note, however, that these are self-reported measures and one should assume that there could be some social desirability bias involved in user's responses.

When thinking more specifically about mental health, this report looked at the findings for SMART goal 7c: *2.5% of WhatsApp users indicate they are more likely to speak to a counsellor about their mental or sexual health needs since being on B-Wise*. For this section, the endline survey asked endline users several subjective and behavioural questions about their needs to speak with a mental health counsellor. Table 11 reports the modal and second most frequent response to these self-reflective questions.

Table 11: Impact on counselling linkages from endline users' perspective

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Behaviour: Linkages to counselling services</i>				
Needed to speak to a counsellor about SRH or MH	Binary	No (52%)	Yes (48%)	499
B-Wise has changed how likely you are to speak to a counsellor about SRH/MH	Likert: 5	Lot more likely (39%)	Little more likely (37%)	191 <sup>19</sup>

Of the endline sample, 48% of users indicated that they felt they needed to speak to a counsellor since joining B-Wise. Again, unfortunately, the baseline survey for the WhatsApp pre-post study did not include questions about users' intention to use counselling services, and thus this report cannot speak to how endline users' need for the mental health services may have changed across the program. However, Table 11 shows that 76% of the endline respondents that were asked, indicated that YAL has increased their likelihood of speaking with a counsellor (with 51% of these saying they are a lot more likely to speak to a counsellor).

When integrating the findings from the qualitative study, this aspect of linking users to services becomes even more complicated. First, when asked about using the referrals, most AYPs indicated that they only linked up with the online counsellor, while those referred to the clinics did not go.

<sup>18</sup> Originally data on users' need for healthcare facilities and uptake of recommendations was going to be gathered through feedback surveys from users engagement with the service finder feature. However, with this feature being paused on the service, an alternate subjective strategy for estimating impact on healthcare seeking before was agreed upon by Reach and EJAF in September of 2023.

<sup>19</sup> This question was only received by those users taking the survey in its second release

Moreover, when the researchers inquired about the counselling referral offered by the WhatsApp chatbot, the AYPs indicated that they did not want a face-to-face interaction but rather someone they could talk to about their specific challenges when they needed it either virtually or over the phone (free call or call back option). This might relate to the earlier cited findings around a major benefit of this platform in terms of how it can increase knowledge and certain attitudes for youth populations (like body image and attitudes on consent); the platform provides a safe space for users to seek out information on sensitive topics, ask questions they feel may incite judgement from family or health care workers, and provide unbiased information rather than misinformation that may come from peers.

***"Like I said, this thing is a robot. So, it gives you direct and factual answers and even more..... People will tell you what they want, what they feel you want to hear and not what is it that they must tell you (i.e., facts)" [Female participant, FGD2]***

This additional analysis demonstrates that further thinking about how to ensure that the services that YAL might suggest to users are truly youth-friendly is needed, and that practitioners are well trained in the ideas of safe spaces and supportive care for youth.

## Activity 2: Provision of a Facebook platform with content and space for peer discussions

Level	Description	Indicators
Output 2.1	Users see content and peer discussions on contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	Total Number of impressions ( <i>number of times any content from the B-Wise Page or about the B-Wise Page entered a person's screen</i> )
Output 2.2	Users post their own concerns seeking peer support or input	Total number of engagements (total number of unique users performing an action on the page per month of use, either a post, share or comment)
Intermediate Outcome 2.1	Users find posted content relevant, interesting and useful	Percentage of users that report finding the content relevant, interesting and useful, Quantitative
Intermediate Outcome 2.2	Users find peer comments relevant, interesting and useful	Percentage of users that report finding the comments relevant, interesting and useful, Quantitative
Outcome 2.1	Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships	<p>Percentage difference in Facebook users who score high on total healthy behaviours in their 1st vs their 3rd month</p> <p>Percentage difference in Facebook users who score high on total self-perceived healthcare outcomes in their 1st vs their 3rd month</p> <p>Percentage of users that believe their knowledge and attitudes have increased since being on the platform</p>
Outcome 2.2	Improved uptake of mental and clinical health services related to sexual and reproductive health	<p>Percentage of Facebook users that indicate they are more likely to visit a clinic or other health facility for their sexual needs since being on B-Wise for 3 months</p> <p>Percentage of Facebook users that indicate they are more likely to visit a counsellor clinic or other health facility for their sexual needs since being on B-Wise for 3 months</p>

To understand if the B-Wise platform effectively provided content in a different online setting and a space for peer discussion to users, the research team analysed data from the relevant results framework indicators shown above as well as specific questions that were included in the Facebook survey and qualitative study. From October 2022, B-Wise Facebook posts covered various topics related to SRH and mental well-being, highlighting the implementation of program activities. As highlighted in the limitation sections above, most of the content was posted between October 2022 and June 2023.

**Output 2.1 - Users see content and peer discussions on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships**

Regarding the audience reached through the B-Wise Facebook component, the page had 29, 242 Facebook followers as of 30 November 2023. Approximately 56% of these followers were women, with women aged 18-24 constituting the largest proportion of the total audience. The paid content<sup>20</sup> reached 9,278,931 people, received 1,564,306 link clicks, and achieved a total of 157,597,280 impressions during the same period.

Table 12: Summary of Facebook page metrics

Metric	Definition	Advertised content
Reach	The total number of people who see your content.	9,278,931
Link clicks	The number of clicks for any given link, whether on a landing page, webpage, or Facebook ad	1,564,306
Impressions	The number of times content is displayed, no matter if it was clicked or not	157,597,280

In order to properly assess user engagement with the different aspects of Facebook content, the team defined categories of exposure to content for the two kinds of users: Facebook group followers, and Facebook users who received paid content. For the Facebook group followers, exposure was categorised as high exposure (every day or a few times a week) and low exposure (every other week or at least once a month). For users who received paid Facebook content, users were categorised (self-reported frequency of exposure in the past week) as high exposure (receiving an ad 5 times and above), low exposure (1 to 4 times), or no exposure (0 times). The data demonstrated

<sup>20</sup> Paid content refers to the paid advertisements that were launched through the B-Wise partner, Avert. The purpose of these advertisements was to both drive people towards the B-Wise Facebook page, but also to act as another method of exposing youth users to appropriate SRH and mental health content.

that a quarter of B-Wise Facebook members were categorised as low exposure, while the majority, constituting 75%, were high exposure. For users engaging with B-wise paid content, 28% were categorised as having no exposure to the content in the past week, while 46% reported low exposure to paid content and 26% reported high exposure to content.

There were no statistically significant associations found between content exposure and user characteristics such as gender, household income, frequency of going hungry, relationship status, HIV status, and previous engagement with sexual and reproductive health topics. These results suggest that sociodemographic characteristics and sexual-related experiences did not influence exposure to B-Wise Facebook posts. The findings also suggest that diverse users were equally attracted to see content and peer discussions on contraceptives, sexual health, HIV and STI, mental health, sexuality, and healthy relationships on the Facebook platform.

As outlined in the TOC, the anticipation was that Facebook platforms would effectively reach the target audience, and the analytics confirmed that the Facebook component did generate the opportunity for users to view SRH content and other users' perspectives on SRH themes. The Facebook component achieved substantial reach, further boosted by additional support from paid social media advertising.

## **Output 2.2: Users post their own concerns seeking peer support or input**

During a 13-month period, B-Wise's paid content was able to elicit a high level of audience engagement from its platform. A total of 8,930,656 post engagements were recorded, which included 245,786 reactions, 3,024 comments/replies, 2,492 shares, and 3,666 saves. Table 13 summarises the most engaging Facebook posts, including the top five posts with a primary educational purpose. The engagement rate for these posts ranged from 7% to 17%.

Table 13: Summary of most engaged topics on Facebook page

Post topic (post type)	Impressions	Reach	Engagement <sup>21</sup>
How do I talk to my boyfriend about using condoms?	196,853	84,063	3,238
Do I still have to ask for consent in a relationship?	151,285	60,617	2,588
Ready for sex - video	384,660	143,264	55,06
Red flags - video	524,837	186,112	90,391
Mental Health Day - video	423,314	254,709	61,392

<sup>21</sup> Examples of engagement- interaction can include actions such as likes, comments, shares, clicks on links, and reactions.

Based on the results of a Facebook survey, it was found that among Facebook users, 41% never shared any posts, and 36% never participated in commenting. A significant percentage of users, 41%, moderately engaged by sharing and commenting 1 to 5 times. Only a smaller percentage of users, 11%, were more actively involved, commenting 6 to 10 times, while 14% commented more than 10 times. For paid content users, 26% never shared any posts, and 36% never commented. The majority of users shared or commented in the range of 1 to 5 times (43% for sharing and 36% for commenting). A smaller fraction of users, 13%, shared posts or questions, and 11% commented on threads six times or more. Table 14 below summarises these findings.

Table 14: Content Engagement Among Facebook Members and Advertised Content Users

	<u>Facebook Members</u>			<u>Facebook advertised content</u>		
	Response	Count	%	Response	Count	%
Frequency: shared a post or question on the B-wise Facebook page about a view or question that you wanted people's input on	Never	42	41	Never	20	26
	1-5 times	44	43	1 - 5 times	43	57
	6-10 times	8	8	6 and above	13	17
	More than 10 times	8	8			
Frequency B-wise Facebook page do you think you have ever commented on	Never	36	35	Never	29	38
	1-5 times	41	40	1 - 5 times	36	47
	6-10 times	11	11	6 and above	11	14
	More than 10 times	14	14			

The rate of sharing and commenting among Facebook users reflects higher levels of mild engagement (sharing a post between 1 and 5 times) and a few instances of high engagement (6+ times) for a smaller number of users, however there is a significant percentage of survey respondents who cite never engaging. Given this analysis of user interaction with the platform, there is no strong indication that the Facebook arm effectively fosters engagement beyond mere exposure. However, when reviewing the literature, previous studies reporting positive program effects have shown engagement rates of around 13% (Kotze et al, 2020), suggesting that the observed participation patterns here, with notable percentages of users actively sharing posts and participating in commenting, align with social media engagement. These findings highlight the platform's opportunity to facilitate meaningful engagement with its



content. Still, more effort must be made to motivate users who engage with the Facebook page at least once to engage more consistently. If the platform can motivate higher engagement, this could provide the Facebook component the opportunity to positively affect young people’s SRH and mental health needs.

**Intermediate Outcome 2.1: Users find posted content or comments relevant, interesting, and useful**

This report uses the data collected through the survey responses from both the Facebook group follower audience as well as the Facebook paid advertising audience to understand the relevance, interest, and usefulness of the content across both audiences. The majority of Facebook users, including both members and paid content consumers, have expressed favourable opinions regarding the relevance, interest, and utility of posts and comments that pertain to sexual health as measured by a set of Likert scale questions. Specifically, a significant proportion of both groups, 75.2% of Facebook members, and 69.3% of paid content consumers, “strongly agreed” that the posts were informative, interesting, and pertinent to their sexual health needs. Furthermore, a considerable number of additional participants in each group expressed “agreement with this assertion”.

Table 15 : Perceived usefulness and relevance of B-wise content

Response	<u>Facebook Member</u>		<u>Paid advert content user</u>		<u>Total</u>		
	Count	%	Count	%	Count	%	
B-Wise relevant, interesting and were useful	Strongly agree	76	75,2	52	69,3	128	72,7
	Agree	21	20,8	19	25,3	40	22,7
	Not sure	4	4,0	3	4,0	7	4,0
	Strongly disagree	0	0,0	1	1,3	1	0,6
B-wise Users' comments relevant, interesting and were useful	Strongly agree	40	39,6	26	34,7	66	37,5
	Agree	45	44,6	37	49,3	82	46,6
	Not sure	15	14,9	10	13,3	25	14,2
	Disagree	1	1,0	2	2,7	3	1,7

**Intermediate outcome 2.2: Users find peer comments relevant, interesting and useful**

Similarly, when evaluating comments made by other users on the B-wise Facebook page, a comparable trend emerged. “Strong agreement” and “agreement” were commonly observed, with 39.6% of Facebook members and 34.7% of paid content consumers strongly agreeing that the comments made by other users were relevant and useful for their sexual health. Taken together, these findings suggest that the content provided on the B-wise Facebook page is perceived positively. Based on this evidence, it can be concluded that the intermediate outcome of the program has been achieved.

**Outcome 2.1: Improved behaviours regarding contraceptives, sexual health, HIV and STI, mental health, sexuality and healthy relationships**

According to the survey data, the majority of participants reported significant improvements in their self-reported knowledge and intentions towards healthier behaviours. Specifically, more than 85% of the respondents reported an increase in their knowledge about condoms, as well as their intentions for consistent condom use, HIV testing frequency, and family planning. A considerable number of respondents attributed these positive changes to the influence of the B-wise platform. Moreover, more than 75% of the participants expressed an increase in their awareness of condoms, intentions for consistent condom use, HIV testing frequency, and changes in their attitudes towards sexual relationships and interactions after engaging with B-wise.

Table 16: Knowledge change and behaviour intention

	Knowledge change and intentions towards healthier behaviours		Participants Reporting Behaviour Change Due to B-wise	
	Count	N	Count	N
<i>Condom Knowledge</i>				
Yes, improved a lot	120	68.18	94	78,3
Yes, improved a little	39	22.16	26	66,7
Stayed the same	16	9.09	-	-
It's a little worse	1	0.57	-	-
<i>Condom use consistency intentions</i>				
Yes - a lot more	99	59.64	75	75,8
Yes - a little more	46	27.71	26	56,5
No - same	16	9.64	-	-
No - a little less	4	2.41	-	-

No - a lot less	1	0.60	-	-
<i>STI test intentions</i>				
Yes - a lot more	111	64.16	81	73,0
Yes - a little more	45	26.01	32	71,1
No - same	14	8.09	-	-
No - a lot less	3	1.73	-	-

Although the program had some influence on adolescents' knowledge, there is less clear evidence to support a direct relationship between exposure to the B-Wise Facebook page and key outcomes such as confidence in taking care of one's health, contraceptive use, condom use, exclusive sexual partners, and STI/HIV testing. This raises questions about the program's overall impact.

Table 17: Facebook page associations with behaviours

	Confidence in taking care of own health	Contraceptive use	Condom use	Exclusive sexual partners	STI/HIV testing	Condom use knowledge
<i>SRH exposure pre-B-Wise</i>						
Somewhat	-1.439 (-0.98)	1.894 (1.42)	0.017 (0.02)	-3.282* (-2.01)	1.733 (1.53)	0.329 (0.24)
Not much	0.543 (0.45)	-0.386 (-0.51)	-0.624 (-0.78)	-0.711 (-0.43)	0.393 (0.50)	1.298 (1.17)
Never	-1.615 (-0.87)	0.737 (0.51)	0.571 (0.44)	-3.490 (-1.47)	1.406 (0.92)	-0.081 (-0.06)
<i>Facebook Exposure</i>						
High exposure	-1.298 (-1.20)	0.987 (1.10)	0.158 (0.22)	0.310 (0.28)	0.819 (0.96)	-1.910 (-1.95)

## **Outcome 2.2: Improved uptake of mental and clinical health services related to sexual and reproductive health**

Regarding changes in the uptake of services, 86% of the participants reported that their exposure to the B-Wise component had improved their likelihood of visiting a clinic. This indicates a positive trend in the participants' intentions for behavioural change, which reflects the influence of the B-Wise Facebook component. The survey results suggest that the program had the potential to contribute to increased knowledge and a change in intentions through user engagement with the platform, but there simply isn't strong enough evidence at this point to state this more definitively. Similarly, there was no significant impact on behaviour change. Our analysis found that a smaller proportion of those who liked the content and found it relevant interacted through posts or comments, but there was no significant correlation between the amount of exposure to the program and behaviour change. This lack of evidence could be due to the prolonged period between program exposure and data collection. Previous studies have reported small but significant positive impacts of mobile health SRH promotion, particularly in enhancing knowledge, changing negative attitudes, and encouraging healthy sexual behaviours such as condom use, abstinence, and STI screening/follow-up. These effects appear consistent across age, gender, country, and intervention dose but are often short-term and not sustained beyond six months. Our inability to establish a link may have been mitigated if data collection had occurred within a few months post-content exposure. To comprehensively assess the program's long-term impact, it may be necessary to extend participant exposure and increase interaction periods, allowing for a more comprehensive observation of how sustained exposure influences adolescent behavioural changes. This should be noted, however, that the question of sustainability of changes in attitudes or behaviours is a research priority for any continuation of the Facebook component in future phases.

### Activity 3: Provision of a service finder tool for SRH and mental health services near to users

Level	Description	Indicators
Output 3	Users see and use service finder tool for SRH and mental health services	<p>Number and percentage of youth that start a service finder session</p> <p>Number and percentage of youth that complete a service finder flow (for the first time) relative to those that started a screening</p>
Intermediate outcome 3	Users find the service finder tool helpful	<p>Percentage of users that rate one of the recommended facilities as reasonably accessible and meets their needs</p> <p>Number and percentage of users that complete more than 1 service finder (given that they completed at least 1 - disaggregated by users review of the screening result)</p>
Outcome 3	Increased linkages to health care facilities	Number and percentage of users that report having sought the service 2 weeks after indicating that they found the finder helpful (vs those that have not yet but want to vs those that won't vs tried but couldn't)

**Output 3 - Users see and use service finder tool for SRH and mental health services**

Launched on 27th of October 2022 with the B-Wise chatbot, the service finder tool allowed users to identify a particular health need and receive a list of the relevant services with approximate distances to each facility. For this component, the team identified SMART goals 15 and 16: *40,000 (40%) users start a service finder session*, and *16,000 (40%) users complete a search*. Unfortunately, the feature had to be deactivated in June 2023 following a contract termination between the NDOH and their partner WitsRHI. The feature has remained deactivated since, though the partner organisation SoulCity has indicated their intention to reactivate the feature in the future. Between launch and June 2023, 1,167 users started a service finder session (3% of all users accessing the platform and 8% of all users subscribed for push messages up to the same date). Of these, 1136 (97%) then went on to complete a search. This demonstrates a relatively low uptake of the service even while it was active.

**Intermediate Outcome 3 - Users find the service finder tool helpful**

In order to determine if users found the service finder tool helpful, the team created SMART goal 18: *2,400 (15%) users complete more than 1 service finder search*. Recalling that the service finder tool was only active between October 2022 and June 2023, we find that 341 users completed multiple service finder searches (or 30% of those completing 1 search). As such, while the absolute number of intended users utilising the screening tool multiple times was not reached, the percentage of repeat searches was double what was initially targeted. This would indicate that for a third of those users that engaged with the service finder, the tool was sufficiently helpful to warrant a second use.

Building on this goal, the research team also created SMART goal 17: *50% of youth that reviewed the results of the service finder rate one of the recommended facilities as reasonably accessible and meeting their needs*. Unfortunately, when the service finder tool was deactivated, so too were the automated surveys that Reach had intended to use to gauge users' experience of the service finder feature. Moreover, due to survey length and concerns with overburdening respondents, no direct equivalent of this type of subjective experience question was asked in the recent WhatsApp endline survey. Nonetheless, some inference on the appropriateness of the service finders' recommendations can be made from a question about users' linkages to recommended clinics, outlined in Outcome 3 below.

**Outcome 3 - Increased linkages to healthcare facilities**

As mentioned above, with the deactivation of the service finder feature, all built-in feedback surveys were also deactivated, and this report cannot completely interpret contributions toward SMART goal 19: *2,500 (2.5%) users report having sought the service 2 weeks after indicating that they found the service finder helpful*. In order to

gather some understanding of users' perceptions of the service finder feature, additional platform review questions were added to the WhatsApp endline survey.

Of those users completing the WhatsApp endline survey, Table 18 reports the modal and second most frequent response to these self-reflective questions.

Table 18: Endline users' feedback on the service finder feature

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Sexual health seeking behaviour</i>				
Needed healthcare facility for SRH or MH	Binary	Yes (64%)	No (36%)	494
Know can use YAL to find clinic based on your needs	Binary	No (51%)	Yes (49%)	500
Used YAL to find clinic	Binary	Never used (62%)	Ever used (38%)	244
Ever visited a service that YAL recommended	Categorical: 4	Yes (71%)	No, too far (14%)	93 <sup>22</sup>

Before analysing users' sentiment of the service it is worth contextualising how the endline sample differs from the average platform user. As already mentioned in Section 3.2, the average endline user is disproportionately female and more likely to be in a relationship at registration. Additionally, Table 18 indicates that endline users appear substantially more engaged with the service finder. Of the endline sample, 49% of users know that YAL provided a means of finding appropriate nearby facilities, and of those, 38% ever used the service finder (ie. 19% of users in need). This is much higher than the 3% of all platform users and 8% of subscribed platform users mentioned above. Noting that engagement with the service finder was ~40% for endline users, once users already indicated knowing about the feature, indicates that even among the endline sample, knowledge of the tool could be improved. However, since the feature was deactivated early into users' journeys this is perhaps to be expected.

Of the 93 endline respondents that did report using the service finder, 71% indicated that they then visited the recommended healthcare facility (i.e. 14% of endline respondents who indicated needing healthcare, or, 9% of all endline respondents). Where the YAL TOC had aimed to have 2.5% of all users visit a recommended healthcare facility, this would indicate that for users similar to the endline respondents the service finder was quite successful in driving behaviour for those in need while active. As such identifying whether differences in uptake of the services is due to endline users having seemingly high knowledge of the service, or due to particularities of their sample is

<sup>22</sup> Restricting to the 93 users that used the clinic finder tool

unclear and would suggest further research would be helpful in the next phase of the project.



