

©Copyright 2019

Trevor N Perrier

Connecting End Users to Domain Experts With Universal Mobile Phone Services

Trevor N Perrier

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Richard Anderson, Chair

Kurtis Heimerl

Keshet Ronen

Program Authorized to Offer Degree:
Paul G. Allen School of Computer Science & Engineering

University of Washington

Abstract

Connecting End Users to Domain Experts With Universal Mobile Phone Services

Trevor N Perrier

Chair of the Supervisory Committee:

Professor Richard Anderson

Paul G. Allen School of Computer Science & Engineering

The potential for communication services available on any mobile phone to engage users, disseminate information, and facilitate behavior change has been well documented. Designing using universal applications based on SMS and voice services enables platforms that reach and empower marginalized segments of the global population. The majority of mobile messaging services for development are either one-way push messaging or use fully automated bi-directional hierarchy based bots. In this thesis, I present the design, deployment and evolution of mWACH – a semi-automated bi-directional SMS platform for global health interventions. This novel system mediates personalized conversations between patients and trained medical professionals through a personalized schedule of automated messages. In collaboration with UW Global Health and Kenyatta National Hospital in Nairobi mWACH has been used in five different projects at seven Kenyan clinics. Since 2013 over 120 thousand automated SMS messages have been sent and we have received over 35 thousand messages from participants with study staff sending an additional 25 thousand messages. This thesis contributes an in depth quantitative analysis of system use and engagement over time and across projects. Results from the system analysis and messaging content will help inform the design of future systems looking to engage patients on health messaging platforms.

TABLE OF CONTENTS

	Page
List of Figures	iii
List of Tables	vii
Glossary	ix
Chapter 1: Introduction	1
1.1 Contributions	2
1.2 Overview of Thesis	4
Chapter 2: Background	5
Chapter 3: The mWACH Project Suite	15
3.1 Overview of mWACH Projects	16
3.2 Evolution of mWACH Projects	21
3.3 Properties of mWACH Projects	27
Chapter 4: Encouragement: mWaCH V1	33
4.1 Overview	33
4.2 Introduction	34
4.3 A Hybrid Computer-Human SMS System	35
4.4 Evaluation	41
4.5 Findings	44
4.6 Discussion	57
4.7 Conclusion	59
Chapter 5: mWaCH XY: Male Partner Engagement	60
5.1 Overview	60

5.2	Introduction	60
5.3	Related Work	62
5.4	Methodology	63
5.5	System Design	68
5.6	Participant characteristics	72
5.7	Measuring Male Engagement	75
5.8	Discussion	83
Chapter 6:	Longitudinal Analysis of Engagement in mWACH X	86
6.1	Baseline Demographics	88
6.2	System Statistics and Analysis	97
6.3	From The Platform’s Perspective	116
Chapter 7:	USSD: The Third Universal App	121
7.1	Overview	121
7.2	Introduction	122
7.3	Universal GSM Applications	124
7.4	USSD Library Design	134
7.5	USSD Use Cases and Personas	136
7.6	Discussion and Future Work	140
7.7	Conclusion	141
Chapter 8:	Conclusion	142
8.1	Contributions	142
8.2	Final Remarks	143
8.3	Acknowledgments	144
Bibliography	146

LIST OF FIGURES

Figure Number	Page
3.1 Depiction of messaging tracks and content during pregnancy, postpartum, and post one year after delivery.	19
3.2 Screen shot of main task oriented dashboard in the updated mWACH platform.	23
3.3 Screen shot of a participants detail page in the updated mWACH platform with the first incoming message unanswered.	23
3.4 Message totals throughout the week in 30 min increments for participant and nurse messages in mWACHX.	26
3.5 Bi-directional flow of SMS messages through the mWACH platform.	32
4.1 The laptop-based interface that the nurse uses to manage all participant communication. The system has a list of searchable contacts on the left, the communication history of the currently selected participant in the center, and mechanisms for tracking the participant’s visit history and health status on the right	38
4.2 Total number of messages sent during each week before and after delivery. Week 0 corresponds to the week in which participants were recorded as giving birth. Week -5 depicts five weeks before giving birth while 5 shows five weeks after giving birth	44
4.3 Graph showing the number and length of conversations (two or more related messages) started by the system and by participants.	51
5.1 System design for three-way SMS conversation in enrolled partner dyads. (1) Automated messages from the system go to both partners. (2) Replies from either partner go only to the system. (3) Replies from study staff by default go only to the original sender, but message and response can optionally be forwarded to the other partner. This design allows study staff to have confidential conversations with either partner.	67
5.2 Phone feature use among participants. Percent of participants who used (bottom/green), had but did not use (middle/yellow), and did not have (top/blue) each phone feature. Participant categories: (n) mothers with partner not enrolled, (e) mothers with enrolled partners, and (p) partners.	75

5.3	Cumulative frequency of SMS messages sent by mothers (grouped by partner status) and the enrolled men. Only 10% of men send more than 20 messages while 50% of women do.	77
5.4	Response times between an outgoing message and an incoming message on a logarithmic time scale. Results are broken down by partners, mothers with enrolled partner, and mothers without an enrolled partner.	79
5.5	The topic of incoming messages by group. All mothers and partners engage with the SMS system around the same topics and at the same frequency. . .	81
6.1	Map of Kenya with the six study sites starred. Two sites are on the outskirts of Nairobi and the other four in Western Kenya near Kisumu. Map from Nations Online Project	90
6.2	Phone sharing breakdown. At baseline 70.3% of participants do not share a phone. Of those that share their phone 16.9% say that they primarily have access to the phone all day.	94
6.3	Comparison of baseline data for mPesa, SMS, Internet, and IVR use between participants at Nairobi and Western sites. (* p<0.01 ** p<0.00001)	95
6.4	Comparison of baseline data for mPesa, SMS, Internet, and IVR use between participants with and without a touchscreen. (* p<0.01 ** p<0.00001) . . .	96
6.5	Comparison of baseline use of specific phone features between participants at Nairobi and Western sites. Non-use of a feature does not mean the phone does not have that ability. (* p<0.01 ** p<0.00001)	96
6.6	Comparison of baseline use of specific phone features between participants with and without a touchscreen. Non-use of a feature does not mean the phone does not have that ability. (* p<0.01 ** p<0.00001)	97
6.7	CDF of SMS totals for the two-way group. Only 3% of participants received fewer than 100 system messages. The median number of participant messages is 23 and the top 10% sent more than 93. A total of 26 (9.4%) participants sent zero messages.	100
6.8	CDF engagement calculated as both message success and weeks active for two-way group. A flat blue line at 1.0 would represent 100% of messages delivered. The median of message delivery is 0.83 and the median percent of weeks active is 0.13. For context a weeks active ratio of 0.5 represents being active for over 52 weeks.	100
6.9	Relationship between percent of system messages delivered and percent of weeks active. No participant is active for more weeks than they successfully receive system messages.	101

6.10	Total number of messages per week from the start of the study (Nov 2015) through the end of May 2019. This count includes all outgoing messages (system + nurse).	102
6.11	CDF of SMS delivery delta (the time between sending a message and receiving a delivery report). The x-axis is log scale in hours and the data is only for SMS messages with a delivery report – thus excluding messages with an unknown status.	104
6.12	Percent of participants who responded per week for the four weeks before and for weeks after the Africa’s Talking system outage Tuesday November 22, 2016.	105
6.13	Violin plot of SMS delivery success rates comparing system messages to the two-way and one-way group and nurse messages to the two-way group.	108
6.14	Heat map of all system messages sent to two-way participants in the study for longer than 100 weeks. Each row is all messages sent to an individual participant and each square is one message color coded by status. Participants are sorted on percent of successful messages, weeks active, and then total messages. The large group at the top are all participants who received more than 80% of messages. The small group at the bottom are all participants who received less than 20% of messages.	110
6.15	CDF of longest miss streak (left) and future SMS delivery success by current longest miss streak (right).	112
6.16	Message delivery (red), weeks active (blue), and messages per week (purple) by quarter since enrollment for quarter in study (left) and centered on delivery week (right).	113
6.17	Percentiles for Message delivery (left), messages per week (middle), and weeks active (right). Percentile lines at 25% (red), 50% (blue), 75% (purple) and 90% (black).	114
6.18	Percentile of participant message length by quarter since enrollment.	115
6.19	Distribution of expected ANC period (signup date to EDD) and difference between EDD and actual delivery.	117
6.20	The paper based ANC register for Mathare North clinic.	118
6.21	Distribution of delivery to system notification delta in weeks on the left and the distribution of EDD to system notification on the right.	118
6.22	The delivery to system notification delta grouped by notification type (sms, phone, visit, other) on the left and by study arm on the right.	119
6.23	CDF of weeks between delivery and system notification (notification delta) split based on study arm.	120

7.1	Comparison of GSM applications design space split by communication medium and interaction method.	132
7.2	Left: Example Maternal Health USSD application running on IdeaMart USSD Simulator; Right: USSD home screen for Indian mobile operator BSNL. . .	136

LIST OF TABLES

Table Number	Page
3.1 Overview of all five mWACH projects with start and stop dates, total number of participants and message totals.	16
4.1 Examples of automated messages sent by the system at (a) 27 weeks pregnant; (b) 8 weeks after birth; and (c) 2 days before a scheduled clinic visit. Note that, as is the case for (b), any messages longer than 140 characters were sent in two parts.	41
5.1 Male partner status of enrolled female participants. Total study enrollment was 260 n_a with 130 participants in the intervention arm n_i . Women’s enrollment is further broken down by male partner status: not invited, invited (but not enrolled), both invited and enrolled, and finally enrolled but sharing a phone.	63
5.2 Distribution of 11469 SMSes sent or received in the study. All messages from a dyad sharing a SIM card or phone are counted as coming from the mother.	66
5.3 Demographic characteristics and technology access by gender. For continuous variables the median and interquartile range (IQR) are shown, for categorical the count and percentage are shown. * paired t-test for continuous, McNemar test for binomial, McNemar-Bowker test for multinomial. ** t-test for continuous, χ^2 for categorical	70
5.4 Percentage of participants self-reporting phone feature use at enrollment based on partner enrollment status. For enrolled participants significant p-values for paired t-test comparing enrolled mothers and partners are indicated at **p<0.01, *p<0.05.	74
5.5 SMS engagement by individuals and couple dyads. Active (column 2) is a count of participants who sent at least one message at anytime during the study. In the Total SMS column the median and inter-quartile range are given.	78
6.1 Breakdown of enrollment (n) and total messaging numbers by study site. The total number of messages System, Participant and Nurse messages are only for Two-Way participants at each site. Response rates are simple fractions based on messaging totals.	91

6.2	Basic demographic baseline data for all participants split based on location (Nairobi or Western sites) and having/not having a touchscreen. For each data point the p-value from a t-test is given.	92
6.3	Baseline technology use data for all participants split based on location (Nairobi or Western sites) and having/not having a touchscreen. For each data point the p-value from a t-test is given.	93
6.4	Breakdown of number of participant messages in response to each of the 39993 outgoing system messages sent. 82.0% of system messages have no participant response while 13.87% have a single participant response and 2.59% have two messages from a participant. There is a long tail of longer conversations. . .	99
6.5	Breakdown of two-way system message status and frequency of common delivery errors. The vast majority unknown status messages are <i>AbsentSubscriber</i> errors from before June 2018.	107
7.1	Third party REST gateway providers and SaaS companies.	123

GLOSSARY

ART: Antiretroviral Therapy

ANC: Antenatal Care

EDD: Estimated Delivery Date – usually calculated based on the last menstrual period

HCI: Human Computer Interaction, a sub-field of computer science

ICT(4)D: Information and Communication Technologies for Development

IVR: Interactive Voice Response, automated voice calls

M4D: Mobile for Development, a field within ICT4D focused on mobile phones

MHEALTH: general term for using mobile phones in medical care

MWACH: Mobile Solutions for Woman, Adolescent and Child Health

PREP: Pre-exposure prophylaxis for HIV

PMTCT: Prevention of mother-to-child transmission (of HIV)

RCT: Randomized Controlled Trial

SBMP: Semi-automated bi-directional messaging platform (such as mWACH)

SMS: Short Message Service protocol for sending a text message

USSD: Unstructured Supplementary Service Data, a session based text mobile phone protocol

Chapter 1

INTRODUCTION

A central paradox in the field of Information Communication Technology for Development (ICTD) is that often the users with the greatest needs also have the least access to technology. In the 2016 report *Digital Dividends* [77] the World Bank shows how access to the information economy increases productivity and efficiency in business, health, and governance. However, that same report indicates that while 60% of the world's population technically has Internet only 15% has reliable access to high speed connections. Those same people and communities living at the margins of the digital economy also face the greatest burdens of disease. For example the Global Burden of Disease report shows that Sub-Saharan Africa is disproportionately effected by communicable, maternal, nutritional, and newborn health issues [107]. This paradox of technology access versus need is replicated in areas of education, business, and utilities throughout the world. The central question for ICTD researchers and practitioners is how does a technology based solution provide any results – or even act as a catalyst for improved health, education, or economic livelihoods.

Flashy new advances in technology that dangle tantalizing benefits in font of researches and practitioners are impractical to implement, scale, and maintain. Such *Charismatic Technologies*, as Morgan Ames calles the One Laptop Per Child and other over-hyped projects [9], generate an abundance of media attention but often fail to live up to their lofty promised goals. To a former Peace Corps Volunteer such as myself, these failures are predictable. First, when it comes to global development there are no universal panaceas only incremental local improvements based on individual effort [178]. Second, for an intervention to be truly universal it must reach users on their own devices with little overhead and operator within existing systems. Research on mobile phone use at the bottom of the pyramid

indicates that only 1 in 4 Kenyans use Internet on their phone [39, 101]. For this reason projects in the area of mobiles for development (M4D) will focus on universally available services such as voice, SMS, and USSD [102]. However, designing services to be deployed over basic communication channels presents both technical and usability challenges that must be carefully considered.

This thesis focuses on using these basic communication channels built into every single mobile phone for reaching end users on their personal devices. I term these universal apps to contrast them with Android or iOS apps that require a specific mobile operating system, a data connection, and usually an account in an application store. A universal app in this context works on every phone from a sub \$10 (USD) basic phone up to the newest \$1,000 (USD) from Google. This is important for programs want to reach as many people as possible, providing digital accessibility and equity across socioeconomic backgrounds.

For this reason SMS and IVR services have been a popular tool for reaching end users in global health and the mobile for development (M4D) world [85, 189, 45]. Many of these M4D services designed to reach people at the bottom of the pyramid provide only one-way 'push' messaging. These types of services do not engage with users on their own terms and therefor may not be as effective as intended [100]. Instead, M4D projects should 'ask, not tell' when trying to engage with end users [112]. In this vein we have developed a semi-automated bidirectional SMS platform designed to be used in Kenyan maternal health clinics. The Mobile Solutions for Woman Adolescent and Child Health (mWACH) platform has been used for five different bi-directional SMS studies. mWACH connects participants directly to health professionals via SMS through personalized weekly automated messages and providing clinicians with an SMS action dashboard for responding to incoming questions and concerns.

1.1 Contributions

In this thesis I present my work developing universal apps for connecting end users to domain experts. The bulk of this work has been in designing, deploying, and evaluating mWACH which is a semi-automated bi-directional SMS platform. This has been a collaborative

project with the University of Washington Global Health Department and Kenyatta National Hospital in Nairobi, Kenya. Currently I have helped deploy mWACH in five different projects sending over 120 thousand SMS messages to more than 2000 participants in peri-urban and semi-rural Kenya. These projects have led to multiple publications in both the HCI4D and global health literature [150, 152, 193, 170, 83, 59].