## Activity 4: Provision of depression/anxiety screening tool for mental health

Level	Description	Indicators
Output 4.1	Users use the depression screening tool	Number and percentage of youth that start a mental health screening Number and percentage of youth that complete a screen
Output 4.2	Users scoring high on screenings are informed of CallBack feature	Number and percentage of youth whose screen result directs them to LoveLife
Intermediate outcome 4.1	Users ask for the LoveLife help-desk to call them back	Number and percentage of youth that ask for LoveLife to call them back
Intermediate outcome 4.2	Users directed to mental health content finish the mental health content syllabus	Percentage users reading at least 60% of the messages if prescribed mental health content in their prescribed content buckets
Outcome 4.1	Improved attitudes and behaviours regarding mental health	Percent of users that score as high risk on their first depression screening that then score low risk at endline
Outcome 4.2	Short-term mitigation of serious psychological stressors for the youth experiencing depression	Percentage of users reporting that they requested a call-back service that actually received one in an automated follow up message Percentage of users that report their call back service as being helpful where they received one

**Output 4 - Users use the depression screening tool**

While the WhatsApp chatbot was launched on 27 October 2022, the addition of onboarding assessments, including the PHQ-4 assessment of depression and/or anxiety, was added later in January of 2023 under the YAL version 2 release. Here, all users who subscribed to regular push messages would receive an invitation to complete the mental health screening within their first week of registration. At this point, SMART goals 20 and 21 were instituted. These goals were that *22,500 (90%) of the users who have subscribed to push messaging start a mental health screening, and, 18,000 (80%) of those users complete the mental health screening.*

Of the 38,825 users that subscribed for outbound messages and would therefore have received the invite to take the PHQ-4 assessment, 11,021 (28%) started a screening which was well below both the targeted number and percentage of users. The mental health screening is the first assessment that all subscribed users are invited to complete, with the invitation to this assessment occurring on users' 2nd day on the platform. This may indicate that the platform needs to develop means of making this invitation more appealing. However, given the large observed drop-off in user engagement as noted in Activity 1, with subscribed users responding to 26% of their first two weeks of subscribed content, it seems possible that the low mental health screening rate is a symptom of the high early attrition rates, common to most m-health programs. As such Reach should focus on improving user retention through the first few days of the program. However, it's also likely that invitations to the mental health screening would need to be sent earlier in the user journey, given the realities of attrition on m-health programs (see Activity 1).

Of these 11,021 users, 10,276 (93%) then complete their screenings. This completion rate is substantially higher than what initially targeted within SMART goal 21 and may indicate that the screening tool is of an appropriate length and is sufficiently engaging for users who do start the screening to complete it.

Building from there, the platform intended not just to support mental health screening but also to direct those users who do show signs of needing mental health support to an appropriate service. In order to ensure that the service was reaching users in need, the research team created SMART goal 22: *9,000 (50%) of those users who complete a mental health screening are directed to LoveLife.* Given that the B-Wise chatbot automatically recommends the LoveLife callback feature to all users screening at-risk of depression/anxiety, this amounts to targeting the number of users that would screen as at-risk for depression or anxiety. Of the 10,276 users that take an initial mental health screening, 7,370 (70%) received scores high enough to be considered "at risk" of depression and/or anxiety under the PHQ-4 assessment methodology (Kroenke et al., 2009), a rate which closely matches the rate seen in the endline sample. These users are automatically recommended for the LoveLife callback feature. This higher-than-anticipated rate of depression/anxiety means that despite fewer users

taking the screening than anticipated, the platform came relatively close to its target of recommending counselling services to 9,000 at-risk users.

**Intermediate Outcome 4.1 - Users ask for the LoveLife help-desk to call them back**

Once the platform directs users to the LoveLife feature for a call from a counsellor, it is up to the user to then engage with the callback feature. Therefore to understand the rate of engagement of the callback feature, SMART goal 23 was introduced: *900 (10%) of those users directed to LoveLife ask for the counselling service to call them back, using the built in LoveLife feature.* Where the logical framework had expected only 10% of users receiving a recommendation to place a call with LoveLife, platform engagement data shows that 26% (1,916) of those users that screened as “at-risk” and were directed to the LoveLife callback feature actually requested a call from LoveLife through the chatbot. Again, this higher-than-anticipated response rate means that, despite engagement rates with the screening tool being lower than targeted ( SMART goal 20) , the chatbot had more than double the targeted number of users reaching out to the counselling service (SMART goal 23). This suggests that, for those platform users that started a mental health screening, when given easy-to-access and confidential opportunities to seek support, youth are likely to utilise that kind of service.

In order to provide some estimate of users' sentiment of YAL's mental health support, the endline survey also gathered users' perceptions of the LoveLife feature. Before considering these responses, it is first important to contextualise the mental health needs of the endline sample. Noting, as before that the endline sample are disproportionately more female and more likely to be in a relationship at registration than the average platform user, Table 19 reports the incidence of mental health challenges, both as the percentage of endline respondents scoring at risk on the PHQ-4 assessment at registration and then at endline. It also reports the percentage of endline respondents that self-reported needing to see a counsellor over the same period. This shows that 14% of endline respondents scored as “at risk” on the PHQ-4 when they registered , 63% then scored “at-risk” on the PHQ-4 at endline, and 48% of the sample indicated that they believed they needed to speak to a counsellor at some point across the program.

Table 19: Endline users' feedback on the LoveLife callback feature

Variable	Variable description and total choices	Modal response or proportion of sample	2nd most frequent or proportion of sample	Total respondents (n)
<i>Need for mental health services</i>				
Prop. of scores PHQ-4 at baseline	Score	Not at risk (86%)	At risk (14%)	489

Prop. of scores on PHQ-4 at endline	Score	At risk (63%)	Not at risk (37%)	489
Self-reported need to speak to a counsellor about SRH or MH	Binary	No (52%)	Yes (48%)	499
<i>Knowledge and use of LoveLife given users' need for counselling</i>				
Knows of LoveLife feature (all endline respondents)	Binary	Yes (52%)	No (48%)	483
Knows of LoveLife feature (given at risk at baseline)	Binary	Yes (57%)	No (43%)	69
Knows of LoveLife feature (given at risk at endline)	Binary	Yes (54%)	No (46%)	317
Knows of LoveLife feature (given self-reported need)	Binary	Yes (55%)	No (45%)	238
Used YAL to request a callback from LoveLife (given knows of feature)	Binary	Never used (63%)	Ever used (37%)	251
Used to request a callback from LoveLife (given knows of feature and need at baseline, endline or self reported)	Categorical: 3	Never used (60%)	Ever used (40%)	401

With this need in mind, Table 19 then also presents users' self-reported knowledge and use of the LoveLife feature, given their various identified mental health needs. It shows that, despite their various identified mental health needs (either by the PHQ-4 or self-perception), only between 52%-57% of the respective groups that expressed some need for counselling knew about the LoveLife feature. As such, simply advertising the tool to users after an "at-risk" PHQ-4 assessment is not sufficient for generating awareness of the feature among those that need it, and the platform would benefit from creating greater awareness of the tool.

That all said, the final 2 rows of Table 19 indicate that, conditional on knowledge of the feature, 37% of endline respondents had requested a callback from LoveLife. As such, the same finding as was seen using the general platform data is corroborated under the somewhat different sample of endline respondents, finding that given users know about the feature, use of LoveLife is higher than was targeted under SMART goal 23, and that exposure to the tool should therefore be prioritised.

### **Intermediate Outcome 4.2 - Users directed to mental health content finish the mental health content syllabus**

In addition to providing users who are at risk with an opportunity to speak to a counsellor, the platform also shares mental health content to support them. SMART goal

24 set the following target: *25% of users read at least 60% of their prescribed mental health content given that they screened as high risk of depression/anxiety.*

As indicated in under activity 1, it is not possible to identify which content pieces users have been prescribed based on their barrier assessments, due to the many versions of the YAL program and the evolution in the set of prescribed content over the intervention period. This limitation has been communicated and agreed upon with EJAF and Reach.

Reach and EJAF agreed on SMART goal 29b as a next best proxy of user engagement with relevant content: *On average, users who screen as at risk of depression/anxiety send messages to the line on 15% of the days that they receive push-messages (for all opted in users active 2 days post registration).* From the launch of the program until the end of November 2023, we find that, on average, those users that scored as “at-risk” on the depression/anxiety screening tool when invited to screen, engage on 55% of the days that they receive push messages, well exceeding the SMART goal. This engagement rate is also substantially higher than the engagement rate for all users. This may indicate that those platform users that are willing to take the mental health screening test differ meaningfully from those users that do not. However, this could also be seen as cursory evidence to support the idea that when users take assessments and have content prioritised in relation to their needs, we see large increases in their engagement rates.

## **Outcome 4.1 - Improved attitudes and behaviours regarding mental health**

Besides aiming to improve attitudes and knowledge regarding sexual health, the YAL platform also aimed to support users’ psychological capacity through content around mental health. Here, the platform targeted a *10% decrease in the proportion of users that score as high risk on their depression/anxiety screening from baseline to endline* (Smart Goal 25). As in Table 6a, Table 20 presents the results of 2 mixed-model specifications for the 502 endline respondents for each psychological capacity barrier of interest separately; i) the proportion of users that screen as depressed and/or anxious on the PHQ-4, ii) the proportion screening as misusing substances Reach’s constructed indicator, and iii) the proportion of users who never/sometimes have someone “to talk to when [facing a] worry or problem”. For ease of interpretation, only the coefficient of time (moving from baseline to endline) is reported, with column 1 reporting the coefficient and p-value on time for a simple regression of time (as a dummy variable) on each outcome of interest. Column 2 then reports the research teams’ preferred model specification including all motivated controls<sup>23</sup>, with coefficients

---

<sup>23</sup> The full list of time invariant variables are; age, gender, HIV status, incidence of weekly hunger (only available at endline), exposure to other SRH content pre YAL, baseline relationship status and log of household income per capita. For an explanation motivating Model 2’s goodness of fit relative to other specifications, please see Appendix 9A

colour-coded as before, and columns 3 and 4 reporting the post-hoc ICC and associated power for each outcome of interest under model 2.

Table 20: Coefficient on time for all psychological barriers from mixed model regressions

Controls included in model specification	Mixed model regressions			
	1	2	2	2
Time invariant controls	-	X	X	X
Urban and province dummies	-	-	-	-
Interaction of time and income	-	X	X	X
<i>Psychological Capacity</i>	<u>Coefficient on time period</u>		<u>ICC</u>	<u>Post-hoc Power</u>
Prop. With Dep/Anxiety	-0.119***	-0.088**	0.17	0.984
Prop. With Substance misuse	0.040*	0.016	0.34	0.407
Prop. With Low connectedness	-0.115***	-0.108***	0.31	0.989
Sample size (range across outcomes)	502	424	424	424

Column 1, shows that there is a large decrease in the proportion of users identified as depressed and/or anxious on the PHQ-4 ( $p < 0.01$ ), suggestive evidence of an increase in substance misuse ( $p < 0.1$ ), and a large decrease in the proportion of users with low social connectedness ( $p < 0.01$ ). Moving to the fully specified model the time coefficient would imply that the proportion of endline respondents with depression and/or anxiety, as well as those with low social connectedness, both improve. The intervention time period is associated with a 9 percentage point ( $p < 0.05$ ) reduction in depression and/or anxiety and an 11 percentage points ( $p < 0.01$ ) reduction in low connectedness, even when controlling for changes in income over time and all available time invariant controls (sans location variables). Given that 76% of endline respondents scored as depressed/anxious baseline, and 78% scored as low socially connected at baseline, this represents a larger than 10 percent decrease in both outcomes. Column 4 indicates that the sample is more than sufficiently powered to detect changes as large as those observed, indicating that this decrease in depression/anxiety and low social connectedness can be extrapolated to users similar to those completing the endline. No statistically significant change in the proportion of users with substance misuse is observed in the fully specified model. It is noted that this result would be under-powered to detect the observed effect size if there were a statistically significant

change, however, this study is only concerned with detecting changes of 10% and larger, as mentioned previously.

When comparing these results from the Model 4 regression to those that resulted from using the Model 5 regression when it was found to be consistent, Model 5 finds a similar result for depression and anxiety demonstrating a reduction in depression/anxiety of 7.4 percentage points and is sufficiently powered at the 92.7% level, though it is not statistically significant. Given it is of a similar magnitude and direction, the downward trend is consistent, but since we lose significance in Model 5 that is consistent for predicting this outcome, more research should be done to confirm the statistically significant result from Model 4. For social connectedness however, Model 5 estimates a moderately lower reduction of 6.7 percentage points. This finding, though, is not statistically significant but powered at the 76.8% level, which leaves this result up for further discussion. Given that Model 5 is consistently able to estimate results during the Monte Carlo simulations and is sufficiently powered to detect a change of the estimated size, the research team believes this renders the results for depression/anxiety and social connectedness in need of further validation. For substance misuse, Model 5's estimate is similar to Model 4 in that it is not statistically significant nor sufficiently powered, therefore there is no effect detected at this time regarding substance misuse.

As a platform designed to assist users with managing their mental health, providing linkages to verified mental health resources and encouraging healthier relationships, it is encouraging to see improvements in these two primary psychological capability outcomes over the intervention period..

While the pre-post nature of the study means readers should be cautious in drawing any causal claims, it is further encouraging to see similar sentiments expressed in the qualitative research. From the FGDs and IDIs, participants reported feeling heard, understood, and supported when discussing mental health issues on the platform. Participants also felt less alone and more hopeful about their future. One of the key results of the perceived impact of the platform cited in the qualitative study is having the vocabulary to talk about mental health and related illnesses. Some of the participants mentioned that they had not been officially diagnosed; however, they perceived themselves as having some kind of mental health illness. They reported that since they started using the platform, they have learnt about the different types of mental illnesses and have a safe space to go to when they need to talk.

**“Not only do I know what I am suffering from by asking the Chatbot questions, I know what the people around me could be struggling with. I know now that people are not crazy when they are depressed”**

**[Female participant, Individual interview]**

They also noted improvements in their mental well-being and reported feeling less alone and more hopeful about their future.

**“I have been using it since September, and it is very good. I trust it. I have mental health issues, and since I have been using this App, I can see progress. The Chatbot is helpful”**

**[Female participant, focus group discussion]**

## **Outcome 4.2 -Short-term mitigation of serious psychological stressors for the youth experiencing depression**

Originally Reach had planned to monitor the efficacy of the LoveLife feature through automated follow-up surveys that would be sent to all users that engaged with the LoveLife feature. These surveys would be used to report on SMART goals 26 and 27: *60% of users reporting that they requested a call-back service that actually received one in an automated follow up message, and, 70% of these users report their call back service as being helpful where they received one.* Unfortunately, these surveys had to be deactivated along with all other automated feedback surveys on the platform, due to concerns about the line’s quality rating at the time. Once the line’s quality rating was high again, subsequent attempts to restart automated feedback surveys were made, but appeared to be associated with a proportion of users blocking the line. It is not clear that automated surveys themselves are causing these dips in line ratings, however future rounds of YAL would benefit from experimenting with multiple versions of the automated feedback surveys, in order to find viable options.

While the survey was active, 291<sup>24</sup> automated surveys were answered. These showed that, for those users that answered the feedback surveys, 100% indicated receiving a callback from LoveLife. And of those 64% found the callback “helpful”.

### **Impact - Improved SRH, mental health, health empowerment, and health persistence for an increasing number and percentage of adolescents and young people reached with the YAL platform.**

Finally then we turn to considering the cumulative effects of changes in all the SRH barriers and behaviours for the TOC’s targeted impact of increasing psychological persistence. As detailed in Activities 1-4, the endline sample demonstrates significant improvements to individual SRH knowledge, SRH attitudes (body image and valuation of consent), mental health (depression/anxiety and social connectedness), and SRH behaviour (condomisation for those respondents not intending to have children in the near future). Following on from this, the YAL TOC would expect to see increases in individuals psychological persistence scores.

---

<sup>24</sup> See the YAL MVP report Dec 2022



As noted previously, the endline sample differs from the average platform user in some key respects, being more likely to be female, be in a relationship at registration and face some key SRH barriers at registration. Therefore, while findings regarding the endline cannot be extrapolated to all users, the endline serves as the only data source by which to investigate changes in persistence while on the platform. Additionally, from a theoretical perspective, demonstrating changes in persistence with at least one relevant sub-sample of the target population, would serve as some useful evidence of YAL's proposed model.

As before, Table 21 below reports the coefficient on time for both the simple and fully specified mixed model regressions of time on the persistence measures of interest (locus of control and self-esteem), with columns 3 and 4 reporting the ICC and post-hoc power calculation of the observed changes in persistence for the 502 endline respondents.

Table 21: Coefficient on time for psychological persistence outcomes

Controls included in model specification	Mixed model regressions			
	1	2	2	2
Time invariant controls	-	X	X	X
Urban and province dummies	-	-	-	-
Interaction of time and income	-	X	X	X
<i>Psychological Persistence</i>	<u>Coefficient on time dummy</u>		<u>ICC</u>	<u>Power</u>
Prop. With External Locus of Control	-0.141***	-0.132***	0.17	0.998
Prop. With Low Self-Esteem	-0.056**	-0.018	0.19	0.677
Sample size (range across outcomes)	502	424	424	424

This shows that the proportion of endline respondents with a predominantly external locus of control decreases significantly in both the simple and fully specified models. The coefficient on time under the fully specified model indicates that even when controlling for baseline demographics and changes in household income over time, the proportion of users with external loci of control decreases by 13.2 percentage points (well above the targeted 10% / 6 percentage point on the sample's baseline rates.) Additionally, column 4 shows that for these large observed changes, the sample is more than sufficiently powered to detect changes as large as this, indicating that the result can be taken as representative of other users similar to those completing the endline survey. While column 1 also finds evidence of a statistically relevant decrease in the proportion of endline respondents with low self-esteem scores, this disappears once accounting for additional controls in column 2.

When comparing these results from the Model 4 regression to those that resulted from using the Model 5 regression when it was found to be consistent, one can see that the results for locus of control are supported given that Model 5 sees a similar 11.4 percentage point decrease in external locus of control. This result was, however, only significant at the 10% level and powered at 72.3%. It therefore agrees in terms of the general direction and relative magnitude of Model 4's result, however it is less certain statistically speaking. Further analysis in the next phase of the project should be done to substantiate the Model 4 results. For low self-esteem, Model 5 similarly finds a result that is a reduction in low self-esteem, but it is not statistically significant and only powered at 70.9%. So while it is encouraging that both models demonstrate a trend downwards in the proportion of users with low self-esteem, it is not confirmed.

## 5) Discussion

### 5.1) Analysis of TOC Validation

As a general introduction to this TOC validation, given the results discussed in the earlier section on the representativeness of the sample of users studied in the different evaluation activities, the findings reported here represent the potential effectiveness of the proposed YAL ToC for users of a significant engagement level (particularly for the findings that are substantiated through the endline survey). Uptake of individual features may have been low (for instance, service finder, etc.), but conditional on engagement, many SMART goals were met in line with the TOC. While this doesn't support the TOC for all users, there's some support for elements of the TOC, conditional on user engagement, which provide motivation that the pilot phase of the YAL program shows promising potential for future phases. However, it is still important to show how the mechanistic pathways do or don't hold true with this more highly engaged subsample, as that population presents the best opportunity for the mechanisms to work. Refinement of the ToC is needed for the next phase of YAL to incorporate this important assumption on reaching higher levels of engagement for outcomes to be achieved.

**TOC Pathway 1 (Activity 1): provision of in-depth content on sexual health, mental health, and healthy relationships, through a WhatsApp-based platform, leads to changes in attitudes, knowledge, behaviour, and ultimately, increased uptake of services.**

Based on the review of evidence, the YAL program has shown preliminary correlations between the provision of SRH and mental health content through the B-Wise chatbot and associated improvements in knowledge and a subset of attitudes (body image and valuation of consent in sexual relationships). The platform was successful in its efforts to reach 100,000 users. The evidence demonstrates that the demographic profile of these users is generally representative of the audiences that YAL was hoping to support, with some statistically significant differences regarding an over-inclusion of the 18-24 age range due to recruitment restrictions early in the program and over-representation of disadvantaged youth as specified by self-reported household income which may be due to measurement error in this self-reported indicator. Given that most respondents across all three studies share that they find the content relevant and useful to them, the report supports the hypothesis that the content YAL provided engages its audience. Thus, users are reading the content and interacting with it sufficiently. The TOC assumes, then, that if we're engaging with the right audience and the content is relevant to them so that they read it, the platform will support improvements in knowledge, attitudes, and behaviours. The evidence partially supports this claim regarding the statistically significant improvements in SRH knowledge, reductions in poor body image and poor consent attitudes, and use of condoms (for users not planning to have children in the next year), however there are certain

attitudes or behaviours that will need further evaluation in the next phases. Moreover, given that for certain of the significant outcomes found under Model 4 lose significance when moving to the Model 5 with a restricted sample, these findings need to be confirmed in the next phase of implementation, ensuring that a larger sample is used to confirm these findings with sufficient power and statistical significance. Qualitatively speaking, participants themselves cite that one of the primary contributions to their changes in knowledge, attitudes, and behaviour is either the B-Wise chatbot or the Facebook page. The user feedback indicates an increase in the intention to use condoms and test for STIs more consistently than they did when they first registered with the platform, though as stated previously, this is specific to engaged users and not the broader platform population. Additionally, the qualitative study revealed other positive changes in SRH behaviours, such as increased comfort and confidence when discussing sensitive topics such as mental health, sexuality, and relationships, and improved communication within their relationships.