The work described in this thesis contributes to the M4D literature by exploring how semi-automated two-way messaging platforms enable novel behavior change interventions. The focus of this work is on the context of engaging pregnant women with health systems; however, the broader design considerations and specific tools built are applicable to other M4D domains such as agriculture, education, and civic engagement.

The three primary research questions and contributions of this work are:

R1 What factors contribute to the design of semi-automated bidirectional SMS platforms.

To answer this we design, build, and deploy multiple semi-automated bidirectional SMS platforms in Kenya and synthesize findings into a framework for understanding requirements and design of such systems

R2 How do Kenyan mothers engage with a bi-directional SMS system over a long period of time?

To answer this we perform quantitative analysis of meta-data surrounding messages to determine engagement and provide feedback for future projects design and needs.

R3 What are the limitations of SMS as a medium for a messaging platform? How can this system be extended to other mediums such as USSD or WhatsApp?

To answer this we explore how the SABM paradigm fits into the Universal App design space with an emphasis on USSD.

1.2 Overview of Thesis

First I present an overview of existing universal applications in the M4D literature in Chapter 2. This background provides context for the mWACH platform I developed and deployed. Chapter 3 presents an overview of mWACH starting with the design process and evolution from an undergraduate capstone project to a system used in multi-year multi-clinic studies in Kenya. This overview briefly summarizes all five mWACH projects as well as describing the properties of a semi-automated bi-directional messaging platform.

Each of the following three chapters present unique details about one of the mWACH projects. In Chapter 4 I present work published in CHI 2015 on the first deployment of mWACH looking at basic messaging patterns and how the open-ended two-way platform was used [150]. Chapter 6 describes how mWACH was modified for a postpartum family planning project that involved male partners as well as mothers. This work was presented at COMPASS 2018 [152]. Chapter 6 is the last project specific chapter and presents a detailed longitudinal analysis of messaging trends and engagement overtime.

In Chapter 7 I introduce work presented at DEV 2015 [151] that describes the potential of USSD to be a third universal app (along with IVR and SMS). Finally in Chapter 8 I present a final discussion on all work in the thesis and offer concluding remarks.

Chapter 2

BACKGROUND

There is a large body of research that explores the feasibility of direct mobile messaging on behavior change outcomes. These projects span the mobile design space and while smartphone applications are one area of research [43], the vast majority focus on communication channels that are universally available on every GSM phone such as IVR systems [98, 143], SMS initiatives [113, 45, 53, 150], and even USSD based projects [123]. In this section I first focus on the underlying theories of behavior change that have informed many of the projects in mHealth. Next, we explore the range of technologies developed for mobile messaging projects and examine the evolution of SMS-based platforms. Third, we review the extensive literature from the medical and HCI communities on individual projects for mobile messaging to better understand what has and has not proved effective. Finally, we look at how messaging projects can remain sustainable and integrate with existing health systems.

2.0.1 Behavior Change Theory

Whether a mobile messaging project is targeting the health of end users [113, 98, 143] or workflow of community health workers (CHWs) [53, 215] the end goal is almost always behavior change. Even in interventions aiming to strengthen health systems through reporting and data flows [14, 181, 17], the adoption of new reporting practices and procedures represents a behavior change. It is important to base the development and design of M4D projects on theories of behavior change. There is a rich history of literature into what factors increase the likelihood of an individual or group adopting new practices, technologies or behaviors. The three theories that have heavily influenced the design and implementation of mobile messaging projects are described in the following subsections. Each of these theories are

highly intertwined and give different insights into the factors that help determine whether or not an individual adopts a certain behavior.

The Theory of Planned Behavior

Based on the *Theory of Reasoned Action* (TRA) [6], the *Theory of Planned Behavior* (TPB) [6, 7] posits that an individual's probability of performing an action is a result of a '*behavior intention*'. This intention is the result of peer's attitudes and beliefs (subjective norms), the individual's attitude towards the behavior, and an individual's perceived behavioral control. As an example of how the TPB works, suppose an NGO is working with a local clinic and wants to increase the rate of immunization in an area through mobile messaging. Messages should be crafted that not only try to convince families that immunization is good (attitudes) but also demonstrate that the community supports this practice (subjective norms) and give easy direct actions that will allow them to immunize their child (perceived control).

The TPB and a sub-theory derived from it known as the *Technology Acceptance Model* (TAM) [46, 200], which posits that an individual's attitude towards a new technology is influenced by the technologies perceived ease of use and the usefulness, have informed many mobile based projects. In analyzing a well known SMS intervention, Smillie et al. [180] argue that by replacing '*technology adoption*' with '*engagement in care*' TAM just as easily applies to health behavior. Kitagawa [103] explores the relative importance of social cues versus personalization in SMS messaging and finds support for social signals reducing risk behavior. Recent work in the HCI literature has shown that when an event is far in the future, people give more weight to attitudes when considering the likelihood they will do a certain task but closer to the event more weight is given to perceived control [185].

Health Belief Model

Another closely related theory of behavior change is the *Health Belief Model* (HBM) [16, 75]. This model was developed by the public health field to help describe attitudes, beliefs, and knowledge that lead to behavior adoption. This theory is more complicated than the TPB

and posits that six factors effect the likelihood of behavior change. These are: perceived susceptibility, perceived severity, perceived benefits, perceived barriers, cues to action, and self-efficacy. The HBM predicts that health messaging can be effective in two ways. First, by carefully crafting a message and responding to concerns with real and valid information, a health messaging campaign changes perceptions around risk, benefits, and barriers. Second, timely messages act as simple cues to action and help remind patients what (and possible when) they should be doing something. Visit reminders are a great example of effective and timely cues to action [82].

FOGG Model

The previous behavior change theories emerged from the health and psychology literature. In the field of HCI and persuasive computing the primary operating theory of behavior change is the Fogg Behavior Model (FBM) [64]. This model was developed based on work exploring how technology can be designed to reward and encourage specific user behavior [63]. The FBM separates the likelihood of an action being performed into three components, which all simultaneously exist for an action to be performed. First, there must be sufficient motivation; second, an agent must have the ability to act; and third, there must be a trigger or stimulus to prompt the action. The FBM can explain many programs built to help change behavior. For example, someone who wants to stay focused may install a browser extension that displays their todo list on a new tab. Now every time they open a new tab a triggering event occurs to remind them to stay on task, however, if there is not sufficient motivation to stay on task then the user can simply ignore their todo list. The field of persuasive technology lies at the intersection of HCI and Ubiquitous Computing and explores how technology can be used to intentionally drive behavior change. The asynchronous nature of SMS means it is well suited to be the trigger in the FBM, particularly when it comes to health interventions [65].

2.0.2 Universal Mobile Messaging Platforms

ICTD projects that must reach people over personal devices with limited digital exclusion cannot assume the presence of a smartphone, a data connection and especially not extra sensors such as a camera or GPS. This requires the design of *Universal Apps* that make use of services uniformly available on every mobile handset. Limiting the design of messaging platforms to the most basic services is fundamental for digital inclusion and equity. Even when an ICTD initiative assumes the presence of a smartphone it is important to consider the cost of mobile data and practices within pre-paid mobile markets [174, 71]. Universal mobile applications are not the most cutting edge or novel and therefore do not receive the same level of attention as their more flashy and widespread counterparts [9, 127]; however, they represent some of the most successful large scale projects such as USSD codes for topping up prepaid mobile phones or M-Pesa [92]. It could also be argued that simply using a mobile phone to connect people is the universal app with the furthest reach people [94, 20].

Mobile For Development Design Space

Mobile applications limited to just text and voice, designed to work with carrier provided services have generally focused on IVR and SMS. A third option for a universal app is Unstructured Supplementary Service Data (USSD). USSD is a protocol defined inside the Global System for Mobile Communications (GSM) standard.¹ USSD is a session based protocol, like IVR, that supports the exchange of text data, like SMS, thus filling a gap in the M4D design space. The USSD interface is less familiar to most users but is quite common for interacting with carriers services. This is done by sending a *star code* (such as *144#) to check an airtime balance with Safaricom in Kenya Figure (a) and accessing interactive menu systems.

There are four primary considerations for choosing between text and voice channels.

¹USSD is not supported on CDMA networks, however GSM covers 90% of the world and is particularly ubiquitous in those regions with the lowest smartphone penetration.

The first, and primary reason, is that for most projects creating and collecting text based content will have lower administrative cost. For example, four hundred preconfigured text messages can be modified with the name of sender and recipient when sending an SMS [150], prerecording these as voice messages means content is not as flexible or customizable. This is a major reason to favor text over voice solutions. However, the second consideration is the literacy proficiency of the end users — which may require that voice be used. A data collection program used by trained CHWs can use SMS [11], while, a program targeting pregnant women from a Mumbai slum should use voice [143]. Another reason to potentially prefer voice over text is that non-Latin script support on low end phones is often poor [121]. Lastly, it is important to consider the bandwidth and cost requirements of each channel. While SMS and USSD have extreme character restrictions (160 and 180 chars respectively) they can easily scale to many thousands of messages an hour. Voice, on the other hand, scales well with larger content but establishing multiple simultaneous voice lines raises costs. There are however extra cost and overhead to find third party providers for SMS and USSD. The cost of USSD setup and dedicated shortcodes can be prohibitive for smaller NGOs working with less than 1000 people, while for a large project sending millions of text messages a week the setup fee is trivial [89]. Such concerns are very dependent on each country, for example voice is cheaper relative to SMS in India while it is the reverse in Kenya.

Evolution of SMS Based Systems

The technology supporting projects using SMS to communicate with end users has changed rapidly and repeatedly in the past decade. A common trend in all this change has been to make it easier and cheaper for small scale projects to test, develop and deploy SMS programs. Initially to automate SMS you needed a business agreement with mobile operators — which was layered in bureaucratic red tape and only feasible for large players such as banks and retailers. One of the first widely used applications for automating SMS communication was FrontlineSMS [12]. FrontlineSMS is a Java based application intended to be run on a desktop or laptop at an NGO office. It sends and receives text messages through a USB connected

mobile phone. This model is similar to open-source IVR systems intended for local deployment and use [198]. FrontlineSMS allowed users to create keyword based SMS applications. These work by identifying simple keywords or regular expressions at the beginning of incoming messages and then performing an action such as entering information into a database and replying with a confirmation. With this model users can easily create SMS based polls, data collection tools, and opt-in based SMS mailing lists.

FrontlineSMS proved to be immensely successful due to the low-cost setup and adaptability [13]. However, once a project grows to the point where a desktop sitting in the corner no longer provides enough reliability, a different technology is required. One solution is to move the processing power out of the NGO's office and onto someone else's servers through a cloud based service. This is what the open source Django based application RapidSMS [130, 158] was designed to do. RapaidSMS provides a programming interface to make they same sort of keyword based applications as FrontlineSMS.

In a cloud environment it is no longer possible to attach a phone to the server as a gateway. Instead HTTP transports for connecting with REST endpoints that send messages text messages are needed. In recent years many third-party services such as Twilio [192] and Nexmo [129] have been created to provide exactly this, enabling outgoing SMS throughout the world for pennies. If you need cheap two-way text messaging then there are multiple country specific providers such as Africa's Talking [187] that provided dedicated value added services for a monthly fee. Another more low cost method was created by Ushahidi called SMSsync which turns any Android phone into an SMS HTTP gateway [194] which allows RapidSMS to be used even in places where third-party providers do not exist [153].

The biggest issue for adoption of RapidSMS is that in order to deploy even a basic SMS poll an experienced web developer is needed. To meet the large demand for messaging applications, multiple companies have filled this gap by essentially providing RapidSMS as a packaged service [124]. For example the team behind FrontlineSMS has created FrontlineCloud which offers even more features than the original software in easily accessible web-based package [69]. Similarly, UNICEF has released an SMS application called RapidPro with a backend designed

to scale to millions of text messages a day and a web-based front end that allows users to graphically create complex SMS based interactions [166].

2.0.3 Engaging Users With SMS and IVR

As mobile phones have become integral parts of everyday life, researchers, public health officials, and practitioners have all sought to reach individuals over personal devices. The key idea is that by automating mobile messaging, the effects of an intervention can easily scale to large segments of a population. The primary means of reaching health workers and end users has been over SMS and IVR. In this section we explore the uses of mobile messaging to engage with individuals and health systems using both one-way and two-way platforms.

One-Way SMS Communication

A large number of prior studies have focused on one-way communication in which a server sends bulk SMS messages to a large number of recipients. Reminder systems fall under this category, including solutions designed to increase medication adherence [117], reduce missed clinic visits [106], or increase the timeliness of health worker home visits [53]. However, a major disadvantage of these systems is that the recipients have no way to send messages back to the system. For example, in one study health workers still received reminders to visit patients that had moved away from the area [53].

Another class of one-way SMS systems allow users to submit data to a server. In most cases, submitted messages must conform to a specific structure to enable automated parsing by the system. This model has been used to report stock levels [14], communicate diagnostic test results [11] or submit data [149]. However, several of these studies find that users struggle to compose properly structured messages and frequently make mistakes [149, 153]. To overcome this limitation, other systems allow users to submit unstructured messages and then employ complex server-side processing to try and make sense of the messages. For example, Ushahidi [141] used server-side natural language processing to try and decipher submitted messages, while Hellström et al. [87] used crowdsourcing to enable users to report

election incidents by SMS. However, these systems do not provide feedback to users and participants reported not using the system because they had no idea if the information they submitted made any difference [87].

Two-Way SMS Communication

To overcome the lack of feedback provided by one-way systems, several projects have designed two-way user-server SMS platforms in which users submit structured queries to a server and receive automated responses. This model has been deployed to provide users with health [130], education [45], and agricultural [199] information, web search [27], and product pricing [134]. However, as with one-way systems that use structured messages, many of these projects report high error rates in the messages submitted to the system. Moreover, training users to compose formatted messages is challenging and hampers the scalability of these systems.

Perhaps the closest match to the semi-automated bi-directional SMS platform created in this thesis work are ‘human-in-the-loop’ systems that require the input of both humans and machines. Prahalad [157] and Joshi [97] have proposed such systems in the context of rural banking and other global development efforts, though not in relation to user engagement or SMS communication. There have also been systems designed for group communication and information sharing such as Tangaza [137] for SMS and Swara [197] for IVR.

Health Outcomes and Behaviors

The large corpus of work around mobile messaging for health behavior change spans the medical, HCI and ICT4D communities. In this section we explore how each of these fields has contributed to our understanding of the methods and effectiveness of mobile messaging projects. This work is not limited to a developing world setting and the ubiquity of SMS has led multiple research teams to design using a simple text channel. The gold standard for proving effectiveness in the medical community is randomized controlled trials (RCTs). Studies comparing SMS or IVR interventions with standard of care are common. Health outcomes targeted by RCTs outside of the developing world include smoking cessation [169, 196], diabetes

care [66, 82], mental health treatment [70], weight loss [177] and asthma management [213, 212]. Even when just using one-way information based messaging many of these studies have found significant and surprising results. For example Chow et al. [34] showed that sending four text messages a week with motivational reminders and behavior change tips for six months significantly lowered cholesterol, blood pressure, and reduced smoking among patients at risk for coronary heart disease. With respect to maternal health, one of the largest interventions has been Text4Baby [146], which started in 2010 and had enrolled over 1.1 million participants by 2018 [21]. Women who were currently pregnant, or had given birth within the last year, could sign-up to receive one-way health messaging via SMS. Text4Baby was not a controlled study but a public health intervention and was deployed in both the United States and Russia.