On the other hand, the linkages between these changes in knowledge, attitudes, and behaviours to demonstrated uptake in clinical services are less clear. This is partly due to the limitations around the survey designs for baseline and endline and the loss of the Service Finder tool, which would have provided additional evidence to validate this aspect of the ToC. However, the data that was collected does demonstrate that 82% of the respondents that were asked indicated that YAL had increased their likelihood of visiting a healthcare facility (with 56% of these saying they are a lot more likely to visit a facility) and 76% of the respondents that were asked indicated that YAL has increased their likelihood of speaking with a counsellor. While these are users' intentions and not evidence of service utilisation, it is a useful finding to support the hypothesis that the YAL platform can help generate additional motivation and capacity for higher engaged platform users to increase their service utilisation.

**TOC Pathway 2 (Activity 2): Provision of a Facebook platform with content and peer discussions on sexual health, mental health, and healthy relationships leads to changes in attitudes, knowledge, behaviour, and ultimately, increased uptake of services.**

Firstly, the program successfully reached a large audience, with over 29,000 Facebook followers and over 9 million people reached through paid content. Additionally, the program generated a high level of user engagement, with over 8.9 million post engagements recorded over a 13-month period. This demonstrates the powerful reach of a social platform such as Facebook, particularly for the audience that YAL intends to support. When looking at user feedback from the Facebook study, a majority of Facebook group followers, 75%, self-report very high levels of engagement, on average accessing the page every day or a few times a week. Similarly, for users who engaged with paid content, 72% of those users reported low to high levels of exposure based on

how many ads they saw in the last week. Thus, the analytics suggest that the Facebook component can provide a significant opportunity for users to view SRH content and other users' perspectives on SRH themes, if those items exist on the Facebook page. However, the report finds that the prevalence of users sharing perspectives on the Facebook page wasn't extensive across survey respondents. For example, among Facebook users who engaged with the Facebook study, 41% never shared any posts, and 36% never participated in commenting, with only 14% of users commenting more than 10 times. This adds some questions to the discussion around the effectiveness of this component, given that the goal is to foster sustained peer discussion and additional clarification for users' questions regarding SRH and mental health issues. However, for users that did engage and interact, many perceived the content positively, with 75.2% of Facebook members and 69.3% of paid content consumers strongly agreeing that the posts were informative, interesting, and pertinent to their sexual health needs.

We also see that users reported significant improvements in their self-reported knowledge and intentions toward healthier behaviours, and 85% of those users attributed these changes to the B-Wise platform (though there could be some conflation on what respondents considered the "platform" potentially including the chatbot service as well). Although the program had some influence on adolescents' knowledge, we did not observe a direct relationship between exposure to the B-Wise Facebook page and key outcomes such as confidence in taking care of one's health, contraceptive use, condom use, exclusive sexual partners, and STI/HIV testing. Thus, this aspect of the ToC will need to be revisited and refined to understand how to foster those behavioural changes more effectively in the next phase.

**TOC Pathway 3 (Activity 3): Provision of in-depth content leads to changes in attitudes, knowledge, behaviour, and ultimately, increased uptake of services.**

Based on the goals set by the team, it appears that the project achieved some of its intended outputs and outcomes. For instance, even though the service finder tool was active for a limited time, the platform got very close to its target of 40% of users in need knowing that there is a tool to help them locate services. In addition, the service finder tool was proven to attract repeat searches from a significant proportion of its users, indicating that it was helpful for the audience of in-need, engaged users. Finally, most users who used the tool ended up visiting the healthcare facilities recommended by the tool, which implies that the tool successfully drove user behaviour among those in need, for the subsample of users who could access the tool while it was active and who are considered engaged users. Thus, given the representativity discussion, while the tool is found to be useful and relevant, more work would need to be done to advocate for the use of this tool with the broader YAL platform population to ensure more users are aware of the tool, know how to use it, and understand the benefits for their individual health and wellbeing. In addition, given the feedback shared within the qualitative focus groups, the next phase of the YAL program should incorporate efforts to validate that

services recommended on the platform are indeed youth-friendly and non-judgmental in order to support actual service uptake.

However, due to the deactivation of the service finder feature, the project was unable to fully measure the long-term intended impact of the tool on increasing linkages to healthcare facilities for a larger proportion of its users, nor did it allow this report to analyse users' experience of the service finder regarding the proximity/quality of the services that were recommended to them. Nonetheless, using the self-reported measures included in the endline survey, the translation of service recommendations to actual user uptake was higher than expected within the endline user population. Overall, the project made some notable progress in achieving this activity's goals, but further research and data collection would need to be built into a future phase to ensure a service finder feature remains active and that those critical service feedback surveys can be sent out and analysed to understand better the quality of services to which the platform is linking its users.

**TOC Pathway 4 (Activity 4): provision of a depression/anxiety screening tool for mental health and associated mental health content leads to changes in improved attitudes and behaviours regarding mental health, and ultimately, increased uptake of mental health services**

In January 2023, a mental health screening tool was added under the YAL version 2 release. The goal was for 90% of users who subscribed to regular push messages to start a mental health screening and 80% of those users to complete it. Of the 38,825 users who received the invite, only 28% started the screening, but 93% of those who started completed it. The platform aimed to direct at-risk users to a support service, and 70% of users who completed the screening received scores high enough to be considered "at risk." Despite not reaching the target for the number of users to start a screening, the platform came relatively close to its target of recommending counselling services to 9,000 at-risk users, given the high completion rate of the users that did start a screening. The goal for users to engage with the callback feature was 10%, but the engagement rate was 26% (1,916) for at-risk users directed to the LoveLife callback feature. This high rate of engagement could be reflective of the profile of users captured in the endline, which has been stated previously as being more highly engaged than the broader YAL platform population, however, it could also be demonstrating an effective approach of linking in-need users to a relevant and useful service. The goals of completing the mental health screening and directing at-risk users to support services were achieved. When looking at the analysis on whether this pathway supported improved attitudes and behaviours regarding mental health, the pre-/post-analysis showed that there were statistically significant improvements in both depression/anxiety (on the PHQ-4) and social connectedness under Model 4, but inconclusive results from the smaller sample captured in Model 5 which would indicate additional research is needed to confirm the impact of this intervention in a more robust sample. There is no change in substance misuse. There is also evidence from the

qualitative study that supports the idea that the B-Wise chatbot service may have helped users in better understanding mental health terminology, supporting them with their own mental health challenges, and providing them a safe space to talk about sensitive or challenging topics. Given these findings, it appears that exposure to the platform is associated with improved mental health outcomes for users like those captured in the pre-post and qualitative studies.

## 5.2) Overall Contribution Analysis Limitations

Overall, the structure of this contribution analysis has generated very useful findings to support improved implementation on YAL in the future. However, there were some limitations that, if addressed in the future, could support similar analyses to be even more enlightening.

First, there were unforeseen changes in the program design's implementation that the TOC did not account for (the loss of the service finder feature, the decline in Facebook content posting, and the lack of peer discussion moderation). To strengthen the outcomes of a program like YAL, implementation should consider how to structure those kinds of critical program activities so that there is more accountability and sustainability built in from the beginning. This is especially important given the promising findings that this analysis highlights for both components of the design.

Second, there were several limitations related to the implementation of the WhatsApp endline survey. For example, given the deviations in program implementation, proxy measures had to be included in the endline survey that were not included in the baseline. This results in the loss of some comparative pre-/post-analysis that would have been helpful in analysing the contribution of the program to observed achievements in those indicators.

Third, this evaluation was not built in a way that it could analyse changes in a group who only experienced version 1 of the YAL program as compared to users who experienced version 2. This is a missed opportunity where we weren't able to see what the additional value of the segmented/tailored approach is and how significantly the segmentation and tailoring improved (or hindered) the types of outcomes a program like YAL can achieve. The team is planning on identifying opportunities to do this kind of non-tailored and tailored program analysis in future programming related to YAL as well as other initiatives undertaken by Reach Digital Health.

## 6) Recommendations

This report identifies several potential areas for improvements in a future iteration of the YAL platform, as revealed through each of the studies conducted to support this report's analysis.

### **Reactivate the Service Finder Tool as soon as possible, or identify an updated offering for YAL users**

The analysis showed very promising results on how a tool like Service Finder can drastically improve youth healthcare service utilisation. Therefore, efforts should be made to deblock the challenges currently faced by the team to reactivate this service. If the tool as previously implemented cannot be revived, the YAL consortium should invest time and resources to design a similar offering so that the platform can continue to close the gap between service need and service utilisation. Moreover, a broader explanation into why certain users who used the Service Finder tool did not then go to the recommended facilities should be undertaken. For example, some users in the qualitative when asked about the lack of uptake of referrals, participants cited discomfort, lack of trust, and lack of confidentiality. The YAL consortium could consider how they might better identify services that meet these criteria and include them in the provider list on the platform.

### **Consider an additional phase of participatory focus groups with certain target audiences within the YAL population**

The report findings indicate in a few ways that more work could be done to strengthen the representation of minority groups (such as the under 17 years old or LGBTQ groups) in the YAL program through either recruitment strategies to ensure users from those groups register with the platform or additional content development through tailoring of existing content or the identification of supplemental content. In addition, anecdotal findings within the qualitative study suggest that additional content focused on supporting young men could be beneficial, especially given that the subpopulation of highly engaged users that participated in the endline had an overrepresentation of young women. If platform content were more tailored to the young male audience, perhaps the YAL platform could motivate higher levels of engagement from young male users.

### **Ensure that the YAL platform, as offered through WhatsApp and Facebook components, is as low-cost as possible and accessible to a diverse range of users**

Participants in the qualitative study cited that the Chatbot should use "free mode" so that more young people can access and use it irrespective of their financial status. Participants argued that free mode would allow youth users from different walks of life and socio-economic statuses to access the platform. Connectivity, particularly in rural areas, is a challenge. Thus, it would be great if B-Wise also had an offline option, even for a limited time, so everyone who needs it can access it. Using free mode with an offline



option reduces the exclusion of young people who cannot afford to buy data to access the platform.

### **Improving the social component of the platform**

The next phase of implementation for YAL should include a moderation mechanism for the Facebook channel to ensure regular posting of thematic content, fostering actual peer-to-peer discussion. This was the intended design for the existing platform, but as discussed above this did not occur to the degree desired. As highlighted in the qualitative study, there is desire and interest to have certain opportunities to discuss these themes within groups of peers. To make this effective, however, there needs to be a sufficiently resourced role with associated accountability to ensure that thematic content is posted regularly, and when users make the effort to post, someone is there to facilitate a discussion around those thoughts or feelings.

## **7) Conclusion**

Overall, this contribution analysis has demonstrated an immense amount of learning regarding the validity of the TOC and the effectiveness of the program's intervention for engaged YAL users. Two of the four pathways (pathways 1 and 4) within the TOC appear to be relatively well supported via the evidence generated across the three studies, at least for engaged users. The Facebook component and the linkage to services via the Service Finder tool were the least supported.

For Pathway 2 focusing on the Facebook component, this was not very well substantiated by the evidence due to a lack of peer discussion and Facebook post engagement from users of the Facebook page as well as a lack of evidence to connect self-reported intentional behaviour changes to actual behavioural outcomes. Given this finding, if there is a similar social component included in future iterations of the YAL program, more effort needs to be made to sustain peer-to-peer engagement and ensure the continued dissemination of relevant SRH and mental health thematic content to drive engagement. The qualitative findings did suggest that youth do value the opportunity for discussing SRH topics within a group, but similarly shared that there isn't always trust that information shared by peers is evidence based. This indicates that there is value in figuring out the proper way to motivate peer discussion while providing support to validate facts that are shared to ensure youth believe that these discussions are legitimate advice or knowledge.

For Pathway 3, there was a limited sample of people that reported finding the service finder helpful (during the period it was active on the platform) in addition to there being some self-reported data that people were somewhat more likely to seek SRH services. However, the existing evidence seems to support the hypothesis that facilitating linkages to care through a service finder feature or connecting a screening outcome to a callback service could increase service utilisation for already engaged users of the

platform. This would lead the research team to believe that additional investment in this branch of the ToC is worthwhile and continued adaptation to optimise this feature in a sustainable way (avoiding the roadblocks that were encountered in this iteration) should be prioritised. Given the findings, it appears that the WhatsApp chatbot contributed more significantly to the observed outcomes related to SRH and mental health knowledge, awareness, and persistence improvements, as well as the observed changes in SRH/Mental health behaviours, than the Facebook component (especially when considering the period after June 2023 when the Facebook page became significantly less active).

In addition to the validation of pathways, there are indications that certain content areas or behavioural nudges within the platform might need to be revisited to improve user exposure to certain themes around contraception and STI testing to reach the platform's desired targets. While there is quantitative evidence that indicates the content is useful and relevant to users and qualitative evidence that participants say they have intentions to take on these behaviours, more work needs to be done to understand the gap between the knowledge about and intention to adopt healthier behaviours and the real-world adoption of those behaviours.

Finally, If we are to look at the COM-B model again, this program design does appear to support better capacity and motivation for users to learn more about their own SRH/mental health, and in some ways, such as the LoveLife call back and the service finder when it was functional, generate opportunities for users. When taken together, these three achievements show that this platform could support behaviour change for sufficiently engaged users, with some clear areas for improvement and growth. Importantly, the achievements that seem to have been produced speak to a more-than-average engaged user as compared to the general YAL user base. Thus, in the next phase of the platform, efforts should be made to drive better engagement through improvements to the user experience, innovation through personalization and diversified content approaches such as gamification and narrative storytelling, and consideration of ways to further incentivize consistent engagement through strategic partnerships with service providers or other benefits to the YAL user. This could help to heighten engagement more broadly across the platform and potentially contribute to more widespread results across the user base.

The YAL platform has identified an important offering for South African youth: an unbiased, judgement-free intervention that gives them any time access to content in formats that are relevant and helpful which, without the platform, they are much less likely to seek through friends, family, or health facilities. Importantly, given that this evaluation report also identifies significant improvements in platform users' measure of persistence (locus of control), there is some hopefulness that the behavioural outcomes achieved through this phase of the program have a higher likelihood of sustaining for the near future. This report demonstrates the importance of continued investment in

digital health interventions, given the positive correlations discussed above, so that we can meet youth where they are, in ways that are most accessible to them.

## 8) References

Bendoly E. System dynamics understanding in projects: information sharing, psychological safety, and performance effects. *Prod. Oper. Manag.*, 23 (8) (2014), pp. 1352-1369, 10.1111/poms.

Beyond Zero, 2019. Populations size estimates for Transgender persons in South Africa 2021. Available at:

<https://beyondzero.org.za/wp-content/uploads/2022/10/J8006-BeyoneZero-TG-Pop-Size-2021-Report-0718-LR.pdf> (Accessed Feb 25, 2024)

Chersich, M.F., Wabiri, N., Risher, K., Shisana, O., Celentano, D., Rehle, T., Evans, M. and Rees, H., 2017. Contraception coverage and methods used among women in South Africa: A national household survey. *South African Medical Journal*, 107(4), pp.307-314.

Craig, A., Rochat, T., Naicker, S.N., Mapanga, W., Mtintsilana, A., Dlamini, S.N., Ware, L.J., Du Toit, J., Draper, C.E., Richter, L. and Norris, S.A., 2022. The prevalence of probable depression and probable anxiety, and associations with adverse childhood experiences and socio-demographics: A national survey in South Africa. *Frontiers in Public Health*, 10, p.986531.

Closson, K., Dietrich, J.J., Beksinska, M., Gibbs, A., Hornschuh, S., Smith, T., Smit, J., Gray, G., Ndung'u, T., Brockman, M. and Kaida, A., 2019. Measuring sexual relationship power equity among young women and young men South Africa: Implications for gender-transformative programming. *PloS one*, 14(9), p.e0221554.

Department of Health, South Africa Demographic and Health Survey, 2016

Druce, K.L., Dixon, W.G. and McBeth, J., 2019. Maximizing engagement in mobile health studies: lessons learned and future directions. *Rheumatic Disease Clinics*, 45(2), pp.159-172.

Ghosh, A. (2022) *Assumptions in OLS Regression - Why do they matter?* Towards Data Science. Available at:

<https://towardsdatascience.com/assumptions-in-ols-regression-why-do-they-matter-9501c800787d> (accessed August 8, 2023)

Imai, K. and Kim, I.S., 2019. When should we use unit fixed effects regression models for causal inference with longitudinal data? *American Journal of Political Science*, 63(2), pp.467-490.

Kroenke K, Spitzer RL, Williams JB, Löwe B. 2009. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics*, 50(6):613-621

McDonagh, L.K., Saunders, J.M., Cassell, J., Bastaki, H., Hartney, T. and Rait, G., 2017. Facilitators and barriers to chlamydia testing in general practice for young people using a theoretical model (COM-B): a systematic review protocol. *BMJ open*, 7(3).

Mungai, K. and Bayat, A., 2019. An overview of trends in depressive symptoms in South Africa. *South African journal of psychology*, 49(4), pp.518-535.

Mngoma, N.F., Ayonrinde, O.A., Fergus, S., Jeeves, A.H. and Jolly, R.J., 2021. Distress, desperation and despair: anxiety, depression and suicidality among rural South African youth. *International review of psychiatry*, 33(1-2), pp.64-74.

Nolen, E. and Panisch, L.S., 2022. The relationship between body appreciation and health behaviors among women and adolescent girls: A scoping review. *Health & Social Work*, 47(2), pp.113-122.

Ng, M.M., Firth, J., Minen, M. and Torous, J., 2019. User engagement in mental health apps: a review of measurement, reporting, and validity. *Psychiatric Services*, 70(7), pp.538-544.

National Youth Development Agency. (2022) State of Youth Report 2022

SALDRU (2023) Souther Africa Labour and Development Research Unit: Income Comparison Tool. Available at: <https://www.saldru.uct.ac.za/income-comparison-tool/> (Accessed: August 8 2023)

Simbayi, L., Zuma, K., Zungu, N., Moyo, S., Marinda, E., Jooste, S., Mabaso, M., Ramlagan, S., North, A., van Zyl, J., Mohlabane, N., Dietrich, C., Naidoo, I., & the SABSSM V Team. (2019). *South African National HIV prevalence, incidence, behaviour and communication survey, 2017*. HSRC Press.

Statistics South Africa. (2022) *Statistical release P0310.1: National Poverty Lines 2022*. Statistics South Africa. Pretoria

Torous, J., Firth, J., Huckvale, K., Larsen, M.E., Cosco, T.D., Carney, R., Chan, S., Pratap, A., Yellowlees, P., Wykes, T. and Keshavan, M., 2018. The emerging imperative for a consensus approach toward the rating and clinical recommendation of mental health apps. *The Journal of nervous and mental disease*, 206(8), pp.662-666.

UCLA: Statistical Consulting Group. (2022) Repeated Measures Analysis with STATA. Available at: <https://stats.oarc.ucla.edu/sas/modules/introduction-to-the-features-of-sas/> (accessed August 8, 2023).

West, R. and Michie, S., 2020. A brief introduction to the COM-B Model of behaviour and the PRIME Theory of motivation [v1]. *Qeios*.

## 9) Appendix A - Regression specification

### MODEL FIT ASSESSMENT

When choosing a regression model, the research team focus on a comparison of models 4 and 5, as reported in the following table below, where model 4 is the fully specified model as in Section 4.3, and model 5 is that same model but with the inclusion of the 2 location variables (which are only available for a sub-sample of users).

Table A1: Coefficient on time for all barriers of interest from mixed model regressions

Controls included in model specification	Mixed model regressions				
	1	2	3	4	5
Time invariant controls	-	X	X	X	X
Urban and province dummies	-	-	X	-	X
Interaction of time and income	-	-	-	X	X

We extended our original model (Model 4) to adjust for location variables (rural versus urban and province). The introduction of these two variables or regressors led to a decrease in model sample size from 502 users completing baseline and endline to only 204 with valid location data. As such, the introduction of the location parameters into the model (i.e., model 5) introduced significant missingness into the model, with both location variables missing for 298 (59.4%) users. The large proportion of missing data in this model (greater than 5% threshold [1,2]) introduces bias to the results of the model with additional regressors due to the reduced sample size. Literature highlights that increasing missingness leads to inconsistency of effect estimates which affects statistical inference significantly [3]. If more than 40% of data are missing (which is the case resulting from the introduction of the location variables) then results should only be considered as hypothesis-generating [4]. Normally, where such challenges are encountered in the literature, one will conduct imputation to impute the missing data for the observations for which the variables are missing following Rubin's recommendations [5]. And even then, the model run on the imputed data would only serve as a form of sensitivity analysis or check. However, in order to construct imputed values, we would need to use other data not currently in the model. In our case, we're already including all demographic variables into the model, and so have no data outside of the model to impute upon. A further assessment of differences in outcomes at baseline between participants with valid location data and those without is shown in Table A2. We found differences in reporting between these groups for a few outcomes namely: Low SRH knowledge, depression/anxiety and low self-esteem.

To evaluate whether the inclusion of the two additional regressors improves the fit of the original model, we fitted model 4 on the number with valid location data (n=204) and compared it with a model that includes location variables. We first used the likelihood ratio test and information criterion indices (AIC and BIC) to compare the two models (see Table A2 below). Based on the results in Table A2, the AIC and BIC for the two models are not significantly different for each outcome, and the high p-value ( $p > 0.05$ ) indicates that the data is consistent with the claim that the extra variables together (not just individually) do not substantially improve model fit.