Using SMS as a health and behavior change intervention to increase antiretroviral adherence has been another large focus of the global health community showing promising results. A significant work in this area is the WelTel Kenya1 study [113]. The WelTel model focuses on HIV positive individuals initiating antiretroviral therapy (ART) and consists of simple two way messaging plus follow up phone calls. Since the WelTel study was so influential in proving how successful a simple SMS intervention could be, it is worth explaining more about it here. Once each week participants in the intervention arm received an SMS containing simply the word *Mambo* (Hello). They were instructed to reply with either *Sawa* (Fine) or *Shida* (Bad) based on if they had a problem or not. Participants who said they had a problem or did not respond within 48 hours received a follow up call from study staff. Importantly, the WelTel Kenya1 study measured not only self-reported adherence rates, but also HIV viral load levels. Analysis showed a significant difference in rates of HIV suppression between the control and intervention groups with 57% of the intervention group successfully suppressing the virus while only 48% of the control group did.

As an early two-way messaging study, the WelTel project has many interesting implications for the design of future interventions. While WelTel was an mHealth solution the technology behind it was exceedingly simple. The paper claims that “health workers used multiple

recipient (bulk) messaging functions to improve efficiency”. This indicates that the *Mambo* messages were being manually sent each week with virtually no automation. The proportion of participants responding with *Shida* was low — averaging just 2.5%. However, a large number of participants were still called since 24% of outgoing messages received no response. Participants reported that cost and other logistical factors such as battery charge or forgetting were primary reasons for not responding rather than health issues.

The promising results of this early SMS based mHealth study have prompted several follow up works. Not all of these studies have shown a significant result such as one aimed at reducing sexual activity after circumcision [135] and even studies that copied the WelTel model in different countries [116] — though an important conclusion from this trial is that allowing for increased open-end engagement through a mobile phone has a positive effect on outcomes [116]. Other implementations of SMS based messaging for global health have shown promise. Zurovac et al. demonstrated that sending information based SMS messages twice a day to health workers significantly improved malaria case management [215]. The wired mothers project used low frequency text messaging to promote clinic delivery among pregnant women [171]. An analysis of Nokia’s mDiabetes program in India — which reached over one million subscribers — showed that fewer intervention participants had a decline in health issues and there was an improvement in diet compared to a control sample that did not opt-in to to intervention [154].

Chapter 3

THE MWACH PROJECT SUITE

In this chapter I describe the design, deployment and evolution of mWACH (Mobile Solutions for Woman, Adolescent and Child Health). mWACH is a collaboration at University of Washington between the Global Center for Integrated Health of Women, Adolescents, and Children (Global WACH) in the Department of Global Health and the ICTD Lab in Paul G. Allen School of Computer Science & Engineering. The complex history of mWACH spans this entire thesis work and a description of its origin and evolution requires its own place. Currently four projects using the mWACH platform have been completed with one still ongoing and several more being planned. All projects were implemented at Kenyan health clinics through the UW Kenya collaboration. This chapter provides a concise summary of the timeline and history for all mWACH projects and provides context for future chapters which focus on the details of three individual projects.

The initial goal of mWACH was to explore a more patient focused care model in mHealth and ICTD. Previous SMS based projects such as DeRenzi et al. [53] and [215] focused either on community health works or had limited engagement with patients and end users. The WelTel [113] model of simple bi-directional SMS with single word messaging threads had proven effective. However, none of these projects provided truly open-ended bi-directional channels between health providers and patients — which from a clinical perspective would provide more personalized care.

It was with this goal in mind that work began on a semi-automated bi-directional SMS platform in the Winter and Spring quarters of 2012. The semi-automated aspect of the platform comes from a human-computer hybrid approach to messaging. A large message bank of pre-written SMS templates is created to be sent at specific time relative to

	Start	Stop	Focus	n	Two-Way	System (2-way)	Nurse (2-way)	Participant (2-way)	Unkown	Duration			
mWACH V1	Aug 2013	Jan 2015	Maternal Health	299	100	7286	3904	484	464	1063	986	0	20 anc / 10 post
mWACH X	Dec 2015	May 2019	PTMTC	824	276	79788	39279	6376	6361	10812	10122	18973	26 anc / 104 post
mWACH X	July 2016	Sep 2017	Family Planning	260	130	7082	6852	3356	3577	3718	3356	6046	14 anc / 25 post
mWACH NEO	Jan 2018	Feb 2019	Neonatal Health	799	–	26021	–	13535	–	19952	–	8010	10 anc / 12 post
PriYa	Feb 2018	Jan 2019	PreP	231	–	5562	–	1063	–	1178	–	1378	31 Weeks
Totals				2115	1436	125739	81618	24814	25000	35660	35594	34407	

Table 3.1: Overview of all five mWACH projects with start and stop dates, total number of participants and message totals.

medically important dates and events. Participants receiving the bi-directional intervention are encouraged to respond to these messages and study staff are able to reply to the questions and concerns of incoming SMS messages.

The first participants were enrolled in version one of mWACH in August of 2013 at Mathare North Health Centre. on the outskirts of Nairobi. As of May 2019 and across all mWACH projects a total of 126,041 automated messages have been sent with 35,594 responses from 2,115 individual participants and 24,814 messages from study staff. The details for each study can be found in Table 3.1. The rest of this chapter details how mWACH went from a single conversation between members of my thesis committee to the platform that is currently in use.

3.1 Overview of mWACH Projects

3.1.1 Encouragement: mWACH V1

The genesis of mWACH is a story of cross-disciplinary collaboration enabled by the unique environment at large research universities. In early 2012, Dr. Jennifer Unger gave a talk on using SMS messaging to connect pregnant women to health professionals in Kenya. Brian DeRenzi, a postdoc in Computer Science and Engineering, attended this talk and since two-way SMS to community health care workers was part of his thesis work he realized that creating a web-based two-way messaging platform would fit the core competencies and research goals of the ICTD Lab. The idea was pitched as an undergraduate capstone project and a student team began working on it in the winter quarter of 2012. This project

eventually became the *Encouragement Platform*, which by the end of Spring quarter was the demonstration version of what would become the two-way mWACH messaging platform.

My first commit on the project was on January 2, 2013 which established a connection to a Kenyan based SMS gateway. This was in preparation for the start of a RCT of the platform to improve pregnancy outcomes at Mathare North Health Centre in Nairobi [193]. The study randomized 300 participants 1:1:1 into a control group who would receive only the standard of care without SMS messages, a one-way group who received weekly SMS messages but could not respond, and a two-way group who received similar messages to the one-way group except that each message ended in a direct question to encourage engagement with the study via SMS.

Besides adding hooks in the system for bi-directional SMS, several other large changes were needed to turn the capstone project into a viable platform for the first mWACH study. The first was to add the one-way and control group to the system since originally only the two-way use case was considered. Second, the registration page needed updating to include relevant medical information and the next clinic visit. Lastly, a message bank of pre-composed SMS templates needed to be added. This message bank was created by Dr. Jennifer Unger based on focus groups with pregnant women and providers in Mathare. There were 159 message templates in this bank going from 20 weeks before delivery until 6 months postpartum.

The first participant was enrolled in August of 2013 and the last SMS message sent in January of 2015. The system sent 3,904 messages to the two-way group and we received 986 messages from participants. The single study nurse sent 464 messages back in response. The mWACH V1 row in Table 3.1 shows further messaging details.

A full analysis of the results of the mWACH V1 RCT is beyond the scope of this thesis but have been published the International Journal of Obstetrics and Gynaecology [193]. This study found that exclusive breastfeeding through 24 weeks is significantly higher in the two-way (62%) vs control (41%). The percent of participants exclusively breastfeeding in the one-way group was significant at 10 and 16 weeks but not 24 weeks. The probability of contraceptive use within 16 weeks of delivery is significantly different between the two-way

(73%) vs control (57%) but this difference was not significant by 24 weeks. The work done for mWACH V1 motivated the future projects.

3.1.2 mWACH X: ART Adherence and PTMTC

The goal of mWACH X was to adapt the V1 platform for a RCT aimed at assisting pregnant women living with HIV. The key focus would be on key behaviors in the prevention of transmission from mother to child (PTMTC) [59]. This involved a complete re-write of the system using *AngularJS* and *Django* to better support the study workflow. The reasons for these changes and the evolution of the platform are described in subsection 3.2. mWACH X started in December 2015 with enrollment initially at three clinics – Mathare North in Nairobi and Ahero and Bondo near Kisumu. To increase the pace of enrollment, in October 2016 enrollment began at three more clinics – Rachuonyo and Siaya near Kisumu and Riruta in Nairobi (See Figure 6.1 for a map of Kenya all mWACH locations). Enrollment for mWACH X ended in May of 2017 with 825 participants in the system – although it was later discovered that one of these participants was a duplicate. A longitudinal analysis of mWACH X is the focus of Chapter 6.

The last recorded delivery was in November of 2017 so the current expected end of messaging is December 2019 with an expected period for collecting end of study surveys. Currently a total of 80,090 SMS messages have been sent by the system with 39,993 to the two-way group and two-way participants have sent back 10,122 messages. The shortcode for running mWACH X was shared with mWACH Neo and has received 27,930 messages from phone numbers not registered with the study. A random sample of the messages from unknown numbers indicates that the vast majority are spam such as ads or promotional material for other SMS campaigns, however, it is likely that some of the messages are from participants using unregistered phone numbers.

Messaging in mWACH X could start as early as 26 weeks before the EDD and went 104 weeks after delivery. For these 130 weeks of messaging there were 1,1149 unique templates in the message bank, which were additionally translated into three languages (English, Swahili,

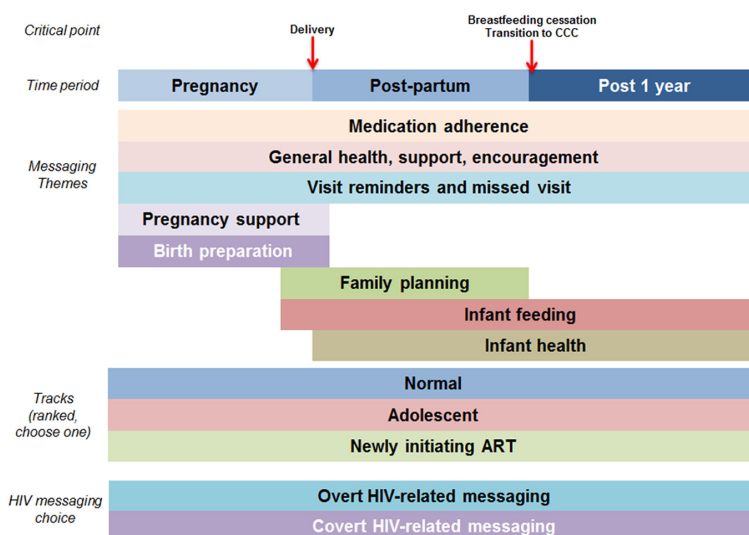


Figure 3.1: Depiction of messaging tracks and content during pregnancy, postpartum, and post one year after delivery.

and Luo) to support all participants. Half of these messages would be for the one-way group and the other half for the two-way group. Besides groups study group participants were also put into messaging tracks based on their status and preferences. Participants who did not share a phone could opt into receiving HIV related content. For participants on the non-HIV messaging track a phrase like *"take care of your health"* would be used instead of *"take your medication"*. The process of designing the HIV messaging track is described in detail elsewhere [170]. Other tracks included one for adolescents, participants having twins, and those just starting ART. These tracks made very minor customization to better tailor content to individuals. A complete timeline of topics and tracks for mWACH X messaging can be found in Figure 3.1.

3.1.3 mWACH XY: Postpartum Family Planning

The goal of mWACH XY was to adapt the mWACH platform into a family planning focused mHealth RCT. Outcomes from this study are beyond the scope of this thesis but have been published elsewhere [83]. This study was conducted from July 2016 through September

2017 at Ahero and Bondo near Kisumu. There was a control and a two-way group with 260 women split 1:1 in each group. A major change in mWACH XY was that women were given the option to also enroll their partner in the study. This three-way conversation over feature phones is described in more detail in Chapter 5. Messaging for mWACH XY began 14 weeks before delivery and ended 25 weeks postpartum. There were no special ANC tracks however all messages had slight variations for enrolled couples vs just the mother. Additionally, if a participant adopted one of six supported a family planning method they were switched into a method specific track. There were a total of 144 templates in the message bank for mWACH XY.

3.1.4 mWACH NEO: Maternal and Neonatal Health

mWACH NEO is a scaled up version of mWACH V1 focusing only on the weeks before and after delivery. As scaled test of the intervention there was no control or one-way group with 799 mothers enrolled from two clinics, Mathare North in Nairobi and Rachuonyo near Kisumu. The project began enrollment in January 2018 and ended in February 2019. There were 62 messages in the message bank going from 10 weeks before EDD until 12 weeks post delivery. In mWACH NEO there was only one messaging track, however, the message bank had nearly double the number of templates as weeks of messaging because variable messaging frequency was used. During ANC one message was sent per week, for the two weeks after delivery one message was sent every single day, for the next two weeks a message was sent every other day, and then a message was sent every three days until 12 weeks after delivery. The goal of this variable messaging frequency was to reach mothers during the most critical times of the neonatal period.

For the time in study each mWACH NEO participant is receiving a much large number of system messages. The system sent a total of 26,021 messages and the 799 participants sent in a total of 19,952. This represents the highest ratio of participant messages to system messages across all five projects. The study nurse in NEO sent 13,535 messages in response to the incoming messages which is a much higher ration then in mWACH X and mWACH V1

but lower than in PrIYA.

3.1.5 PrIYA: PreP Retention

The PrIYA project focused on young women at risk of HIV exposure who initiated HIV Pre-exposure prophylaxis (PrEP) [155]. Women initiating PrEP at two hospitals in Western Kenya were eligible to receive PrEP focused messages in an effort to increase their PrEP adherence. Since the project followed many of the core philosophies of previous mWACH projects – mainly having semi-automated bidirectional SMS messaging linking patients with health experts – modifying the code-base for this project was easy.

Messaging for PrIYA was based off of the PrEP initiation data and continued for 31 weeks. Since there was only one track there were only 34 messages in the message bank. A total of 231 women were registered and messaging lasted from February 2018 through January 2019. All participants received two-way messaging with the system sending 5,562 messages, participants sending 1,178 and the study nurse sending 1,063. A detailed analysis of the PrIYA messaging is outside the scope of this thesis, but it is interesting to note that PrIYA has the highest ratio of participant to nurse messages.

3.2 Evolution of mWACH Projects

This section details the changes made to the mWACH platform over the course of all five studies. Besides improvements to make it easier for study staff to reply to messages and track participants, the system evolved into a more general platform making it easier to adapt to the needs of new studies. For now the target use case is still small interventions or studies scaling to less than 2000 participants. Larger NGO or government based semi-automated bi-directional messaging deployments will use platforms like RapidPro [166] or CommCare [26]. The niche within M4D that the mWACH system fills is a flexible research platform for testing ideas at large enough scale to measure impact and effectiveness.

The original capstone project has everything in a single view illustrated in Figure 4.1). The three main components are:

1. A searchable participant list on the left, ordered by first and last name, most recent message, or upcoming visit dates.
2. A message pane in the center with a text area for sending messages above the conversation history for the selected participant.
3. A list of medical details and visit history to provide context to study staff reviewing messages.