We further conducted a Wald test on the model with additional location variables to assess if removing the location variables compromises or reduces the fit of the model. Based on the literature, if the Wald test shows insignificance on the additional variables, then all the fit measures (e.g., AIC-BIC, residual variance) will not be significantly different between models with and without the location variables [6, 7]. The Wald test is used to assess if the parameters of added variables in the model are simultaneously equal to zero. Our findings (see Table A3) show that the parameters are not significantly different from zero ( $p > 0.05$ ) except for the poor SRH attitudes outcome model, which strongly suggests that removing them from the model will not substantially reduce the fit of that model. These findings align with the observations in the likelihood ratio test and information criterion indices and present that there is no sufficient evidence to motivate the addition of the location variables in these models.

**Table A2: Likelihood ratio test comparing model 4a to model 5**

	Outcome of interest	N	ll(model)	df	AIC	BIC	LR (Chi-square) test statistic	p-value
<b>Outcome 1</b>	Used contraception	266	-175.072	13	376.1431	422.7285	1.18	0.5535
	Used contraception*	266	-174.48	15	378.9601	432.7125		
<b>Outcome 2</b>	Used condom	260	-174.092	13	374.1837	420.4725	2.39	0.3022
	Used condom*	260	-172.895	15	375.7904	429.2006		
<b>Outcome 3</b>	1 or fewer partners	287	-3.90545	13	33.81091	81.38418	0.94	0.6246
	1 or fewer partners*	287	-3.43485	15	36.86969	91.76193		
<b>Outcome 4</b>	Ever tested STI	282	-108.758	13	243.5167	290.8615	0.34	0.8426
	Ever tested STI*	282	-108.587	15	247.1742	301.8028		
<b>Outcome 5</b>	Poor self-perceived healthcare	287	-196.537	13	419.0747	466.648	0.13	0.9368
	Poor self-perceived healthcare*	287	-196.472	15	422.9441	477.8363		
<b>Outcome 6</b>	Used contraception (given no plan for children in the next year)	196	-122.486	13	270.9721	313.5876	0.4	0.8191
	Used contraception (given no plan for children in the next year)*	196	-122.287	15	274.573	323.7447		
<b>Outcome 7</b>	Used condom (given no plan for children in the next year)	192	-123.852	13	273.7038	316.0512	2.23	0.328
	Used condom (given no plan for children in the next year)*	192	-122.737	15	275.4742	324.3366		



<b>Outcome 8</b>	Low SRH knowledge	287	-25.1585	13	76.31703	123.8903	3.04	0.219
	Low SRH knowledge*	287	-23.6399	15	77.27981	132.172		
<b>Outcome 9</b>	Knowledge condoms reduce STI risk	287	63.0885	13	-100.177	-52.6037	2.46	0.2919
	Knowledge condoms reduce STI risk*	287	63.1731	15	-96.3462	-41.4539		
<b>Outcome 10</b>	Knowledge exclusive partners reduce STI risk	283	-118.679	13	263.3585	310.7493	0.68	0.7135
	Knowledge exclusive partners reduce STI risk*	283	-117.448	15	264.8956	319.5773		
<b>Outcome 11</b>	Knowledge use of ineffective contraceptives	266	-125.047	13	276.0935	322.6789	0.12	0.9408
	Knowledge use of ineffective contraceptives*	266	-124.709	15	279.4182	333.1707		
<b>Outcome 12</b>	Poor body image	287	-157.925	13	341.8503	389.4236	0.12	0.9408
	Poor body image*	287	-157.864	15	345.7282	400.6204		
<b>Outcome 13</b>	Poor gender attitudes	287	-153.264	13	332.5275	380.1007	5.28	0.0715
	Poor gender attitudes*	287	-150.625	15	331.2501	386.1423		
<b>Outcome 14</b>	Poor consent attitudes	287	-106.84	13	239.6802	287.2535	2.2	0.3331
	Poor consent attitudes*	287	-105.741	15	241.4813	296.3735		
<b>Outcome 15</b>	Depressed or Anxious	287	-137.186	13	300.3717	347.945	0.21	0.8991
	Depressed or Anxious*	287	-137.08	15	304.1591	359.0513		
<b>Outcome 16</b>	Misusing substances	287	-143.738	13	313.4755	361.0488	2.64	0.2677
	Misusing substances*	287	-142.42	15	314.8399	369.7321		

<b>Outcome 17</b>	Consent1 low	287	35.08365	13	-44.1673	3.4060	1.22	0.5422
	Consent1 low*	287	35.34639	15	-40.6928	14.1995		
<b>Outcome 18</b>	Consent2 low	251	-110.492	13	246.9842	292.8151	0.15	0.9281
	Consent2 low*	251	-109.88	15	249.7598	302.6416		
<b>Outcome 19</b>	Body Image1 low	287	-165.354	13	356.7079	404.2812	1.89	0.388
	Body Image1 low*	287	-165.279	15	360.5587	415.4509		
<b>Outcome 20</b>	Body Image2 low	287	-181.54	13	389.079	436.6522	0.7	0.706
	Body Image2 low*	287	-180.595	15	391.1896	446.0819		
<b>Outcome 21</b>	Low social connection	287	-162.542	13	351.0838	398.6571	0.98	0.6134
	Low social connection*	287	-162.053	15	354.1064	408.9986		
<b>Outcome 22</b>	Locus of control (low)	287	-158.346	13	342.6915	390.2648	0.7	0.706
	Locus of control (low)*	287	-157.998	15	345.9952	400.8874		
<b>Outcome 23</b>	Low self-esteem	287	-101.69	13	229.3809	276.9542	1	0.607
	Low self-esteem*	287	-101.191	15	232.3825	287.2748		

\* Models with location variables are shown by (\*); AIC - Akaike Information Criterion; BI-Bayesian Information Criterion, ll(model)-log-likelihood value. The high p-value indicates that the data is consistent with the claim that the extra variables together (not just individually) do not substantially improve model fit.

**Table A3: Wald test for the significance of the additional parameters (full model)**

\*Testing that the parameters of added variables in the full model are simultaneously equal to zero

Outcome	Parameter	Chi-square statistic	p-value
Used contraception	Urban	1.09	0.5804
	Province		
Used condom	Urban	2.24	0.3262
	Province		
1 or fewer partners	Urban	0.89	0.6412
	Province		
Ever tested STI	Urban	0.32	0.8537
	Province		
Poor self-perceived healthcare	Urban	0.13	0.9393
	Province		
Used contraception (given no plan for children in the next year)	Urban	0.35	0.8408
	Province		
Used condom (given no plan for children in the next year)	Urban	2.04	0.3609
	Province		
low SRH knowledge	Urban	2.87	0.2379
	Province		
Poor body image	Urban	0.11	0.9449
	Province		
Poor SRH attitudes	Urban	<b>7.17</b>	<b>0.0278</b>
	Province		
Poor gender attitudes	Urban	<b>5.06</b>	<b>0.0795</b>
	Province		
Knowledge condoms reduce STI risk	Urban	0.16	0.9216
	Province		
Knowledge exclusive partners reduce STI risk	Urban	2.33	0.312

	Province		
Knowledge use of ineffective contraceptives	Urban	0.61	0.7377
	Province		
Poor consent attitudes	Urban	2.08	0.3542
	Province		
Depressed or Anxious	Urban	0.2	0.9051
	Province		
Misusing substances	Urban	2.49	0.2873
	Province		
Consent1 low	Urban	0.48	0.7872
	Province		
Consent2 low	Urban	1.16	0.5592
	Province		
Body image1 low	Urban	0.15	0.9276
	Province		
Body image2 low	Urban	1.79	0.4096
	Province		
Low social connection	Urban	0.91	0.633
	Province		
Locus of control (low)	Urban	0.66	0.7185
	Province		
Low self-esteem	Urban	0.91	0.6331
	Province		

**Table A4: Model specification**

Model	Definition	N (sample size)
Model 4	All control variables except for location variables	502
Model 4a	Only users with location data, using all control variables except for location variables	204

Model 5	Only users with location data, using all control variables including location variables	204
---------	---	-----

\*The inclusion of the location parameters into the model (i.e., model 5) introduces significant missingness, with location variables missing for 59.4% (n=298) of the 502 users.

To further assess model fit for each outcome, we estimated the adjusted R-squared from our linear probability mixed models (LPM) using the Rights and Sterba [10] framework. Table A5 below compares the estimates of adjusted R-squared values for all model variants in Table A4 fitted for each outcome. In LPM models, adjusted R-squared is interpreted as the fraction of the variance explained due to the difference between conditional means of the two groups of the binary outcome [11]. However, the standard use of this estimate as a descriptive tool for goodness-of-fit is similar to other definitions where the dependent variable is continuous. There were no large changes in the values of adjusted R-squared for the majority of the outcomes across the different model variants, except for the sub-index of knowledge regarding condoms and STI transmission and the binary variable for sex positivity (see Figure 1).

**Table A5: Estimates of adjusted R-squared values for model variants (4, 4a, and 5)**

Outcome	Adjusted R-squared measures		
	Model 4	Model 4a	Model 5
Used contraception	0.30	0.25	0.25
Used condom	0.41	0.44	0.44
1 or fewer partners	0.22	0.33	0.34
Ever tested STI	0.41	0.37	0.38
Poor self-perceived healthcare	0.29	0.21	0.22
Used contraception (given no plan for children in the next year)	0.28	0.29	0.30
Used condom (given no plan for children in the next year)	0.40	0.47	0.47
Low SRH knowledge	0.43	0.45	0.46

<b>Knowledge condoms reduce STI risk</b>	<b>0.34</b>	<b>0.67</b>	<b>0.68</b>
Knowledge exclusive partners reduce STI risk	0.34	0.34	0.35
Knowledge use of ineffective contraceptives	0.40	0.35	0.35
Poor body image	0.40	0.34	0.35
<b>Poor sex positivity</b>	<b>0.32</b>	<b>0.61</b>	<b>0.61</b>
Poor gender attitudes	0.41	0.49	0.50
Poor consent attitudes	0.25	0.18	0.18
Depressed or Anxious	0.20	0.17	0.18
Misusing substances	0.37	0.43	0.43
Low social connection	0.36	0.41	0.41
Low locus of control	0.23	0.28	0.29
<u>Low self-esteem</u>	<u>0.21</u>	<u>0.15</u>	<u>0.15</u>

In addition to the above, we also conducted two sets of Monte Carlo simulations [12] to investigate the actual coverage probability of confidence intervals and to test the significance of the bias of the confidence interval (or coefficients) for a fixed sample size. Confidence intervals represent the variability in the simulation by providing a range of likely values for an estimated parameter [13]. The coverage of confidence intervals is the percentage of times the confidence intervals include the true value of the estimated parameter. In the first set of Monte Carlo simulations, as depicted in Table A6, we assessed the coverage of the slope for the period (time) variable in each of the models and assumed that the estimate from the full sample model 4 variant is the true estimate. For each outcome, we simulated 1000 random samples using the model 4 variant and obtained the realization of the estimator for each sample. We then use the realizations to approximate the actual small sample distribution (mean-estimator of the true slope\*, and standard deviation) of the estimator and check properties, such as coverage probabilities or bias of the confidence intervals. We then repeated this process using model 4a and 5 variants using the slope for the period variable in model 4 variant as the true estimate.

Table A6 shows the results of the Monte Carlo simulations where the results of model 4 are assumed to be the true parameters. The average value of the slope for the period for the first set of iterations using the model 4 variant is very close to the true estimate or slope for all the outcomes. The bias or size distortion of the 95% confidence intervals using estimated coverage was not significantly different from zero at the 5% level for all outcomes (p-values greater than 0.05). Therefore, the first set of simulations produced unbiased estimates of the true slope for all outcomes. Repeating this process using model variants 4a and 5—reduced sample— showed that the confidence intervals were not consistent (biased) for the majority of the outcomes (p<0.05). This inconsistency could be a result of the missing data (the remaining sample has a different distribution).

In the second set of Monte Carlo simulations (as shown in Table A7), we conducted 200 simulations per outcome using model 5 (which includes location variables) and assuming that the slope for the period variable estimated from the model 5 variant is the true slope. This is done to assess the internal consistency of the slope estimate after repeated sampling with replacement. Table A7 shows the results of this exercise, with several confidence interval estimates not consistent at the 5% level, for over half of the outcomes. As such for all outcomes that are inconsistent under model 5, model 4 serves as the preferred fully specified model and is reported as such in the main report.

Comparing the results of the actual regressions estimates provided by models 4 and 5, we find that that 6 of the 10 outcomes that were consistent under model 5 returned the same results as was found by model 4. Model 4 and 5 agreed on a significant decrease of ~14.5 percentage points in the proportion of users with depression/anxiety ( $p < 0.05$ ). The models also both found large changes in locus of control (13 and 11.5 percentage points). While this was only significant at the 10 percent level for model 5, this is likely due to the fact that post hoc power calculations showed that model 5 was only powered at the 70% level. The models also concurred on finding no significant change in; contraception (for all users and those not planning on having children), substance misuse or self-esteem.

Then there were 4 outcomes where the regression results of models 4 and 5 differed, despite the Monte Carlo simulations of each being consistent. These were; attitudes towards consent, the knowledge index, the specific knowledge question about efficacious contraceptive options used, and social connectedness. And the difference between regression results for models 4 and 5 are represented in Table A10. For consent, the size of the coefficient remains relatively constant (8 percentage points as opposed to 7.5) but model 5's estimate is statistically insignificant. However model 5 is only powered at the 30% level (see Table A10). Therefore, given that the coefficient does not substantially change and model 4 is adequately powered, it seems appropriate to conclude in favour of model 4 as regards consent. For the knowledge index overall, the specific knowledge question about contraceptive methods and the social connectedness index, the coefficient on time does change more meaningfully moving to model 5 (dropping to a 3, 3 and 7 percentage point reduction respectively), with all becoming statistically insignificant. Post-hoc power calculations find that model 5 is only powered at 33%, 23% and 77% for these outcomes.

This presents a complicated situation to interpret. While both models are consistent, since the knowledge results are so substantially under-powered, it seems possible that the large reduction in sample size implied by model 5, affects its ability to detect the same result as found with power under model 4. As such this paper concludes in favour of model 4 for both knowledge outcomes. Social connectedness on the other hand is both adequately powered and consistent under model 5. As such, model 4 and 5, therefore, imply different conclusions. As such this paper concludes that there is mixed evidence regarding changes in social connectedness and a clear conclusion cannot be drawn. Depending on the theoretical validity of the programme having differential associations on connectedness based on location, a reader should prefer model 4 or 5 respectively.



**Table A6: Monte Carlo simulation results using model 4 estimates as the true parameters.**

Model outcome	Model 4 (true estimate)	MC simulation based on Model 4 [mc=1000]					MC simulation based on Model 4a [mc=1000]					MC simulation based on Model 5 [mc=1000]				
	_b[period]	_b[period]	95% Confidence intervals		test statistic	p-value	_b[period]	95% Confidence intervals		test statistic	p-value	_b[period]	95% Confidence intervals		test statistic	p-value
total_good_behavs_end	0.092	0.094	-0.144	0.269	-0.360	0.722	-0.077	-0.370	0.157	-4.256	<0.001	-0.076	-0.368	0.144	-4.256	<0.001
beh_contraception	0.032	0.033	-0.031	0.094	-1.658	0.100	-0.020	-0.083	0.046	-5.759	0.000	-0.023	-0.085	0.045	-6.258	<0.001
beh_condom	0.079	0.079	0.028	0.116	-0.898	0.377	-0.042	-0.114	0.012	-	-	-0.041	-0.117	0.011	-	-
beh_partners_less_2	-0.037	-0.037	-0.052	-0.027	-0.360	0.722	-0.0394	-0.068	-0.020	1.167	0.246	-0.038	-0.068	-0.019	2.132	0.035
beh_STI_test	0.038	0.038	0.007	0.059	-0.858	0.247	0.019	-0.021	0.051	-3.367	0.001	0.019	-0.017	0.051	-2.985	0.004
beh_SPHC_low	-0.109	-0.109	-0.172	-0.067	-0.892	0.177	-0.073	-0.138	-0.003	-2.143	0.035	-0.074	-0.152	0.013	-1.391	0.167
total_good_beh_nokid	0.157	0.159	-0.070	-0.070	-0.813	0.313	-0.010	-0.300	0.227	-3.909	<0.001	-0.004	-0.272	0.211	-4.256	<0.001
beh_contra_nokid	0.054	0.054	-0.009	0.096	-0.846	0.317	-0.004	-0.067	0.062	-7.283	<0.001	-0.005	-0.066	0.053	-8.955	<0.001
beh_condom_nokid	0.107	0.108	0.065	0.151	-1.508	0.138	-0.017	-0.073	0.040	-	-	-0.014	-0.072	0.045	-	-
knowledge_srh_low	-0.075	-0.075	-0.102	-0.048	-1.167	0.249	0.032	0.011	0.054	-	-	0.035	0.014	0.057	-	-
knowledge1_srh_condoms	0.034	0.034	0.007	0.048	-1.167	0.249	0.007	-0.011	0.025	-15.084	<0.001	0.007	-0.014	0.029	-9.929	<0.001
knowledge2_srh_exclusive	0.058	0.059	0.028	0.089	-0.774	0.443	0.058	0.010	0.106	-	-	0.054	0.017	0.090	-	-
knowledge9_srh_contra	0.100	0.100	0.072	0.129	-0.360	0.722	0.040	-0.009	0.089	-14.472	<0.001	0.039	0.000	0.079	-28.191	<0.001
att_BI_low	-0.130	-0.130	-0.164	-0.095	-0.898	0.377	-0.096	-0.145	-0.047	-4.596	<0.001	-0.097	-0.146	-0.047	-4.256	<0.001
att_srh_low	-0.018	-0.018	-0.039	0.004	-0.898	0.377	0.042	-0.001	0.085	-19.942	<0.001	0.048	0.016	0.081	-66.096	<0.001
att_gender_low	-0.010	-0.010	-0.044	0.025	-0.898	0.377	0.016	-0.039	0.070	-2.143	0.035	0.019	-0.028	0.067	-4.256	<0.001
att_consent_low	-0.080	-0.080	-0.105	-0.055	-0.360	0.722	-0.079	-0.125	-0.033	2.132	0.035	-0.076	-0.113	-0.039	4.010	<0.001
psy_depanx_low	0.548	0.548	0.518	0.579	-1.320	0.197	0.600	0.554	0.646	-18.656	<0.001	0.600	0.551	0.648	-10.451	<0.001
psy_Sub Abuse	0.016	0.016	-0.014	0.046	-0.360	0.722	0.050	0.003	0.098	-4.596	<0.001	0.052	0.005	0.100	-5.263	<0.001
psy_connect_low	-0.109	-0.109	-0.143	-0.074	-0.898	0.377	-0.069	-0.225	0.088	0.010	0.999	-0.072	-0.122	-0.022	-4.931	<0.001
pers_LOC_low	-0.132	-0.132	-0.171	-0.092	-0.360	0.722	-0.122	-0.170	-0.073	0.508	0.613	-0.123	-0.172	-0.075	4.000	<0.001
pers_SE_low	-0.018	-0.018	-0.048	0.013	-0.360	0.722	-0.069	-0.115	-0.023	-11.171	<0.001	-0.068	-0.107	-0.028	-15.240	<0.001

**Table A7: Monte Carlo simulation results using model variant 5 as the true model.**

Model outcome	Model 5 (true estimate)		MC simulation based on Model 5 [mc=200]			
	_b[period]	_b[period]	95% Confidence intervals		test statistic	p-value
total_good_behavs_end	-0.076	-0.080	-0.406	0.247	4.010	<b>0.000</b>
beh_contraception	-0.024	0.024	-0.169	0.121	-1.100	0.274
beh_condom	-0.040	0.042	-0.196	0.112	2.132	<b>0.035</b>
beh_partners_less_2	-0.038	-0.038	-0.127	0.051	4.100	<b>0.000</b>
beh_STI_test	0.019	0.017	-0.102	0.136	2.131	<b>0.034</b>
beh_SPHC_low	0.075	-0.074	-0.225	0.076	4.011	<b>0.000</b>
total_good_beh_nokid	-0.005	0.002	-0.307	0.311	0.720	0.472
beh_contra_nokid	-0.006	-0.002	-0.158	0.154	0.712	0.465
beh_condom_nokid	-0.012	-0.014	-0.163	0.135	2.132	<b>0.036</b>
knowledge_srh_low	0.036	0.037	-0.049	0.123	0.148	0.883
knowledge1_srh_condoms	0.008	0.005	-0.072	0.081	1.500	0.140
knowledge2_srh_exclusive	0.054	0.057	-0.057	0.171	1.768	<b>0.079</b>
knowledge9_srh_contra	0.039	0.042	-0.089	0.172	1.151	0.251
att_BI_low	-0.097	-0.098	-0.234	0.038	4.050	<b>0.000</b>
att_srh_low	0.049	0.046	-0.066	0.159	2.168	<b>0.033</b>
att_gender_low	0.020	0.017	-0.121	0.156	2.203	<b>0.030</b>

att_consent_low	-0.077	-0.074	-0.189	0.042	1.151	0.251
dep_anx_new_low	0.599	0.603	0.474	0.732	0.720	0.472
psy_SubAbuse	0.053	0.056	-0.079	0.191	0.623	0.534
psy_connect_low	-0.071	-0.066	-0.218	0.085	1.178	0.240
pers_LOC_low	-0.124	-0.119	-0.257	0.018	0.756	0.451
pers_SE_low	-0.068	-0.069	-0.181	0.043	1.227	0.221