While this interface worked well in testing – where less than ten mock participants were used – it did not scale well to having 100 two-way participants. One major issue with this design is that the message status toggled from *read* to *unread* when viewing and all visual hints that a message had not been responded to disappeared. It was not uncommon for the study nurse to read a message but need to get additional information before responding. The cognitive load of remembering to respond was shifted from the system to study staff and it was easy to forget to follow up via SMS.

The central paradigm guiding the design of the system re-write for mWACH X was a task dashboard as shown in Figure 3.2. This made it very easy to tell if there were currently pending messages or translations. Reading a message would not toggle its pending status – instead a message had to be explicitly replied to or ignored. This can be seen in Figure 3.3 where the first message is marked as pending and has a reply or dismiss button at the bottom. In this way a reply would not accidentally be forgotten if responding required extra information or a staff discussion. This has the added benefit of allowing the system to record exactly which incoming message a nurse is replying too. However, there is not a perfect one-to-one correspondence of replies since sometimes a nurse may reply to two incoming messages at once, or alternatively use more than one message to reply to a single message.

The task dashboard also listed pending message translations and past clinic visits to confirm. Both of these were added based on experiences from the mWACH V1 study. Message translation was added for two reasons. First, it was often easier for study staff to read cleaned

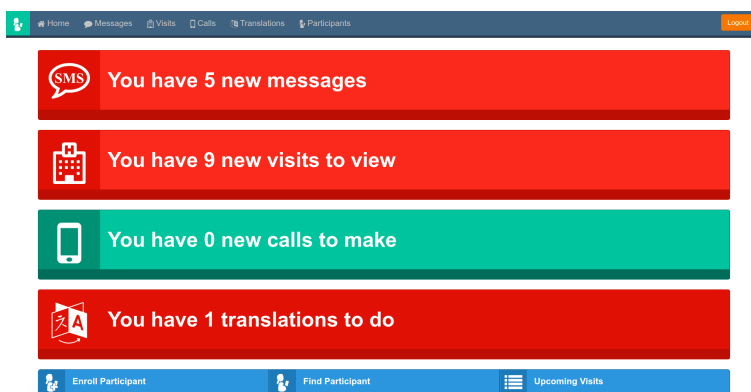


Figure 3.2: Screen shot of main task oriented dashboard in the updated mWACH platform.

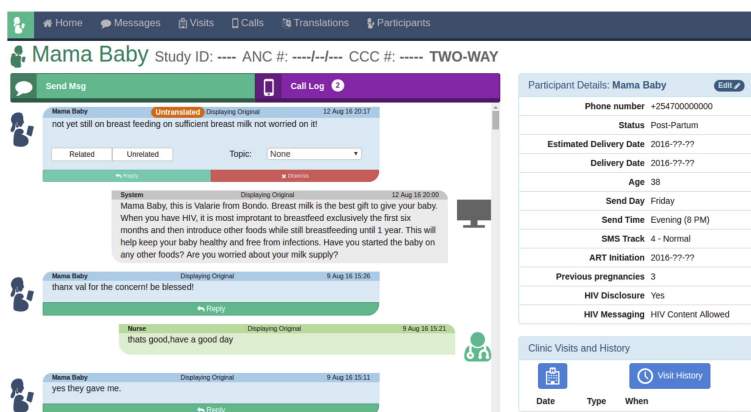


Figure 3.3: Screen shot of a participants detail page in the updated mWACH platform with the first incoming message unanswered.

English versions of messages whose original versions contained texting abbreviations and slang. Second, with the addition of the Western sites in mWACH X, Dholuo would be added as an additional language option and not all staff were fluent in it. Translating messages as the study went along also distributed this workload and allowed US based team members to review messages.

Another lesson from mWACH V1 was that keeping track of clinic visits when running a system that was tangential to daily clinic operations is difficult. Participants were supposed to receive SMS visit reminders two days before an upcoming visit. In order to do this the next visit date for all participants was tracked. Clinic dates were recorded in clinic registers along with other details of the visit. Study staff would inspect these clinic registries looking for records of visits from participants. Not surprisingly, this method of tracking visits was not timely or accurate. In mWACH V1 the system would automatically send out a message two days after a visit if it had not been marked as *attended* in the system. Many people received missed visit reminders who had actually attended their visit but the system had not updated accordingly. In mWACH X the pending visit dashboard was created to not only show upcoming visits but to list visits that needed verification. Missed visit reminders were only sent if the clinic register had been checked and no record of the visit found.

Several other important changes were made to the system during the mWACH X rewrite based on lessons from V1. One major oversight in the original design was how to handle infant deaths. Not wanting to send messages about caring for a child after the child died, we turned messaging off for these participants. However, stopping messages in this situation seemed callous as well. For this reason we developed a loss track which had four weeks of bereavement messages and then continued messaging but focused on the mother's health only. When an infant death was reported participants could now opt in to continuing to receive messages.

Another change to make the messaging system more accurate was the addition of a call two weeks after the expected delivery date. The pregnancy focused messaging stopped at this time and switching someone over to the postpartum messages as soon as possible was

important. This phone call was designed to catch participants who did not report to the system when they delivered and were still receiving pregnancy messages. If they were still pregnant at this call a follow up call would be scheduled two weeks later.

In V1 we had observed participants using multiple phone numbers. This was for a variety of reasons, either they lost their SIM card and got a new one, had multiple SIM cards to start with, or were responding to system messages using a partner or family member's phone. In mWACH X, we made it so that a participant could have multiple phone numbers. Incoming messages from any of these numbers would be associated with that participant. However, outgoing messages would only be sent to one number. Lastly, in mWACH X, the system was designed to support multiple clinic sites with different nursing staff responding to messages at each clinic.

For the mWACH XY project, two major changes were added to the system. mWACH XY started a year after mWACH X, and one thing we noticed in the data was that if participants sent messages on Friday afternoon, they would not get a response until Monday morning when study staff came into work. This can be seen in Figure 3.4 which shows the message totals for participants and nurses throughout the week in 30 minute intervals. The red line for participants has a spike at 8am, 1pm, and 8pm which is when system messages went out. This even occurs on the weekend when the blue line for total nurse messages is flat. While this was expected behavior and the mWACH platform was not designed to be an emergency system with fast response times, the long delay seemed unnecessary. For mWACH XY, participants could not select Friday or Saturday to receive messages. The other major change in mWACH XY was to add partners into the conversation. This is explained in more detail in Chapter 5.

In NEO the only significant change was to add a variable number of messages per week. The focus of NEO was on connecting mothers to the health system during the critical days and weeks after delivery. To help facilitate this, a message was sent everyday for the two weeks after delivery, every two days for the next two weeks, and then every three days until the end of the study at twelve weeks past delivery. Having a variable number of messages per week enabled the mWACH platform to connect with participants when most needed. Other

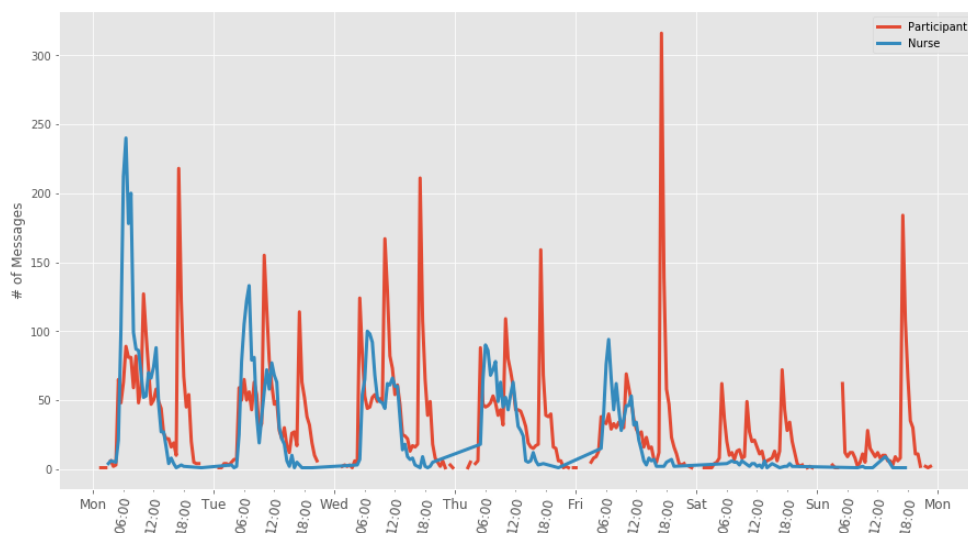


Figure 3.4: Message totals throughout the week in 30 min increments for participant and nurse messages in mWACHX.

critical time periods when this could be used is right after starting ART or if the system knows when a mother is transition from ANC HIV care to the general clinic extra messages could be added during high loss to follow up event.