**Table A8: log-likelihood for the model, Akaike's information criterion and Bayesian information criterion**

Model (outcomes)	Reduced parameter model					Full model (reduced sample due to missing)				
	N	ll(model)	df	AIC	BIC	N	ll(model)	df	AIC	BIC
Used contraception	676	-484.71	13	995.4	1054.1	266	-205.9	15	441.7	495.5
Used condom	670	-482.16	13	990.3	1048.9	260	-203.5	15	436.9	490.3
1 or fewer partners	749	-15.18	13	56.4	116.4	287	-43.4	15	116.8	171.7
Ever tested STI	726	-301.91	13	629.8	689.5	282	-143.5	15	316.9	371.5
Poor self-perceived healthcare	749	-545.64	13	1117.3	1177.3	287	-228.1	15	486.1	541.0
Used contraception (given no plan for children in the next year)	491	-342.88	13	711.8	766.3	196	-151.9	15	333.7	382.9
Used condom (given no plan for children in the next year)	489	-353.31	13	732.6	787.1	192	-151.5	15	333.0	381.8
low SRH knowledge	749	-353.97	13	733.9	794.0	287	-62.4	15	154.8	209.7

# REACH

## Health Made Possible

Poor body image	749	-448.24	13	922.5	982.5	287	-190.8	15	411.7	466.6
Poor SRH attitudes	749	-296.95	13	619.9	679.9	287	-124.3	15	278.6	333.5
Poor gender attitudes	749	-480.36	13	986.7	1046.8	287	-183.5	15	397.0	451.9
Poor consent attitudes	749	-343.28	13	712.6	772.6	287	-141.6	15	313.2	368.1
Depressed or Anxious	749	-427.97	13	881.9	942.0	287	-171.5	15	373.1	428.0
Misusing substances	748	-414.27	13	854.5	914.6	287	-175.8	15	381.6	436.5
Low social connection	749	-457.60	13	941.2	1001.2	287	-194.7	15	419.4	474.3
Locus of control (low)	749	-462.70	13	951.4	1011.4	287	-191.2	15	412.3	467.2
Low self-esteem	749	-360.94	13	747.9	807.9	287	-137.3	15	304.6	359.5

**Table A9: Differences in outcome scores at baseline.**

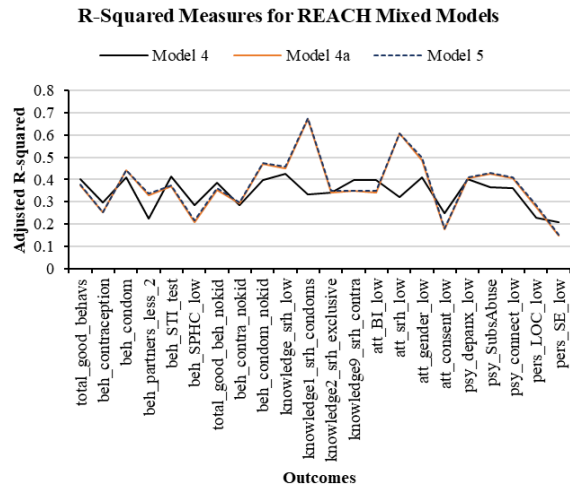
	Valid location data	Missing location data	Test	p-value
Model outcome	N=204	N=298		
total_good_behavs	3.00 (2.00-4.00)	3.00 (2.00-4.00)	Wilcoxon rank	0.380
beh_contraception	118 (64.1)	166 (63.6)	Chi-square	0.910
beh_condom	96 (52.7)	149 (56.4)	Chi-square	0.440
beh_partners_less_2	188 (92.2)	277 (93.0)	Chi-square	0.740
beh_STI_test	166 (83.0)	247 (87.0)	Chi-square	0.220
beh_SPHC_low	89 (43.6)	130 (43.6)	Chi-square	0.999
total_good_beh_nokid	3.00 (2.00-4.00)	4.00 (2.00-4.00)	Wilcoxon rank	0.260
beh_contra_nokid	90 (67.7)	132 (69.5)	Chi-square	0.730
beh_condom_nokid	77 (57.9)	120 (62.2)	Chi-square	0.440
knowledge_srh_low	19 ( 9.3)	65 (21.8)	Chi-square	<0.001
knowledge1_srh_condoms	193 (94.6)	286 (96.0)	Chi-square	0.470
knowledge2_srh_exclusive	162 (80.6)	-	-	
knowledge9_srh_contra	143 (76.9)	215 (80.2)	Chi-square	0.390
att_BI_low	46 (22.5)	66 (22.1)	Chi-square	0.920
att_srh_low	34 (16.7)	40 (13.4)	Chi-square	0.310
att_gender_low	55 (27.0)	99 (33.2)	Chi-square	0.140

# REACH

## Health Made Possible

att_consent_low	29 (14.2)	43 (14.4)	Chi-square	0.950
dep_anx_new_low	114 (55.9)	197 (66.1)	Chi-square	0.020
psy_SubAbuse	52 (25.5)	68 (22.8)	Chi-square	0.490
psy_connect_low	132 (64.7)	200 (67.1)	Chi-square	0.580
consent1_low	11 ( 5.4)	20 ( 6.7)	Chi-square	0.550
consent2_low	31 (17.2)	47 (17.7)	Chi-square	0.900
bodyimage1_low	58 (28.4)	86 (28.9)	Chi-square	0.920
bodyimage2_low	68 (33.3)	102 (34.2)	Chi-square	0.840
pers_LOC_low	44 (21.6)	66 (22.1)	Chi-square	0.880
pers_SE_low	25 (12.3)	61 (20.5)	Chi-square	0.016

**Figure A1: Plot of adjusted R-squared values**



**POST-HOC POWER ANALYSIS**

Finally, we conducted post-hoc power analysis to assess the observed power using the two-sample paired-proportions test (McNemar's test). Since the interclass correlation (ICC) generally seem to vary wildly in replications of the same study [8], even when based on large samples, we calculated the observed power based on a range of values of ICC (in light of the observed model ICC) and present multiple corresponding power estimates (see Table A9). The post-hoc test was conducted to re-examine the assumptions made *a priori*, based on the observed data and to provide both resolution to our prior misconceptions and guide for designing future follow-up studies. Under all three ICC conditions (minimum, model, and optimistic scenario), the findings in Table A4 show that power estimates for *'used condom, low SRH knowledge, poor body image, poor consent attitudes, depressed or anxious, low social connection, and locus of control (low)'* outcome models were higher than the 80% standard power threshold [9], while the power estimates for remaining outcomes models were lower. These findings point to being underpowered to conduct analysis related to the following outcomes: *'used contraception, 1 or fewer partners, ever tested STI, poor self-perceived healthcare, used contraception (given no plan for children in the next year), used condom (given no plan for children in the next year), poor SRH attitudes, poor gender attitudes, misusing substances, and low self-esteem'*. Therefore, any additional follow-up will consider these findings to ensure we have adequate sample size to achieve the minimum desired power.



**Table A10: Post-hoc observed power tests: two-sample paired-proportions test (McNemar's test)**

Outcome			rho/ICC			n	Power estimate		
	P+1	P1+	minimum ICC	model ICC	optimistic ICC		minimum	model	optimistic
Used contraception	0.36	0.41	0.27	0.29	0.30	390	0.388	0.397	0.401
Used condom	0.45	0.54	0.36	0.39	0.41	386	<b>0.876</b>	<b>0.891</b>	<b>0.901</b>
1 or fewer partners	0.07	0.06	0.21	0.22	0.23	424	0.081	0.081	0.081
Ever tested STI	0.15	0.20	0.38	0.39	0.41	412	0.665	0.671	0.686
Poor self-perceived healthcare	0.56	0.50	0.23	0.26	0.27	424	0.512	0.528	0.533
Used contraception (given no plan for children in the next year)	0.31	0.37	0.25	0.26	0.28	286	0.414	0.418	0.428
Used condom (given no plan for children in the next year)	0.40	0.49	0.34	0.37	0.39	283	0.751	0.771	0.783
Low SRH knowledge	0.83	0.77	0.38	0.42	0.44	424	<b>0.801</b>	<b>0.813</b>	<b>0.826</b>
Knowledge condoms STIs	0.04	0.05	0.04	0.05	0.05	424	0.133	0.136	0.138
Knowledge partners STIs	0.19	0.21	0.19	0.21	0.22	165	0.084	0.085	0.086
Knowledge effective contra.	0.21	0.29	0.21	0.29	0.30	392	<b>0.895</b>	<b>0.904</b>	<b>0.913</b>
Poor body image	0.78	0.63	0.33	0.37	0.38	424	<b>0.998</b>	<b>0.999</b>	<b>0.999</b>
Poor SRH attitudes	0.85	0.84	0.25	0.29	0.31	424	0.075	0.076	0.077
Poor gender attitudes	0.69	0.67	0.36	0.39	0.41	424	0.122	0.126	0.128
Poor consent attitudes	0.86	0.78	0.17	0.21	0.22	424	<b>0.912</b>	<b>0.925</b>	<b>0.927</b>
Consent1 low	0.94	0.95	0.12	0.13	0.14	424	<b>0.104</b>	<b>0.105</b>	<b>0.106</b>
Consent2 low	0.82	0.72	0.23	0.24	0.26	392	<b>0.965</b>	<b>0.967</b>	<b>0.971</b>
Body image1 low	0.71	0.59	0.36	0.37	0.39	424	<b>0.985</b>	<b>0.986</b>	<b>0.988</b>
Body image2 low	0.66	0.55	0.35	0.37	0.39	424	<b>0.982</b>	<b>0.984</b>	<b>0.987</b>
Misusing substances	0.76	0.80	0.30	0.34	0.35	424	0.388	0.407	0.412
Low social connection	0.34	0.23	0.30	0.31	0.33	424	<b>0.988</b>	<b>0.989</b>	<b>0.991</b>

Locus of control (low)	0.78	0.64	0.14	0.17	0.18	424	<b>0.998</b>	<b>0.998</b>	<b>0.998</b>
Low self-esteem	0.83	0.77	0.15	0.19	0.20	424	0.656	0.677	0.682

\* P+1 is the proportion of success for each outcome at endline, P1+ is the proportion of success for each outcome at baseline.  
 \* n is the total number of pairs/clusters. ICC is the intraclass correlation which corresponds to the correlation of measurements within the same cluster or individual in this case.  
 \* We used the McNemar's test in Stata software program.  
 \* Minimum and optimistic ICC are 5% percent deviations (negative and positive respectively) from the model ICC

**Table A11: Comparison of regression results and power for results which are consistent under model 4 and model 5**

VARIABLES	Consistency		Model 4		Coefficient	Model 5		Post-hoc power
	Model 4	Model 5	Number of groups	Number of groups		Coefficient	Post-hoc power	
Contraception	X	X	390		0.032	157	-0.022	0.089
Contraception (given no child plans)	X	X	286		0.052	116	-0.001	0.086
Condom use	X		386		0.080*			
Condom use (given no child plans)	X		283		0.106**			
Less than 2 partners	X		424		-0.038*			
Ever test for STI	X		412		0.038			
Low Self-Perceived Healthcare	X		424		-0.109**			
Low knowledge (index)	X	X	424		-0.075**	168	0.036	0.326
Low condom knowledge	X	X	424		-0.034*	168	-0.007	0.068
Low exclusivity knowledge	X		165		-0.061			
Low contraception knowledge	X	X	392		-0.100***	157	-0.030	0.232
Low body image (index)	X		424		-0.130***			
Body image (not feel good about self)	X	~	424		-0.105***			
Body image (not feel good about body)	X	~	424		-0.079**			
Low sex positivity	X		424		-0.017			
Low gender attitudes	X		424		-0.011			
Low consent attitudes	X	X	424		-0.080**	168	-0.075	0.314
Consent (agree in right to force sex)	X	~	424		0.034			
Consent (comfortable saying no to sex)	X	~	392		-0.103***			
Depressed/anxious	X	X	424		-0.088***	168	-0.074	0.927
Substance misuse	X	X	424		0.016	168	0.053	0.521
Low social connectedness	X	X	424		-0.108***	168	-0.067	0.768
Low Locus of Control	X	X	424		-0.132***	168	-0.114*	0.723
Low self esteem	X	X	424		-0.018	168	-0.063	0.709

## Appendix A - References

- [1]. Alice, M., 2015. Imputing missing data with R; MICE package. *Data Science Plus*.
- [2]. Schafer, J.L., 1999. Multiple imputation: a primer. *Statistical methods in medical research*, 8(1), pp.3-15.
- [3]. Mishra, S. and Khare, D., 2014. On comparative performance of multiple imputation methods for moderate to large proportions of missing data in clinical trials: a simulation study. *J Med Stat Inform*, 2(1), p.9.
- [4]. Dong, Y. and Peng, C.Y.J., 2013. Principled missing data methods for researchers. *SpringerPlus*, 2, pp.1-17.
- [5]. Rubin, D.B., 1976. Inference and missing data. *Biometrika*, 63(3), pp.581-592.
- [6]. Hanchane, S. and Mostafa, T., 2012. Solving endogeneity problems in multilevel estimation: an example using education production functions. *Journal of applied statistics*, 39(5), pp.1101-1114.
- [7]. Bell, A., Fairbrother, M. and Jones, K., 2019. Fixed and random effects models: making an informed choice. *Quality & quantity*, 53, pp.1051-1074.
- [8]. Mehta S, Bastero-Caballero RF, Sun Y, Zhu R, Murphy DK, Hardas B, Koch G. Performance of intraclass correlation coefficient (ICC) as a reliability index under various distributions in scale reliability studies. *Stat Med*. 2018 Aug 15;37(18):2734-2752. <https://www.doi:10.1002/sim.7679>.
- [9]. Serdar, C.C., Cihan, M., Yücel, D. and Serdar, M.A., 2021. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. *Biochemia medica*, 31(1), pp.27-53.
- [10] Rights, J. D., and S. K. Sterba. 2023b. R-squared measures for multilevel models with three or more levels. *Multivariate Behavioral Research*, 58: 340-367
- [11] Gronau, R., 1998. A Useful Interpretation of R-Squared in Binary Choice Models (Or, Have We Dismissed the Good Old R-Squared; Prematurely).
- [12] Flowers-Cano, R.S., Ortiz-Gómez, R., León-Jiménez, J.E., López Rivera, R. and Perera Cruz, L.A., 2018. Comparison of bootstrap confidence intervals using Monte Carlo simulations. *Water*, 10(2), p.166.
- [13] Adkins, L.C. and Gade, M.N., 2012. Monte Carlo experiments using Stata: a primer with examples. In 30th Anniversary Edition (pp. 429-477). Emerald Group Publishing Limited.

## 10) Appendix B - Survey Instruments

### 10.1) WhatsApp Baseline Survey Questions

Section 1: Demographics		
No.	Question	Response Options
101	Gender	1. Woman 2. Man 3. Non-binary or transgender 4. Prefer not to say
102	What is your age? <i>Age in years</i>  -----	<i>Strictly held between 18-24 years old</i>
103	Are you currently in a relationship ?	1. Yes, seeing someone 2. No, I'm single 3. It's complicated
104	Are you HIV positive?	1. Yes 2. No 3. Rather not say 4. Skip this question
105	(Ask if 104 is Yes) Do you take Antiretroviral Therapy (ART) medication on a regular basis?	1. Yes 2. No 3. Rather not say 4. Skip this question
106	(ask if 104 is No) Do you take Pre-Exposure Prophylaxis (PrEP) medication on a regular basis?	1. Yes 2. No 3. Rather not say 4. Skip this question
107	Are you in South Africa?	1. Yes 2. No 3. Rather not say 4. Skip this question
109	Have you been part of the Young Africa Live Pilot survey before?	1. Yes 2. No 3. Rather not say 4. Skip this question
110	Have you been exposed to the Young Africa Live platform and its content before?	1. Yes 2. No 3. Rather not say 4. Skip this question
111	What is the total monthly income of your whole household?	1. No income 2. R1 - R400 3. R401 - R800 4. R801 - R1600 5. R1 601 - R3200 6. R3 201 - R6400 7. R6 401 - R12800 8. R12 801 - R25600 9. R25 601 - R51200

		10. R51 201 – R102 400 11. R102 401 – R204 800 12. R204 801 or more
112	In the past 7 days, how many days did you go hungry?	1. None 2. 1-2 3. 3-4 4. 5-7 5. Rather not say 6. Skip this question
<b>Section 2: Locus of Control:</b> <i>Looking after you and your health</i> <i>I've got a few questions to help me figure out where you're at when it comes to taking care of your love and health needs.</i>		
201	<i>Can I start by asking how much you agree or disagree with some statements about you, your life, and your health?</i>	01: OK, let's start! 02: I can't right now
202	I'm my own boss. 😊	1. Does not apply at all 2. Applies somewhat 3. Applies 4. Applies a lot 5. Applies completely 6. I don't understand 7. Skip
203	If I work hard, I will be successful. 🤔	1. Does not apply at all 2. Applies somewhat 3. Applies 4. Applies a lot 5. Applies completely 6. I don't understand 7. Skip
204	Whether at work or in my personal life, what I do mainly depends on other people. 🧑🧑	1. Does not apply at all 2. Applies somewhat 3. Applies 4. Applies a lot 5. Applies completely 6. I don't understand 7. Skip
205	Fate often gets in the way of my plans.	1. Does not apply at all 2. Applies somewhat 3. Applies 4. Applies a lot 5. Applies completely 6. I don't understand 7. Skip
<b>Section 4: Self Esteem</b>		
401	I feel that I am a person of worth, at least on an equal plane with others.	1. Strongly agree 2. Agree 3. Disagree 4. Strongly disagree 5. I don't understand 6. Skip
402	I feel that I have a number of good qualities.	1. Strongly agree 2. Agree

		<ol style="list-style-type: none"> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
403	All in all, I am inclined to feel that I am a failure.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
404	I am able to do things as well as most other people.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
405	I feel I do not have much to be proud of.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
406	I take a positive attitude toward myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
407	On the whole, I am satisfied with myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
408	I wish I could have more respect for myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
409	I certainly feel useless at times.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
410	At times I think I am no good at all.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 5: Connectedness</b>		

501	Do you have someone to talk to when you have a worry or problem?	<ol style="list-style-type: none"> <li>1. Never</li> <li>2. Some of the time</li> <li>3. Most of the time</li> <li>4. All the time</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 6: Body Image</b>		
601	I feel good about myself	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Sometimes</li> <li>4. I don't understand</li> <li>5. Skip</li> </ol>
602	I feel good about my body	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Sometimes</li> <li>4. I don't understand</li> <li>5. Skip</li> </ol>
<b>Section 7: Anxiety</b>		
701	<p><i>Over the last two weeks, how often have you been bothered by the following problems?</i></p> <p>1. <i>Feeling nervous, anxious or on edge</i></p>	<ol style="list-style-type: none"> <li>1. "Not at all"</li> <li>2. Several days</li> <li>3. More than half the days</li> <li>4. Nearly every day</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
702	2. Not being able to stop or control worrying	<ol style="list-style-type: none"> <li>1. "Not at all"</li> <li>2. Several days</li> <li>3. More than half the days</li> <li>4. Nearly every day</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 8: Depression</b>		
801	<p>Over the last two weeks, how often have you been bothered by the following problems?</p> <p>Feeling down, depressed or hopeless</p>	<ol style="list-style-type: none"> <li>1. "Not at all"</li> <li>2. Several days</li> <li>3. More than half the days</li> <li>4. Nearly every day</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
802	Little interest or pleasure in doing things	<ol style="list-style-type: none"> <li>1. "Not at all"</li> <li>2. Several days</li> <li>3. More than half the days</li> <li>4. Nearly every day</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 9: Self-Perceived Healthcare</b>		
901	How good a job do you feel you are doing in taking care of your health?	<ol style="list-style-type: none"> <li>1. Excellent</li> <li>2. Very Good</li> <li>3. Good</li> <li>4. Fair</li> <li>5. Poor</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
902	When I have a health need (e.g. contraception, flu symptoms), I go to my closest clinic	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Sometimes</li> <li>4. I don't understand</li> <li>5. Skip</li> </ol>
<b>Section 10: Sexual Reproductive Health Literacy</b>		

1001	<p>Is the following statement true or false?</p> <p>People can reduce the risk of getting sexual transmitted infections (STIs) by using condoms every time they have sexual intercourse.</p>	<ol style="list-style-type: none"> <li>1. True</li> <li>2. False</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1002	<p>Is the following statement true or false?</p> <p>People can reduce the risk of getting sexually transmitted diseases by only having sex with one partner who isn't infected and who has no other partners.</p>	<ol style="list-style-type: none"> <li>1. True</li> <li>2. False</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1003	<p>If I'm sexually active, I am able to insist on using condoms when I have sex.</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I'm not sexually active</li> <li>7. I don't understand</li> <li>8. Skip</li> </ol>
1004	<p>My sexual needs or desires are important.</p>	<ol style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1005	<p>I think it would be important to focus on my own pleasure as well as my partner's during sexual experiences.</p>	<ol style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1006	<p>I expect to enjoy sex.</p>	<ol style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1007	<p>The last time you had sex, did you or your partner do or use something to avoid or delay getting pregnant?</p>	<p>Yes No Don't remember Haven't had sex yet I don't understand Skip</p>
1008	<p>Did you use a condom last time you had penetrative sex?</p>	<p>Yes No Don't remember Haven't had sex yet</p>



		I don't understand Skip
1009	What's been the MAIN way you or your partner have tried to delay or avoid getting pregnant?	<ul style="list-style-type: none"> <li>1 - Contraceptive Pill</li> <li>2 - IUD (intrauterine device)</li> <li>3 - male condom</li> <li>4 - female condom</li> <li>5 - Contraceptive Injection</li> <li>6 - Contraceptive Implants</li> <li>7 - diaphragm</li> <li>9 - pulling out (withdrawal method)</li> <li>11 - standard days rhythm method</li> <li>sterilisation</li> <li>15 - exclusive</li> <li>16.- breastfeeding within the first 6 months of child birth</li> <li>17- Haven't had sex yet</li> <li>18- I don't understand</li> <li>19- Skip</li> </ul>
1010	How many sexual partners did you have over the last month?	<ul style="list-style-type: none"> <li>1. None</li> <li>2. One</li> <li>3. More than one</li> </ul>
1011	Ok. You can tell me how many sexual partners you had here. If "more than one" to 8	<Enter any number>
1012	Have you ever been tested for sexually transmitted infections (STIs) and HIV?	<ul style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. I don't understand</li> <li>5. Skip</li> </ul>
<b>Section 11: Gender Attitudes</b>		
1101	"How do you feel about each statement? There are no right or wrong answers. Would you say that you agree, somewhat agree or disagree with the following statements?" There are times when a woman deserves to be beaten	<ul style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>
1102	It's a woman's responsibility to avoid getting pregnant	<ul style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>
1103	A man and a woman should decide together what type of contraceptive to use	<ul style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>

1104	If a guy gets women pregnant, child is responsibility of both	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
<b>Section 12: Sexual Consent</b>		
1201	<p>"Robert and Samantha have been dating for 5 years and love each other very much.</p> <p>Every year on Robert's birthday, Samantha promises him sex for his birthday. This year, Samantha tells Robert that she is too tired for sex. To what extent do you agree with this statement: Robert has the right to force Samantha to have sex."</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1202	"If you're in a relationship, which of these statements describes you best?"	<ol style="list-style-type: none"> <li>1. I'm cool with telling bae no if they want to have sex but I don't.</li> <li>2. I find it hard to say no to bae if bae wants to have sex but I don't.</li> <li>3. I'm not sure how I feel about saying no when bae wants to have sex and I don't.</li> <li>4. I'm not in a relationship</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 13: Alcohol and Substance Abuse</b>		
1301	<p>Have you ever felt guilty about drinking or drug use?</p> <p>Have you ever felt you needed to cut down on your drinking or drug use?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1302	Have people annoyed you by criticising your drinking or drug use?	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1303	Have you ever felt you needed a drink or used drugs first thing in the morning (eye-opener)	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>

## 10.2) WhatsApp Endline Survey Questions

No.	Question	Response Options
<b>Section 1: Demographics</b>		
101	Are you seeing someone special right now?	1. Yes, seeing someone 2. No, I'm single 3. It's complicated 4. Rather not say 5. Skip this question
102	What is the total monthly income of your whole household?	1. No income 2. R1 - R400 3. R401 - R800 4. R801 - R1600 5. R1 601 - R3200 6. R3 201 - R6400 7. R6 401 - R12800 8. R12 801 - R25600 9. R25 601 - R51200 10. R51 201 - R102 400 11. R102 401 - R204 800 12. R204 801 or more 13. Rather not say 14. Skip this question
103a	How many people (including yourself) live in the household now? Don't forget to include babies. (If you're unsure - this counts as anyone sleeping the house 4 nights in the past week).	1 - Just me 2 - Two people 3 - Three people 4 - Four people 5 - Five people 6 - Six people 7 - Seven people 8 - Eight or more Rather not say Skip question  <i>(If answer = 8 THEN question 103b            ELSE proceed to question 104)</i>
103b	Okay - you said there are 8 or more people in your household. How many people (including yourself) live in the household now? Don't forget to include babies. (If you're unsure - this counts as anyone sleeping in the house 4 nights in the past week).	8 - Including me 9 - Nine people 10 - Ten people 11 - Eleven people 12 - Twelve people 13 - Thirteen people 14 - Fourteen people 15 - Fifteen or more Rather not say Skip question

104	In the past 7 days, how many days did you go hungry?	<ol style="list-style-type: none"> <li>1. None</li> <li>2. 1-2</li> <li>3. 3-4</li> <li>4. 5-7</li> <li>5. Rather not say</li> <li>6. Skip this question</li> </ol>
105	What province do you live in?	<ol style="list-style-type: none"> <li>1. Eastern Cape</li> <li>2. Free Stata</li> <li>3. Gauteng</li> <li>4. KwaZulu-Natal</li> <li>5. Limpopo</li> <li>6. Mpumalanga</li> <li>7. Northern Cape</li> <li>8. North-west</li> <li>9. Western Cape</li> <li>I don't understand</li> <li>Skip</li> </ol>
106	What type of area are you living in?	<ol style="list-style-type: none"> <li>1. Traditional area/chiefdom</li> <li>2. Urban area/town</li> <li>3. Farm / rural area</li> <li>I don't understand</li> <li>Skip</li> </ol>
107	Before joining B-Wise, how often did you have discussions or interact with content about sexual topics?	<ol style="list-style-type: none"> <li>1. A lot</li> <li>2. Somewhat</li> <li>3. Not much</li> <li>4. Never</li> <li>Skip</li> </ol>
<b>Section 2: User experiences</b>		
201	You have received a lot of content from B-Wise. Did B-Wise send you content that related to your sexual needs?	<ol style="list-style-type: none"> <li>1. Related extremely well</li> <li>2. Related well</li> <li>3. Related fine</li> <li>4. Related a little</li> <li>5. Didn't relate at all</li> <li>I don't understand</li> <li>Skip</li> </ol>
202	For the content that B-Wise sent you that related to your needs, was the content that B-Wise sent you interesting?	<ol style="list-style-type: none"> <li>1. It was extremely interesting</li> <li>2. It was quite interesting</li> <li>3. It was kind of interesting</li> <li>4. It was not really interesting</li> <li>5. It was extremely uninteresting</li> <li>I don't understand</li> <li>Skip</li> </ol>
203	How useful did you find the information the B-Wise sent you for managing your sexual health and relationship needs?	<ol style="list-style-type: none"> <li>1. It was extremely useful</li> <li>2. It was quite useful</li> <li>3. It was kind of useful</li> <li>4. It was not really useful</li> <li>5. Not at all useful</li> <li>I don't understand</li> <li>Skip</li> </ol>

204	Did you know that B-Wise is also on Facebook (WA if on FB survey)	<ol style="list-style-type: none"> <li>1. Yes, and I used it every week</li> <li>2. Yes, and I used it every month</li> <li>3. Yes, I used it, but not much</li> <li>4. Yes, but I never used it</li> <li>5. No I didn't know that</li> </ol> I don't understand Skip
205	Since joining B-Wise, have you ever felt like you needed to see a medical service about your sexual health?	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> </ol> I don't understand Skip  (IF answer = YES then question 206 ELSE question 209)
206	How many times have you visited a clinic or other health facility for your sexual health since joining B-Wise? (We know it may be hard to remember, we'd appreciate your best guess)	<ol style="list-style-type: none"> <li>0. None</li> <li>1.</li> <li>2.</li> <li>3.</li> <li>4.</li> <li>5.</li> <li>6.</li> <li>7.</li> <li>8 or more</li> </ol> I don't understand Skip (IF answer = 0 then question 207 IF answer >=1 then question 208 ELSE question 209)
207	Was there a reason you didn't go to the clinic or other health facility?	<ol style="list-style-type: none"> <li>1. I didn't know where to go</li> <li>2. I couldn't find the time to go</li> <li>3. I couldn't find the money to go</li> <li>4. I was afraid of being judged at the facility</li> <li>5. I was afraid of being mistreated at the facility</li> <li>6. I got help elsewhere</li> <li>7. I no longer felt I needed to go</li> </ol> I don't understand Skip  (For all answers move to question 209)
208	When you visited the clinic or other health facility, what was the outcome? (If you had different experiences, please pick the response that was true most of the time).	<ol style="list-style-type: none"> <li>1. I got help (either a diagnosis, medication or treatment)</li> <li>2. I was seen by a nurse/doctor but they don't know what's wrong</li> <li>3. I went to the clinic/facility but was not seen by a nurse or doctor</li> </ol> I don't understand Skip
209	Do you think that your time on B-Wise has changed how likely you are to visit a clinic or health facility for your sexual and mental health?	<ol style="list-style-type: none"> <li>1. A lot more likely</li> <li>2. Little more likely</li> <li>3. No change</li> <li>4. Little less likely</li> <li>5. A lot less likely</li> </ol> I don't understand Skip

210	Since joining B-Wise, have you ever felt like you needed to speak to a counsellor about your mental or sexual health?	1. Yes 2. No I don't understand Skip
211	Did you know you can use the B-Wise WhatsApp chatbot to ask LoveLife (a counselling group) to call you back?	1. Yes, I used it and I got help 2. Yes, I used it but didn't get help 3. Yes, but I never used it 4. No, I didn't know that 5. No, but I never needed help I don't understand Skip
212	Other than using LoveLife, have you visited another mental or sexual health counsellor (since joining B-Wise) and if so, how many times? (We know it may be hard to remember, we'd appreciate your best guess)	0. None - LoveLife was enough 1. 2. 3. 4. 5. 6 or more 7. None but I needed to I don't understand Skip  (IF answer = 7 then question 213 ELSE question 214)
213	What was the main reason you didn't speak to a counsellor about your mental or sexual health?	1. I didn't know where to go 2. I couldn't find the time to go 3. I couldn't find the money to go 4. I was afraid of being judged at the facility 5. I was afraid of being mistreated at the facility 6. I got help elsewhere 7. I no longer felt I needed to go I don't understand Skip
214	Do you think that your time on B-Wise has changed how likely you are to speak to a counsellor about your mental or sexual health?	1. A lot more likely 2. Little more likely 3. No change 4. Little less likely 5. A lot less likely I don't understand Skip
<b>Section 3: Locus of Control</b>		
	Great. Now for the next few questions, I'm going to ask how much you agree or disagree with some statements about you, your life, and your health?	01: OK, let's start! 02: I can't right now
301	I'm my own boss. 😎	1. Does not apply at all 2. Applies somewhat 3. Applies 4. Applies a lot 5. Applies completely 6. I don't understand 7. Skip

302	If I work hard, I will be successful. 😊	<ol style="list-style-type: none"> <li>1. Does not apply at all</li> <li>2. Applies somewhat</li> <li>3. Applies</li> <li>4. Applies a lot</li> <li>5. Applies completely</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
303	Whether at work or in my personal life, what I do mainly depends on other people. 🙏	<ol style="list-style-type: none"> <li>1. Does not apply at all</li> <li>2. Applies somewhat</li> <li>3. Applies</li> <li>4. Applies a lot</li> <li>5. Applies completely</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
304	Fate often gets in the way of my plans.	<ol style="list-style-type: none"> <li>1. Does not apply at all</li> <li>2. Applies somewhat</li> <li>3. Applies</li> <li>4. Applies a lot</li> <li>5. Applies completely</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
<b>Section 4: Self Esteem</b>		
401	I feel that I am a person of worth, at least on an equal plane with others.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
402	I feel that I have a number of good qualities.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
403	All in all, I am inclined to feel that I am a failure.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
404	I am able to do things as well as most other people.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
405	I feel I do not have much to be proud of.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>

406	I take a positive attitude toward myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
407	On the whole, I am satisfied with myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
408	I wish I could have more respect for myself.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
409	I certainly feel useless at times.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
410	At times I think I am no good at all.	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Disagree</li> <li>4. Strongly disagree</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 5: Connectedness</b>		
501	Do you have someone to talk to when you have a worry or problem?	<ol style="list-style-type: none"> <li>1. Never</li> <li>2. Some of the time</li> <li>3. Most of the time</li> <li>4. All the time</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
<b>Section 6: Body Image</b>		
601	I feel good about myself	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Sometimes</li> <li>4. I don't understand</li> <li>5. Skip</li> </ol>
602	I feel good about my body	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Sometimes</li> <li>4. I don't understand</li> <li>5. Skip</li> </ol>
<b>Section 7: Anxiety</b>		



701	Over the last two weeks, how often have you been bothered by the following problems? 1. Feeling nervous, anxious or on edge	1. "Not at all" 2. Several days 3. More than half the days 4. Nearly every day 5. I don't understand 6. Skip
702	2. Not being able to stop or control worrying	1. "Not at all" 2. Several days 3. More than half the days 4. Nearly every day 5. I don't understand 6. Skip
<b>Section 8: Depression</b>		
801	Over the last two weeks, how often have you been bothered by the following problems? Feeling down, depressed or hopeless	1. "Not at all" 2. Several days 3. More than half the days 4. Nearly every day 5. I don't understand 6. Skip
802	Little interest or pleasure in doing things	1. "Not at all" 2. Several days 3. More than half the days 4. Nearly every day 5. I don't understand 6. Skip
<b>Section 9: Self-Perceived Healthcare</b>		
901	How good a job do you feel you are doing in taking care of your health?	1. Excellent 2. Very Good 3. Good 4. Fair 5. Poor 6. I don't understand 7. Skip
902	When I have a health need (e.g. contraception, flu symptoms), I go to my closest clinic	1. Yes 2. No 3. Sometimes 4. I don't understand 5. Skip
<b>Section 10: Sexual Reproductive Health Literacy</b>		
1001	Is the following statement true or false?  People can reduce the risk of getting sexual transmitted infections (STIs) by using condoms every time they have sexual intercourse.	1. True 2. False 3. I don't understand 4. Skip
1002	Is the following statement true or false?  People can reduce the risk of getting sexually transmitted diseases by only having sex with one partner who isn't infected and who has no other partners.	1. True 2. False 3. I don't understand 4. Skip

1003	If I'm sexually active, I am able to insist on using condoms when I have sex.	<ul style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I'm not sexually active</li> <li>7. I don't understand</li> <li>8. Skip</li> </ul>
1004	My sexual needs or desires are important.	<ul style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>
1005	I think it would be important to focus on my own pleasure as well as my partner's during sexual experiences.	<ul style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>
1006	I expect to enjoy sex.	<ul style="list-style-type: none"> <li>1. Not at all true</li> <li>2. A little true</li> <li>3. Kind of true</li> <li>4. Very true</li> <li>5. Extremely true</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>
1007	The last time you had sex, did you or your partner do or use something to avoid or delay getting pregnant?	<ul style="list-style-type: none"> <li>Yes</li> <li>No</li> <li>Don't remember</li> <li>Haven't had sex yet</li> <li>I don't understand</li> <li>Skip</li> </ul>
1008	Did you use a condom last time you had penetrative sex?	<ul style="list-style-type: none"> <li>Yes</li> <li>No</li> <li>Don't remember</li> <li>Haven't had sex yet</li> <li>I don't understand</li> <li>Skip</li> </ul>

1009	What's been the MAIN way you or your partner have tried to delay or avoid getting pregnant?	<ul style="list-style-type: none"> <li>1 - Contraceptive Pill</li> <li>2 - IUD (intrauterine device)</li> <li>3 - male condom</li> <li>4 - female condom</li> <li>5 - Contraceptive Injection</li> <li>6 - Contraceptive Implants</li> <li>7 - diaphragm</li> <li>9 - pulling out (withdrawal method)</li> <li>11 - standard days rhythm method</li> <li>12- sterilisation</li> <li>15 - exclusive</li> <li>16.- breastfeeding within the first 6 months of child birth</li> <li>17- Haven't had sex yet</li> <li>18- I don't understand</li> <li>19- Skip</li> </ul>
1010	How many sexual partners did you have over the last month?	<ul style="list-style-type: none"> <li>1. None</li> <li>2. One</li> <li>3. More than one</li> </ul>
1011	Ok. You can tell me how many sexual partners you had here. If "more than one" to 8	<Enter any number>
1012	Have you ever been tested for sexually transmitted infections (STIs) and HIV?	<ul style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. I don't understand</li> <li>5. Skip</li> </ul>
1013	Over the past 5 months, do you think that your knowledge about the importance of using condoms has changed?	<ul style="list-style-type: none"> <li>1. Yes, improved a lot</li> <li>2. Yes, improved a little</li> <li>3. Stayed the same</li> <li>4. It's a little worse</li> <li>5. It's a lot worse</li> <li>I don't understand</li> <li>Skip</li> </ul>
1014	Which of these has most influenced your knowledge about using condoms?	<ul style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> <li>8. No change</li> </ul>
1015	Are you planning to have a child within the next year?	<ul style="list-style-type: none"> <li>1. Yes</li> <li>2. Maybe</li> <li>3. No</li> <li>4. Skip</li> </ul>
1016	Do you now plan to use condoms more consistently than you did 5 months ago?	<ul style="list-style-type: none"> <li>1. Yes - a lot more</li> <li>2. Yes - a little more</li> <li>3. No - same</li> <li>4. No - a little less</li> <li>5. No - a lot less</li> <li>6. I don't understand</li> <li>7. Skip</li> </ul>

1017	Do you now plan to go for STI or HIV tests more often than you did 5 months ago?	<ol style="list-style-type: none"> <li>1. Yes - a lot more</li> <li>2. Yes - a little more</li> <li>3. No - same</li> <li>4. No - a little less</li> <li>5. No - a lot less</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1018	Which of these has most influenced your plans to use condoms or test for STIs/HIV?	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> </ol>
<b>Section 11: Gender Attitudes</b>		
1101	"How do you feel about each statement? There are no right or wrong answers. Would you say that you agree, somewhat agree or disagree with the following statements?" There are times when a woman deserves to be beaten	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1102	It's a woman's responsibility to avoid getting pregnant	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1103	A man and a woman should decide together what type of contraceptive to use	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1104	If a guy gets women pregnant, child is responsibility of both	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
<b>Section 12: Sexual Consent</b>		

1201	<p>"Robert and Samantha have been dating for 5 years and love each other very much.</p> <p>Every year on Robert's birthday, Samantha promises him sex for his birthday. This year, Samantha tells Robert that she is too tired for sex. To what extent do you agree with this statement: Robert has the right to force Samantha to have sex."</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
1202	<p>"If you're in a relationship, which of these statements describes you best?"</p>	<ol style="list-style-type: none"> <li>1. I'm cool with telling bae no if they want to have sex but I don't.</li> <li>2. I find it hard to say no to bae if bae wants to have sex but I don't.</li> <li>3. I'm not sure how I feel about saying no when bae wants to have sex and I don't.</li> <li>4. I'm not in a relationship</li> <li>5. I don't understand</li> <li>6. Skip</li> </ol>
1203	<p>Which of these has most influenced your attitudes about sexual relationships and interactions?</p>	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> </ol>
<b>Section 13: Alcohol and Substance Abuse</b>		
1301	<p>Have you ever felt guilty about drinking or drug use?</p> <p>Have you ever felt you needed to cut down on your drinking or drug use?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1302	<p>Have you ever felt you needed to cut down on your drinking or drug use?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1303	<p>Have people annoyed you by criticising your drinking or drug use?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>
1304	<p>Have you ever felt you needed a drink or used drugs first thing in the morning (eye-opener)</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. I don't understand</li> <li>4. Skip</li> </ol>

### 10.3) Facebook - B-Wise page Survey Questions

Question number	Relevant questions	List of responses
1	Are you in South Africa?	1. Yes 2. No 3. Rather not say 4. Skip this question (IF NOT YES, THEN SEND MESSAGE EXPLAINING CANNOT BE A PART OF THE STUDY AS EXPLAINED IN CONSENT FORM)
2	What is your age? Age in years	Strictly held between 18-24 years old (IF BELOW 18 OR OVER 24, THEN SEND MESSAGE EXPLAINING CANNOT BE A PART OF THE STUDY AS EXPLAINED IN CONSENT FORM)
3	Roughly, when did you first visit the B-Wise Facebook page?	1. From July 2023 and after 2. June 2023 and before  (IF AFTER JUNE 2023, THEN SEND MESSAGE EXPLAINING CANNOT BE A PART OF THE STUDY AS EXPLAINED IN CONSENT FORM)
4	What gender do you identify as?	1. Woman 2. Man 3. Non-binary or transgender 4. Prefer not to say
5	What is the total monthly income of your whole household?	1. No income 2. R1 - R400 3. R401 - R800 4. R801 - R1600 5. R1 601 - R3200 6. R3 201 - R6400 7. R6 401 - R12800 8. R12 801 - R25600 9. R25 601 - R51200 10. R51 201 - R102 400 11. R102 401 - R204 800 12. R204 801 or more 13. Skip this question

6	<p>In the past 7 days, how many days did you go hungry?</p>	<ol style="list-style-type: none"> <li>1. None</li> <li>2. 1-2</li> <li>3. 3-4</li> <li>4. 5-7</li> <li>5. Rather not say</li> <li>6. Skip this question</li> </ol>
7	<p>Are you currently in a relationship ?</p>	<ol style="list-style-type: none"> <li>1. Yes, seeing someone</li> <li>2. No, I'm single</li> <li>3. It's complicated</li> <li>4. Skip this question</li> </ol>
8	<p>Are you HIV positive?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Rather not say</li> <li>4. Skip this question</li> </ol>
9	<p>Before joining B-Wise, how often did you have discussions or interact with content about sexual topics?</p>	<ol style="list-style-type: none"> <li>1. A lot</li> <li>2. Somewhat</li> <li>3. Not much</li> <li>4. Never</li> <li>Skip</li> </ol>
10	<p>Reflecting on your participation on Bwise Facebook, if you had to guess how many times have you ever shared a post or question on the Bwise Facebook page about a view or question that you wanted people's input on??</p>	<ol style="list-style-type: none"> <li>a. Never</li> <li>b. 1-5 times</li> <li>c. 6-10 times</li> <li>d. More than 10 times</li> </ol>
11	<p>Reflecting on your participation in Bwise Facebook, if you had to guess how many threads on the Bwise Facebook page do you think you have ever commented on??</p>	<ol style="list-style-type: none"> <li>a. Never</li> <li>b. 1-5 times</li> <li>c. 6-10 times</li> <li>d. More than 10 times</li> </ol>
12	<p>Looking back 6 months ago, how frequently would you say you visited Bwise Facebook page?</p>	<ol style="list-style-type: none"> <li>a. Everyday</li> <li>b. Once or twice a week</li> <li>c. Every other week</li> <li>d. Once a month</li> </ol>

13	<p>Thinking about all of the posts that you've read on the B Wise facebook page, how much do you agree with the following statement? The posts on B Wise are normally relevant, interesting and were useful for my sexual health needs?</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. Skip this question</li> </ol>
14	<p>Thinking about all of the comments that you've seen other users make on the B Wise Facebook page, how much do you agree with the following statement? Other users' comments on B Wise are normally relevant, interesting and were useful for my sexual health needs?</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. Skip this question</li> </ol>
15	<p>How good a job do you feel you are doing in taking care of your health?</p>	<ol style="list-style-type: none"> <li>1. Excellent</li> <li>2. Very Good</li> <li>3. Good</li> <li>4. Fair</li> <li>5. Poor</li> <li>6. Skip this question</li> </ol>
16	<p>The last time you had sex, did you or your partner do or use something to avoid or delay getting pregnant?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Don't remember</li> <li>4. Haven't had sex yet</li> <li>5. Skip this question</li> </ol>
17	<p>Did you use a condom last time you had penetrative sex?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. Skip this question</li> </ol>
18	<p>How many sexual partners did you have over the last month?</p>	<ol style="list-style-type: none"> <li>1. None</li> <li>2. One</li> <li>3. Two to three</li> <li>4. More than three</li> <li>5. Skip this question</li> </ol>
19	<p>Have you ever been tested for Sexually Transmitted Infections (STIs) and HIV?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. Skip this question</li> </ol>



20	Over the past 12 months, do you think that your knowledge about the importance of using condoms has changed?	<ol style="list-style-type: none"> <li>1. Yes, improved a lot</li> <li>2. Yes, improved a little</li> <li>3. Stayed the same</li> <li>4. It's a little worse</li> <li>5. It's a lot worse</li> <li>I don't understand</li> <li>Skip</li> </ol>
21	Which of these has most influenced your knowledge about using condoms?	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> <li>8. No change</li> </ol>
22	Are you planning to have a child within the next year?	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. Maybe</li> <li>3. No</li> <li>4. Skip</li> </ol>
23	Do you now plan to use condoms more consistently than you did 12 months ago?	<ol style="list-style-type: none"> <li>1. Yes - a lot more</li> <li>2. Yes - a little more</li> <li>3. No - same</li> <li>4. No - a little less</li> <li>5. No - a lot less</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
24	Do you now plan to go for STI or HIV tests more often than you did 12 months ago?	<ol style="list-style-type: none"> <li>1. Yes - a lot more</li> <li>2. Yes - a little more</li> <li>3. No - same</li> <li>4. No - a little less</li> <li>5. No - a lot less</li> <li>6. I don't understand</li> <li>7. Skip</li> </ol>
25	Which of these has most influenced your plans to use condoms or test for STIs/HIV?	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> </ol>
26	Which of these has most influenced your attitudes about sexual relationships and interactions?	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-Wise chatbot / facebook</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> </ol>
27	How many times have you visited a clinic or other health facility for your sexual health since joining B-Wise? (We know it may be hard to remember, we'd appreciate your best guess)	<ol style="list-style-type: none"> <li>0. None</li> <li>1.</li> <li>2.</li> <li>3.</li> <li>4.</li> <li>5.</li> <li>6.</li> </ol>

		7. 8 or more I don't understand Skip
28	Do you think that your time on Bwise has changed how likely you are to visit a clinic or health facility for your sexual and mental health?	1. A lot more likely 2. Little more likely 3. No change 4. Little less likely 5. A lot less likely I don't understand Skip
29	Did you know you can use the Bwise WhatsApp chatbot to ask LoveLife (a counselling group) to call you back?	1. Yes, I used it and I got help 2. Yes, I used it but didn't get help 3. Yes, but I never used it 4. No, I didn't know that 5. No, but I never needed help I don't understand Skip
30	Other than using LoveLife, have you visited another mental or sexual health counsellor (since joining Bwise) and if so, how many times? (We know it may be hard to remember, we'd appreciate your best guess)	0. None - LoveLife was enough 1.1 time 2.2 times 3.3 times 4.4 times 5.5 times 6 or more 7. None but I needed to I don't understand Skip
31	Do you think that your time on Bwise has changed how likely you are to speak to a counsellor about your mental or sexual health?	1. A lot more likely 2. Little more likely 3. No change 4. Little less likely 5. A lot less likely I don't understand Skip
33a	That's great to hear! Would you mind telling us what cell phone number you used to register with the chatbot? We will only use this information to understand how people on Facebook engage with the chatbot. We will never share the cellphone number or use it for marketing.	Valid cell number (Or skip this question)

34	<p>Thank you for taking part in this survey. In order to be compensated R15 airtime, please indicate the South African cellphone number that you would like us to send the airtime to. Please note we will only be able to send the airtime to number with a +27 area code.</p>	<p>Valid cell number (Or skip this question)</p>
----	---	--

## 10.4) Facebook - Avert Content Survey Questions

Question number	Relevant questions	List of responses
1	Are you in South Africa?	1. Yes 2. No 3. Rather not say 4. Skip this question (IF NOT YES, THEN SEND MESSAGE EXPLAINING CANNOT BE A PART OF THE STUDY AS EXPLAINED IN CONSENT FORM)
2	What is your age?  Age in years	Strictly held between 18-24 years old  (IF BELOW 18 OR OVER 24, THEN SEND MESSAGE EXPLAINING CANNOT BE A PART OF THE STUDY AS EXPLAINED IN CONSENT FORM)
3	Roughly, how many months ago did you first engage with a B-wise Facebook post?	Strictly numeric (0 and greater)
4	What gender do you identify as?	1. Woman 2. Man

		<ul style="list-style-type: none"> <li>3. Non-binary or transgender</li> <li>4. Prefer not to say</li> </ul>
5	What is the total monthly income of your whole household?	<ul style="list-style-type: none"> <li>1. No income</li> <li>2. R1 - R400</li> <li>3. R401 - R800</li> <li>4. R801 - R1600</li> <li>5. R1 601 - R3200</li> <li>6. R3 201 - R6400</li> <li>7. R6 401 - R12800</li> <li>8. R12 801 - R25600</li> <li>9. R25 601 - R51200</li> <li>10. R51 201 - R102 400</li> <li>11. R102 401 - R204 800</li> <li>12. R204 801 or more</li> <li>13. Skip this question</li> </ul>
6	In the past 7 days, how many days did you go hungry?	<ul style="list-style-type: none"> <li>1. None</li> <li>2. 1-2</li> <li>3. 3-4</li> <li>4. 5-7</li> <li>5. Rather not say</li> <li>6. Skip this question</li> </ul>
7	Are you currently in a relationship?	<ul style="list-style-type: none"> <li>1. Yes, seeing someone</li> <li>2. No, I'm single</li> <li>3. It's complicated</li> <li>4. Skip this question</li> </ul>
8	Are you HIV positive?	<ul style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Rather not say</li> <li>4. Skip this question</li> </ul>
9	Before you started engaging with B-wise Facebook post(s), how often did you have discussions or interact with content about sexual topics?	<ul style="list-style-type: none"> <li>1. A lot</li> <li>2. Somewhat</li> <li>3. Not much</li> </ul>

		4. Never Skip
10	In the past 5 months, if you had to guess how many times have you ever shared a B-wise Facebook posts or question on the of any of the B-wise Facebook posts about a view or question that you wanted people's input on?	Strictly numeric (0 and greater) (or skip this question)
11	In the past month 5 months, if you had to guess how many of the B-wise Facebook posts do you think you have commented on?	Strictly numeric (0 and greater) (or skip this question)
12	If you had to guess, how many days this week, would you say you've seen B-wise Facebook posts?	Strictly numeric (0-7) (over skip this question)
13	In the past 5 months, have you seen any of this content?	Select from Avert ads that were posted during the duration of interest  <ol style="list-style-type: none"> <li>1. Tips for having great sex</li> <li>2. Things to know about Sexually Transmitted Infections (STIs)</li> <li>3. Family planning/Birth Control</li> <li>4. Caring for your mental health</li> <li>5. Safe sex including using condoms</li> </ol>
14	Thinking about all of the B-wise Facebook posts that you have read in the last 5 months; how much do you agree with the following statement?  The B-wise Facebook posts are normally relevant, interesting and were useful for my sexual health needs?	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. Skip this question</li> </ol>
15	To what extent do you believe the B-wise Facebook posts encourages safe sex practices among adolescents?	<p>Strongly Disagree</p> <p>Disagree</p> <p>Neutral</p> <p>Agree</p> <p>Strongly Agree</p>

16	<p>Thinking about all of the comments that you've seen other users make on the B-wise Facebook posts the last 5 months, how much do you agree with the following statement?</p> <p>Other users' comments on B-wise Facebook posts are normally relevant, interesting and were useful for my sexual health needs?</p>	<ol style="list-style-type: none"> <li>1. Strongly agree</li> <li>2. Agree</li> <li>3. Not sure</li> <li>4. Disagree</li> <li>5. Strongly disagree</li> <li>6. Skip this question</li> </ol>
17	<p>How good a job do you feel you are doing in taking care of your health?</p>	<ol style="list-style-type: none"> <li>1. Excellent</li> <li>2. Very Good</li> <li>3. Good</li> <li>4. Fair</li> <li>5. Poor</li> <li>6. Skip this question</li> </ol>
18	<p>The last time you had sex, did you or your partner do or use something to avoid or delay getting pregnant?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Don't remember</li> <li>4. Haven't had sex yet</li> <li>5. Skip this question</li> </ol>
19	<p>Did you use a condom last time you had penetrative sex?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. Skip this question</li> </ol>
20	<p>How many sexual partners did you have over the last month?</p>	<ol style="list-style-type: none"> <li>1. None</li> <li>2. One</li> <li>3. Two to three</li> <li>4. More than three</li> <li>5. Skip this question</li> </ol>
21	<p>Have you ever been tested for Sexually Transmitted Infections (STIs) and HIV?</p>	<ol style="list-style-type: none"> <li>1. Yes</li> <li>2. No</li> <li>3. Haven't had sex yet</li> <li>4. Skip this question</li> </ol>
22	<p>Over the past 5 months, do you think that your knowledge about the importance of using condoms has changed?</p>	<ol style="list-style-type: none"> <li>1. Yes, improved a lot</li> <li>2. Yes, improved a little</li> <li>3. Stayed the same</li> </ol>

		<p>4. It's a little worse</p> <p>5. It's a lot worse</p> <p>I don't understand</p> <p>Skip</p>
23	Which of these has most influenced your knowledge about using condoms?	<p>1. Internet / social media</p> <p>2. B-wise chatbot / facebook</p> <p>3. My friends / partner</p> <p>4. At school / university</p> <p>5. Health facility</p> <p>6. TV / radio</p> <p>7. Other</p> <p>8. No change</p>
24	Are you planning to have a child within the next year?	<p>1. Yes</p> <p>2. Maybe</p> <p>3. No</p> <p>4. Skip</p>
25	Do you now plan to use condoms more consistently than you did 5 months ago?	<p>1. Yes - a lot more</p> <p>2. Yes - a little more</p> <p>3. No - same</p> <p>4. No - a little less</p> <p>5. No - a lot less</p> <p>6. I don't understand</p> <p>7. Skip</p>
26	Do you now plan to go for STI or HIV tests more often than you did one month ago?	<p>1. Yes - a lot more</p> <p>2. Yes - a little more</p> <p>3. No - same</p> <p>4. No - a little less</p> <p>5. No - a lot less</p> <p>6. I don't understand</p> <p>7. Skip</p>
27	Which of these has most influenced your plans to use condoms or test for STIs/HIV?	<p>1. Internet / social media</p> <p>2. B-wise chatbot / facebook</p> <p>3. My friends / partner</p> <p>4. At school / university</p> <p>5. Health facility</p> <p>6. TV / radio</p>

		7. Other
28	Which of these has most influenced your attitudes about sexual relationships and interactions?	<ol style="list-style-type: none"> <li>1. Internet / social media</li> <li>2. B-wise chatbot / facebook/ B-wise Facebook post</li> <li>3. My friends / partner</li> <li>4. At school / university</li> <li>5. Health facility</li> <li>6. TV / radio</li> <li>7. Other</li> </ol>
29	How many times have you visited a clinic or other health facility for your sexual health since your first interaction with B-wise Facebook post? (We know it may be hard to remember, we'd appreciate your best guess)	<ol style="list-style-type: none"> <li>0. None</li> <li>1.</li> <li>2.</li> <li>3.</li> <li>4.</li> <li>5.</li> <li>6.</li> <li>7.</li> <li>8 or more</li> <li>I don't understand</li> <li>Skip</li> </ol>
30	Do you think that your interaction with B-wise Facebook post(s) has changed how likely you are to visit a clinic or health facility for your sexual and mental health?	<ol style="list-style-type: none"> <li>1. A lot more likely</li> <li>2. Little more likely</li> <li>3. No change</li> <li>4. Little less likely</li> <li>5. A lot less likely</li> <li>I don't understand</li> <li>Skip</li> </ol>
31	Did you know you can use the B-wise WhatsApp chatbot to ask LoveLife (a counselling group) to call you back?	<ol style="list-style-type: none"> <li>1. Yes, I used it and I got help</li> <li>2. Yes, I used it but didn't get help</li> <li>3. Yes, but I never used it</li> <li>4. No, I didn't know that</li> <li>5. No, but I never needed help</li> <li>I don't understand</li> <li>Skip</li> </ol>



32	Other than using LoveLife, have you visited another mental or sexual health counsellor (since joining B-wise) and if so, how many times? (We know it may be hard to remember, we'd appreciate your best guess)	<p>0. None - LoveLife was enough</p> <p>1.</p> <p>2.</p> <p>3.</p> <p>4.</p> <p>5.</p> <p>6 or more</p> <p>7. None but I needed to</p> <p>I don't understand</p> <p>Skip</p>
33	Do you think that your interaction with B-wise Facebook post(s) has changed how likely you are to speak to a counsellor about your mental or sexual health?	<p>1. A lot more likely</p> <p>2. Little more likely</p> <p>3. No change</p> <p>4. Little less likely</p> <p>5. A lot less likely</p> <p>I don't understand</p> <p>Skip</p>
33a	That's great to hear! Would you mind telling us what cell phone number you used to register with the chatbot? We will only use this information to understand how people on Facebook engage with the chatbot. We will never share the cellphone number or use it for marketing.	Valid cell number (Or skip this question)
35	Thank you for taking part in this survey. In order to be compensated R15 airtime, please indicate the South African cellphone number that you would like us to send the airtime to. Please note we will only be able to send the airtime to number with a +27 area code.	Valid cell number (Or skip this question)

## 11) Appendix C - Additional report detail

### 11.1) WhatsApp Chatbot Pre-Post Study

## Baseline and Endline Survey Enrolment

Of the 1999 users invited to participate over this period, 1295 (65%) consented to and completed the baseline survey, receiving R30 airtime upon completion<sup>25</sup> as compensation for their time. The baseline survey gathered information on users' registration levels of; i) SRH knowledge, ii) SRH attitudes, iii) psychological capacity, iv) psychological resilience, and v) SRH behaviours, as well as several demographic controls<sup>26</sup>.

All baseline users were then invited to complete the 75-question, voluntary endline survey 5 months later. Between November 23rd-29th, 302 of the 1295 baseline completers consented to and completed the endline survey, receiving R50 airtime upon completion as compensation for their time. Shortly, thereafter it was discovered that, due to a coding error, 10 questions had not been presented to the respondents<sup>27</sup>. As such the endline was redeployed to all remaining baseline completers, gathering a further 200 endline responses between 18 December 2023 and 16 January 2024. To estimate changes in users' barriers and outcomes over time, all of the same questions that were asked in the baseline were included in the endline regarding; i) SRH knowledge, ii) SRH attitudes, iii) psychological capacity, iv) psychological resilience and v) SRH behaviours. The endline asked a few demographic questions to capture exogenous changes in the sample over time, as well as a wide range of questions about users' experience of the platform and their subjective opinion of the effect of the WhatsApp chatbot<sup>28</sup>.

Ethical approval for both surveys was obtained from the Pharma Ethics Committee before data collection, ensuring that the surveys were voluntary and sufficiently anonymous. Both surveys and their invites were conducted entirely over the WhatsApp chatbot, with users' incremental responses securely stored within Reach's Amazon S3 data lake. All data remained within South African borders. This data was collected and anonymised by Reach's data science team, before being analysed with Stata 14.

## Endline's Demographic Representativity of Platform Users

Column 1 of Table A5 below, presents summary statistics of all onboarding questions for all users accessing the platform who are aged 15-24 that subscribed to push-notifications (ie. the target group YAL could have an effect on), with Columns 2 and 3 then presenting these same summary statistics of registration data for users that are not captured in the endline survey and those that are captured in the endline. Columns 4 and 5 then report the difference between these groups and the p-value associated with a paired t-test. For ease of interpretation, both here and in all tables of

<sup>25</sup> While this may have presented selection effects, comparisons between users enrolling in the baseline sample and those declining the baseline invite show no statistically significant differences in income. Implying that this compensation may not have differentially incentivised users.

<sup>26</sup> To see all baseline survey instruments, refer to Appendix B

<sup>27</sup> 4 questions related to self-reported clinic-seeking behaviours, 3 questions related to counselling-seeking behaviours, 2 questions related to users' location level data and 1 related to SRH knowledge. These are questions 206-209, 212-214, 105, 107 and 1002 in Appendix B.

<sup>28</sup> To see all endline survey instruments, refer to Appendix B

this report, stars in column 5 represent statistical significance with p-values of p<0.1, p<.05 and p<0.01 represented as \*, \*\* and \*\*\* respectively.

Table A5: Summary statistics of registration data for subscribed targeted-platform and endline users

Variable	All platform users (1)	Did not do endline (2)	Did endline (3)	Difference (3-2)	P-value (3 vs 2)
<b>Demographics</b>					
Age (restricted to 15-24)	20.40	20.39	20.64	0.24	**
Under 20 of those in target age	37%	37%	34%	3%	
Female as opposed to male	55%	55%	69%	14%	***
Relationship status is “in relationship” at registration	55%	55%	59%	5%	**
Relationship status is “complicated” at registration	20%	20%	22%	2%	
Relationship status is “single” at registration	25%	25%	19%	-6%	***
Household income at registration	R2084.40	R2081.12	R2168.27	R87.15	
<hr/>					
Total users	32422	31922	502		

Column 5 shows that there are statistically significant differences between the average platform user (within the target age group) and those users that go on to complete the endline survey. In terms of average age, the two groups differs by only 0.24 years. Given that the endline is restricted to users 18 years or older, this indicates the relative lack you minors on the platform. Indeed, there is no statistically significant difference between the groups in terms of the proportion of users under 20 years old (36% and 34% respectively). There is no statistically significant difference in average household income between general target users and those completing the endline survey, with the both groups having a median response of no income. While the income variable should be seen with caution given the measurement error points raised earlier, there is not presently any evidence to indicate that income level affects users enrolment in the surveys.

While the evidence suggests that the endline survey is relatively representative of the platform’s target users in terms of age and income, there are significant differences between the groups in terms of gender and relationship status. Of users completing the endline, 69% are women, whereas 55% of target platform users are women (p<0.01). Additionally, users in the endline sample are slightly more likely to be in a relationship at registration (59% vs 55%, p<0.05) and less likely to be single (25% vs 19%, p<0.01).

Finally, we can decompose these demographic differences between users that chose to not complete the baseline survey (self-selection) and users who then did not complete the endline survey (attrition). Table A6 presents summary statistics similar to Table A5, except that it compares users that have only completed the baseline survey against those that also completed the endline survey. Additionally, it provides these comparisons for some other demographic fields not covered in onboarding.

Table A6: Demographic analysis of attrition from baseline to endline

Variable	All baseline respondents (1)	Did baseline only (2)	Did baseline and endline (3)	Difference (2-3)	P-value (2 vs 3)
<b>Demographics</b>					
Age	20.67	20.66	20.65	-0.02	
Below 20 years old	33%	32%	34%	2%	
Female as opposed to male	61%	55%	69%	14%	***
HIV positive	4 <sup>29</sup> %	4%	3%	-1%	
Relationship status is (in relationship) at baseline	56%	54%	59%	6%	**
Relationship status is (complicated) at baseline	21%	20%	22%	2%	
Relationship status is (single) at baseline	23%	26%	19%	-7%	***
Total household income at baseline	R2275.96	R2344.60	R2168.27	R-176.33	
Total users	1295	793	502		

Column 5 indicates that there has been no selective attrition from baseline to endline, in terms of age or average household income<sup>30 31</sup>. Whereas, users that go on to complete the baseline are 6 percentage points more likely to have been in a relationship, 7 percentage points less likely to have been single, and 14 percentage points more likely to identify as female at registration than those that just complete the baseline survey. These account for almost all of the differences noted in Table A5, implying that these differences are largely due to selective attrition as opposed to self-selection into the baseline.

<sup>29</sup> Note that this is very close to the national estimate of 5% of youth being HIV positive (NYDA, 2022)

<sup>30</sup> With the median response for both also being no income.

<sup>31</sup> Unfortunately, no household size data was gathered at registration, as such income for all registration data is captured only at the household level.

### Endline's Representativity of young people's SRH needs (population and platform)

Referring back to Section 3.3, columns 2, 3, 4 and 5 of Table 2 compare the starting SRH needs of users that only took the baseline and those users that then went on to also complete the endline. Of the 15 SRH needs considered included in Table 2, 5 variables are statistically different beyond the 5 percent level, 2 variables show statistical differences at the 10 percent level, and 8 variables show no statistical evidence of systematic attrition.

In terms of barriers, there are no statistical differences in the proportion of users with insufficient psychological capacity for any of the 3 variables between the two groups. However, a higher proportion of users in the endline sample appeared to have low initial SRH knowledge scores (23%) than those that did not complete the endline (12%). This indicates that of the endline sample is skewed toward users with poor initial SRH knowledge scores. Given the variability of SRH knowledge measures in the literature it is not possible to say whether this endline sample is more representative of the national population or not, however, given this program's focus of improving SRH knowledge as a primary goal, having bias towards users with low SRH knowledge appears useful.

In terms of relevant attitudes, the proportion of users with poor sex positivity or beliefs about consent remains constant between the groups, however, the proportion of users with body image issues or poor gender attitudes does differ significantly. For body image and gender attitudes, respectively, there is a 6 and 7 percentage point difference between users completing the endline and those that do not ( $p < 0.05$  and  $p < 0.01$ ). This indicates that the endline sample is perhaps slightly biased towards users with poor initial body images, and slightly biased away from users with poor initial gender attitudes, as compared to the best proxy of the target-platform user (the baseline).

Lastly, Table 2 also shows suggestive evidence that endline users have slightly better baseline scores regarding the behavioural and persistence-based outcomes of interest. Users that continue to endline are slightly more likely at baseline to; have used any form of contraception at their last sexual encounter, had 1 or fewer total sexual partners in the last month<sup>32</sup>, and have ever tested for an STI or HIV. However, only one of these is significant at the 5 percent level, and the actual differences in proportions when compared to the overall proportions are small in all cases (between 3 and 6 percentage points). For persistence there is no statistically significant difference in self-esteem, however, endline respondents are significantly less likely to have had predominantly external loci of control at baseline (36% vs 44%  $p < 0.01$ ).

Recall that the baseline survey serves as the only proxy of the subscribed target populations' SRH needs on the platform as well as the demographic biases of the endline sample.

### Estimation strategy

In selecting between estimation strategies, this study ultimately opted to employ mixed-model regressions to perform this more precise estimation. The decision to

---

<sup>32</sup> Though this difference only exists in the binary form, not when considering users' average number of sexual partners.

estimate the effect of time through mixed-model regressions is motivated by several considerations.

- Pooled OLS fails to recognise the paired nature of the data, such that errors between a user's result at baseline and endline are not independent and results are biased (Ghosh, 2022).
- Individual fixed effects / first differencing causes the time variable to fall out of the model making estimating the effect of the program through the coefficient on time non-viable (Imai and Kim, 2019).
- The preferred choice in the literature is then between ANCOVA and repeated measures mixed-model regressions (UCLA: Statistical Consulting Group, 2022). Both are essentially linear estimators, that account for variation both within subjects and between subjects. One benefit of mixed-model regressions is that they do not assume any particular within-subject covariance structure, whereas ANCOVA requires that the within-subject covariance matrix be "compound symmetrical" (UCLA: Statistical Consulting Group, 2022)<sup>33</sup>

## 11.2) Facebook Cross-sectional Study

### Sample size and methods

Randomly selected registered users of the B-Wise WhatsApp platform could view the invitation to the Facebook survey. Reach limited the number of participants to a total of 200 respondents. Considering the continuous outcome of users' frequency of engagements, 167 engagements were deemed sufficient for the survey to estimate within a 7% margin of error (with a 95% confidence level) of the true parameter for an assumed population of 20,000 B-Wise Facebook users, following the methodology suggested by Dunn & Clark. Therefore, the inclusion of 178 respondents provided the study with adequate precision, allowing for a slight buffer to account for minor unforeseen circumstances. The survey was in English, given that the B-Wise platform was strictly in English. A total of 178 usable surveys were received.

### Data tools

We administered two questionnaires –one for Facebook page members and another for participants who have experienced B-wise paid posts. While both questionnaires shared similar content, they were tailored to capture these two groups' specific experiences and perspectives. By administering these tailored questionnaires, we aimed to gather insights into the distinct experiences and perceptions of Facebook page members and those exposed to paid posts, contributing to a comprehensive understanding of the impact of the B-Wise Facebook platform on its audience.

### Data Collection

A short invitation to participate in a Google Forms-based survey was shared to randomly selected registered B-Wise WhatsApp members. Interested users were required to confirm their age (18-24) and residency in South Africa before undergoing a

---

<sup>33</sup> I.e. that there is a shared variance at all periods, and that variance is constant across subjects at different time periods.

comprehensive consent process. The full online survey consent procedure explicitly stated that there were no consequences for choosing not to participate. The survey allowed 178 individuals aged 18-24 to participate, taking approximately 5 minutes to complete and offering R15 airtime as compensation. This survey was conducted entirely on Google Forms, incurring minimal cost to the user. To be eligible for invitation to the study, a user registering on the YAL/B-Wise platform must meet all of the following requirements:

- Be living within South Africa
- Be 18-24 years old
- Have voluntarily agreed to receive regular messaging from the B-Wise/YAL platform.

**Data analysis**

The unit of analysis was the individual respondent (18-24 years old). All data was securely gathered through the password-protected Google Forms service. Only this study's Principal Investigators (PIs) had access to these responses and only downloaded and stored the data in CSV format onto password-protected and secure local computers. Access to and analysis of the data was conducted only by the listed team members in this study.

After getting the data, we loaded it into STATA 18 for analysis. We used frequencies and proportions to summarize categorical data, looking at demographic details and exposure to SRH before B-wise prevalence rates. Using the chi-square test, we compared categorical exposure variables with outcome variables, showing results as Odds Ratios (OR) with 95% confidence intervals (CI). We set significance levels at 5% ( $P < 0.05$ ).

### 11.3) Qualitative Focus group discussions

**Data management and analysis**

All IDIs and FGDs were recorded and transcribed over a voice recorder. The data was collected in Sepedi, Setswana, Isizulu, and/or simple English. Confidentiality of data was maintained to ensure that ethical standards were maintained by de-identifying the data. Hence, all reports are de-identified, and all participant information, such as consent forms, is stored in locked file cabinets in areas with access limited to staff. Data is stored on a One-Drive account that is password protected, with access given to limited staff in the project.

Thematic Analysis (TA) was used to examine patterns and organise data into themes. The process undertook the following steps: familiarization of content, generation of initial themes, refining themes, and finally definition of and naming themes. The evaluation team used TA to understand the findings on the effectiveness of YAL's mobile health intervention, document its impact, and share key recommendations.

**Data tools**

WhatsApp was chosen as the most preferred Application (hereafter referred to as the App) since most participants already use it to engage with the Chatbot, and it is affordable regarding data consumption. When participants struggled to connect to the discussions, researchers sent the discussion guiding questions, and the participants returned voice notes with the responses. The consultants asked some follow-up questions, and they responded to those, providing examples or clarity as needed.

**Sampling**

A two-stage sampling strategy was utilised. In the first stage, Reach launched a location survey to better understand the distribution of the platform users. Before the formal invitation to the focus groups, all eligible users on the platform were sent a short, four-question preparatory survey over the B-Wise WhatsApp Chatbot. The survey only captured users' a) province of residence, b) city/town of residence, c) type of living area and d) interest in participating in focus groups. After the location survey, Reach shared a dataset about the current users interested in the FGDs. The dataset had the following information:

- The name the user gives the WhatsApp chatbot as a personalisation in onboarding. This excluded the real names of the users.
- The user's age and relationship status at registration
- Gender identity
- Opt-out status
- Total number of messages the user sent YAL; and,
- The date a user joined the line, the most recent date a user sent YAL messages, and the last date a user's contact information was updated by our system.

The second stage of sampling involved purposive sampling of the participants. The location survey identified 102 AYPs from Gauteng who indicated interest in participating and 46 from KwaZulu Natal. The selected participants were invited to the study, and their contact details were shared with the consultants. The consultants contacted them to explain the study further and to schedule a time for an individual interview or focus group discussion. The study participants were divided into the following groups: (a) two groups of males aged between 20 – 24 years old, (b) two groups of females aged between 19 – 24 years old, (c) two groups of 19 – 23 years old individuals mixed along sex lines and (d) three groups individuals aged between 19 – 24 years old mixed sex lines. This group included youth who had used Chatbot less than 30 times regardless of when they had joined and those who had used Chatbot less than 50 times regardless of when they had joined. All focus groups included youth who were in or out of school, those who were employed, and those who were not in school, not employed, or not in training. Table A7 and A8 summarise the participants by data collection methods.



Table A7: Summary of FGD participants of the study

Sex of Group	Age range	Data collection method	No. of groups	Total no. of participants
Young men	20 - 24	WhatsApp	2	9
Young women	19 - 24	WhatsApp	2	11
Combined, young men and young women	19 - 23	In Person	2	12
Combined, young men and young women	19 - 24	WhatsApp	3	12

Table A8: Overall summary of participants by data collection method

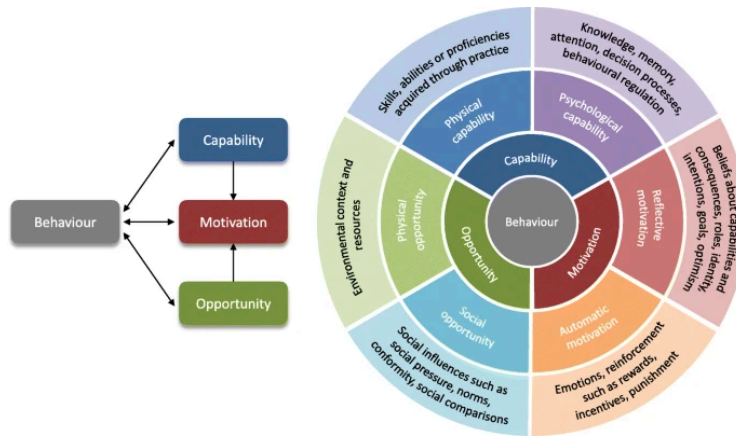
	Online (WhatsApp or Virtual FGD/IDI)	In-person	IDI
Young men	14	5	0
Young women	18	7	9
Total	32	12	9

## 11.4) The YAL TOC and instrument construction

### The COM-B model of behaviour change

To improve the sexual, reproductive, and mental health of youth in South Africa, YAL's program was designed based on the COM-B behavioural change model (see Figure A1), which posits that interventions that impact individuals' capability, opportunity, and motivation can lead to improved behaviours. Capability refers to an individual's knowledge, skills, and ability to engage in the behaviour. Opportunity refers to factors that enable individuals to execute a specific behaviour. Motivation refers to an individual's disposition to want to do the behaviour instead of treating it as a taxing necessity (West and Michie, 2020).

Figure A1: COM-B behavioural Model (McDonagh et al., 2017)



Ultimately, the YAL program aims to improve both young people’s SRH persistence measures (as a psychological construct) and, subsequently, their SRH outcomes - getting youth to adopt behaviours that serve their SRH needs where they can do so. Each of the interventions detailed in the “YAL Theory of Change” is ultimately in service of improving one of the 4 outcomes of interest: i) SRH behaviours, ii) SRH persistence, iii) uptake of SRH services, and iv) information on SRH services.

## Instrument definitions

Figure A2 below lists the conceptual factors selected by Reach for assessment in this pilot phase of the YAL program. It collects the factors under the categories of knowledge, attitudes, and psychological capacity and identifies each construct as either relevant to Motivation, Opportunity, or Capability. Note that all knowledge and attitude variables are thought to most affect motivation, while psychological resources are considered capabilities. No opportunities are listed since it is not believed that the mobile intervention can affect opportunities (except for the service finder and LoveLife tools which create opportunities to access SRH services). However, in general, which category a construct falls under can be seen from various perspectives.

In order to assess the validity of the COM-B based TOC, each conceptual factor is proxied by an indicator (a number of questions that attempt to estimate the given construct for a particular user). The figure lists the total number of questions making up each instrument, and provides a short description of each question. In order to see the exact wording of each question, as well as its available responses, please use the associated question reference to the baseline survey questions presented in Annex B.

**Figure A2: Description of TOC constructs and relevant indicators**

Construct	COM-B classification	Instrument questions
<i>Knowledge</i>		
SRH knowledge	Motivation	<ul style="list-style-type: none"> <li>- Aware that condomisation reduces risk of STIs (Q1001)</li> <li>- Aware that sexual exclusivity reduces risk of STIs (Q1002)</li> <li>- Selects an effective form of contraception (Q1009)</li> </ul>
<i>SRH Attitudes</i>		
Body image	Motivation	<ul style="list-style-type: none"> <li>- Feel good about one's self (Q601)</li> <li>- Feel good about one's body (Q602)</li> </ul>
Sex positive attitudes	Motivation	<ul style="list-style-type: none"> <li>- Believe one's sexual needs and desires are important (Q1004)</li> <li>- Believe it is important to focus on both own and partner's pleasure during sex (Q1005)</li> <li>- Can insist on condom use (Q1003)</li> <li>- Expects to enjoy sex (Q1006)</li> </ul>
Gender equality within sexual relationships	Motivation	<ul style="list-style-type: none"> <li>- Believes there are times where violence against women is justified (Q1101)</li> <li>- Believes it's a woman's responsibility to avoid getting pregnant (Q1102)</li> <li>- Believe partners should decide together on preferred form of contraception (Q1103)</li> </ul>

Valuation of consent in sexual relationships	Motivation	<ul style="list-style-type: none"> <li>- Believe that men share the responsibility of children (Q1104)</li> <li>- Believes acceptable to force sex in a presented vignette (Q1201)</li> <li>- Can express disinterest in having sex (Q1202)</li> </ul>
<i>Psychological Capability</i>		
Depression/anxiety (PHQ-4)	Capability	<ul style="list-style-type: none"> <li>- Frequency of feeling nervous, anxious or on edge (Q601)</li> <li>- Frequency of unable to stop worrying (Q602)</li> <li>- Frequency of feeling down, depressed or hopeless (Q701)</li> <li>- Frequency of little pleasure in doing things (Q702)</li> </ul>
Alcohol and substance misuse	Capability	<ul style="list-style-type: none"> <li>- Felt guilty about one's drinking or drug use (Q1301)</li> <li>- Been annoyed by people criticising one's drinking or drug use (Q1301)</li> <li>- Needed a drink or to use drugs first thing in the morning (Q1301)</li> <li>- Ever felt guilty about one's drinking or drug use (Q1301)</li> </ul>
Social connectedness	Capability	<ul style="list-style-type: none"> <li>- Frequency with which can contact to talk to when has a worry or problem (Q501)</li> </ul>

## 11.5) Activity 1 Supplementary Analysis

### Output 1 - The COM-B model of behaviour change

Regarding the provision of content, the platform set a goal of 25% of users to link between the WhatsApp chatbot and the Facebook channels (SMART Goal 2). Unfortunately, due to Meta's privacy policies, individual-level data on traffic to the B-Wise Facebook page is unavailable. As a next best approximation, the WhatsApp endline survey gathered information on users' reported awareness and use of the various components of the YAL platform, acknowledging the limitations of the endline sample as relatively female and more likely to be in a relationship at registration. Table A9 reports the modal and second most frequent response to two short questions of users engagement with the Facebook page as well as their perceptions of the WhatsApp content (analysis in section Intermediary Outcome 1.1)

Table A9: Endline users' feedback on the platform and features

Variable	Variable description and total choices	Modal response and relative frequency	2nd most frequent response	Total respondents (n)
<i>Facebook page</i>				
Know about B-Wise Facebook page	Binary	Yes (71%)	No (29%)	498
Frequency of B-Wise Facebook page use	Categorical: 4	Not much (49%)	Weekly (33%)	338 <sup>34</sup>
<i>WhatsApp content</i>				
Content related to your sexual needs	Likert: 5	Very related (46%)	Related well (27%)	493
Content was interesting when it related	Likert: 5	Very interesting (57%)	Quite interesting (25%)	498
Content was useful for managing your SRH and relationship needs	Likert: 5	Extremely useful (63%)	Quite useful (25%)	501

Table A9 shows that 71% of endline respondents from the WhatsApp chatbot indicate knowing about the B-Wise Facebook page, and 67% indicated having ever visited the Facebook page. From the qualitative study, the research team found that most participants started using the platform after seeing it advertised on Facebook. These findings indicate that for these groups there was a clear linkage between the use of the chatbot and awareness of the Facebook page, and vice versa. Unfortunately, when analysing whether that visit then translated into regular engagement with the Facebook page, the research team found that for users that had visited the page at least once, 5% never revisited the page and 49% visited the page "not much". This finding was to be anticipated, given the difficulties the project has faced with maintaining regular content sharing and moderation on the B-Wise page across the five months. Surprisingly, however, 33% of users who have ever visited the page indicate that they do so weekly (the highest response possible), and another 18% visited the page monthly. Since the frequency of posts by the page was low for much of the intervention period, these users may be visiting the page as a first step to directly messaging the B-Wise Facebook team (a means of interaction that the B-Wise team indicates has remained relatively active).

**Intermediary Outcome 1.2 - Improved knowledge and attitudes regarding contraceptives, sexual health, HIV and STI, sexuality, and healthy relationships.**

Table A10 below reports on changes in the proportion of users with substantial barriers to SRH (low knowledge or poor attitudes) from baseline to endline for all 502 users that completed both the baseline and endline surveys. Respectively, columns 1, 2 and 3

<sup>34</sup> Given that users indicated any knowledge of the Facebook page

report baseline values, endline values, and the absolute difference between these for the sample, while columns 4 and 5 compare these proportions, reporting the odds ratio and exact p-value from an associated McNemar test.

Table A10: Changes in intermediary outcomes from baseline to endline

Variable	(1) Baseline	(2) Endline	(3) Absolute Change	(4) Odds ratio	(5) P-value
<i>Knowledge</i>					
Low knowledge (Less than 1.5 correct on 3 SRH knowledge questions)	23%	17%	6%	0.51	***
<i>Attitude</i>					
Poor body image	37%	22%	-14%	0.29	***
Poor sex positivity	16%	15%	-1%	0.88	
Poor gender	33%	31%	-3%	0.79	
Poor consent valuation	22%	14%	-8%	0.44	***

This would indicate that there is a significant decrease in the proportion of users with low SRH knowledge, large decreases in the proportion of users with a poor body image or poor valuation of consent in sexual relationships, but no observable change in the proportion of users with poor sex-positive or gender equality attitudes. However, it is important to consider what other factors may have changed over the same time. Of all the demographic variables gathered in the pre-post study, only household income and relationship status are typically considered variable over time. Table A11 is similar in structure to Table A10 but reports on changes in these demographic variables from baseline to endline for all 502 users that completed both the baseline and endline surveys. Additionally, where household income is measured in continuous form, for this one variable, columns 4 and 5 report the t-statistic and associated p-value from a paired t-test of the mean of income in the two periods.

Table A11: Changes in SRH demographics and capacity from baseline to endline

Variable	(1) Baseline	(2) Endline	(3) Absolute Change	(4) Odds ratio /T-stat	(5) P-value
<i>Demographic</i>					
Relationship status is (in relationship)	60%	69%	9%	1.96	***
Relationship status is (complicated)	22%	13%	-9%	0.47	***
Relationship status is (single)	19%	19%	0%	1	

Total household income	2168	4155	1987	3.56	***
Household income per capita (inferred) <sup>35</sup>	832.07	1171.11	339.05	0.85	

Table A11 indicates that between baseline and endline, users' relationship statuses change slightly, with the proportion of users in a committed relationship at endline increasing by 9 percentage points from baseline ( $p < 0.01$ ), while the proportion of users in "complicated situations" decreases by the same amount ( $p < 0.01$ ). Given that a lot of the content is focused on educating users on healthy relationships, it is plausible that these changes are at least partially related to the engagement with the platform, as such, these changes should be thought of as at least partly endogenous to the intervention.

Additionally, Table A11 indicates that average household income increases substantially from baseline to endline ( $p < 0.01$ ), although no statistically significant difference is seen once accounting for household size<sup>36</sup>. Given that YAL does not look to directly affect factors that may be thought to increase household income, this should be seen as an exogenous shock that needs to be accounted for in the estimation of the changes due to time. To control for the variation of income over the intervention period, as well as other time-invariant demographic factors that may be associated with differential trends for sub-groups, the research team believes that a paired subjects mixed model linear regression<sup>37</sup> is more appropriate than a straight McNemar test for estimating the effect of time on each outcome variable of interest<sup>38</sup>.

<sup>35</sup> Unfortunately, no household size data was gathered at baseline, as such income per capita for baseline is made by assuming that household size is constant across time for all users. However, it is likely reasonable to assume that household size is constant for the vast majority of houses over just 5-6 months.

<sup>36</sup> A possible explanation for the significance on total household income but not on household income per capita is that the largest increases were reported for users in larger households, such that once adjusted to the per person scale and balanced against little change in household income in smaller households, this culminated in an insignificant increase in income. See footnote above on why the income and income per capita variables should be seen with some caution.

<sup>37</sup> Gomila (2021) demonstrates that employing linear relationships for causal inference on binary outcomes is often unbiased and favourable, with the additional benefit that the coefficient on linear regressions is easily interpretable.

<sup>38</sup> For a justification of the application of the paired subjects mixed model regressions, please see section 3.2 - Estimation Strategy