Over the course of five studies the mWACH platform has evolved to meet the needs of a diverse set of requirements. Making sure the system can run with minimal hardware requirements (e.g. a basic install of RapidPro requires at least five servers) has been important for adapting to the needs of each project. However, each change required an in-depth knowledge of both the code-base and the deployment context – such knowledge is hard to find in one person. We are currently working with software consultants to standardize common modules from all five past projects and make future modifications more streamlined. This process will hopefully make it easier for future global health projects to use mWACH as a base for SMS studies.

3.3 Properties of mWACH Projects

In this section I define the properties of semi-automated bi-directional messaging platforms (SBMP) like mWACH. These properties are general and include more than just SMS systems. IVR systems like TAMA [98], USSD systems, or full web-based EMRs are also included. The three key properties that make such a system are:

- sending patients scheduled automated messages with health and behavior change content;
- allowing patients to ask free form questions;
- providing feedback and answers from domain experts in real time.

Creating a system with all three things requires cross-disciplinary collaboration between medical experts, technologist, and clinic administrators.

In the following four sub-sections I discuss how features of mWACH which enable all three properties of a SBMP. The first two features – scheduled message templates and tracking the state of the world – enable semi-automation. The second two features – the practitioner dashboard and transport layer – enable bi-directional messaging.

3.3.1 Scheduled Message Templates

The first step to semi-automation in SBMP is the design of an SMS template message bank. Designing the messaging schedule and content requires a deep understanding of the context the messages are used in. For all of our projects countless hours went into crafting the most appropriate messages. This is a careful balancing act between being too general versus too specific. For example in MXY a family planning track was created for tubal ligation and only one participant selected this family planning method and received these messages. While an extreme example of specific targeting this shows the amount of preparation that went into creating the message templates.

The limiting factor for the size of the message bank is not the complexity of the software system, rather, it is the time and expertise required to create the messages. The size of the message bank is directly proportional to both the the duration of the study and the number of different messaging tracks and groups. In mWACH X there were 1149 messages templates (3447 if each language is counted separately) for 130 weeks. The extra messages come from small changes for all the track variations.

For example, 10 weeks after delivery, the standard message someone who hasn't opted in to HIV messaging is:

{name}, this is {nurse} from {clinic}. Your health is very important to care for your baby. Continue to take time each day for your health. Are you having any challenges this week?

If they had opted into HIV messaging they would receive the following messages (changes in italics):

{name}, this is {nurse} from {clinic}. Your health is very important to care for your baby. *Continue to take your medications each day, even when the baby stops taking medications. They will keep you healthy and protect your baby. Are you having any challenges with your medications?*

Participants whose child had died received variations of these messages with references to their baby removed and emphasis placed on their own health. Additionally, another version of both messages was created for participants who had twins with *baby* changed to *babies*. Only one participant had twins and she opted into HIV messaging so the non-HIV version of all twins messages was never used (though in a larger sample size it might be). For messages sent a year after delivery there was also a version for mothers who were pregnant again after their initial pregnancy for the study. This track was added during the study when it was realized we needed it. The biggest messaging difference in the second pregnancy track is that

family planning related messages were changed. The above messages are the two-way version but for all variants there was also a one-way version that did not have the last question. For all projects messages were also translated into multiple languages by Kenyan study staff. Figure 3.1 shows the full timeline of message topics and tracks for mWACH X.

As just this one message for 10 weeks after delivery illustrates, there is a large workload for carefully crafting and translating all messages. There is a trade-off between targeting the message templates to an individuals situation and maximizing how often each message is used. Some of the extra tracks are clearly needed – like the infant loss track, which allows mothers to continue receiving supportive messages in a time of need. However, the need for a twins track, while grammatically correct, has diminishing returns. The important takeaway from these two tracks, as well as the second pregnancy track, is that at large enough scale, uncommon situations must be considered in the original designs.

The message bank also had multiple versions of visit reminders, to be sent before and after scheduled clinic visits. During the course of the study, we also sent once off messages to select groups of participants. These messages were sent for things like clinic specific strikes, national holidays, and study announcements. This ability was added to the system after the study started and could only be done manually from a command line. Future deployments of mWACH should include once-off messaging in the admin web interface.

3.3.2 *Tracking the state of the world*

The second step to semi-automation in SBMP is knowing when to send which message to whom. Each template in the message bank has a *send base* and *send offset* associated with it. The base can be several different dates: *study sign-up*, *due date*, *delivery date*, *loss date*, *family planning start date*, *visit date*. Each project will have different send bases depending on what outcomes the messages are targeting.

Regardless of what send bases are used, updating the messaging system when events in the real world change the participant’s status is key. Keeping the systems view of the world synchronized with the state of the world is a central challenge for any SBMP. When what

the system understands to be true is different from reality, the wrong message will go out and the value of creating highly targeted messages is lost. For example, if the system does not learn about a delivery until three months after the baby was born all messages for the twelve critical weeks in the early postpartum period will be skipped.

All of our projects have been co-located with clinics, but operating in parallel to standard of care. Study staff had offices onsite but separate from normal clinic workflow meaning all status updates and clinical information needed to be transcribed manually into the system. Not surprisingly this results in the mWACH system potentially holding an inaccurate version of the current participant status. Thus, at critical times such as after a delivery or in the case of an infant loss a participant continues to receive messages from based on their old status.

Our recommendation for minimizing these errors and maintaining better synchronization between the system and the state of the world is for any SMBP to operate within the clinic workflow. Tighter integration between standard of care and the system means that events are captured as they occur, reducing the amount of work needed to maintain system state. Integrating into existing workflows is important if SMBPs like mWACH are to scale beyond small trials.

In many ways the mWACH system acts as a simple electronic medical record (EMR) for each participant. None of the clinics we deployed mWACH at used an EMR at the time the studies started. However, some of the bigger clinics are starting to track participants using an EMR. Connecting mWACH to this data source would solve most of the system state synchronization issues. Full EMR's are notoriously hard to integrate with health systems. One model that has been successful, is for the project specific data collection platform to act as a mini-EMR. Two such examples are MomConnect [176] and 99 Dots [41]. Information from both of these projects is not only used to provide project level services but also shared with the local health systems for improved monitoring and evaluation.

3.3.3 Practitioner Dashboard

A central component of bi-directional messaging is the practitioner dashboard. As detailed in Section 3.2 this view switched from a conversational paradigm to a task oriented dashboard. The dashboard treats incoming messages as pending actions waiting for a response. All messages are expected to be responded to, but can easily be dismissed without a response if needed.

When responding to messages a participants detail page shows relevant context for the conversation. The participant's status and brief medical history is shown along with the most recent messages to or from that participant. There is also a notes section allowing staff to make comments on the participants record. Providing all this context is important because the same study nurse does not always respond to messages. It is often important to be reminded of previous messages and know when someone is due or delivered before responding. Exploring features such as making the notes system more visible and usable or adding frequently asked questions would make it even easier for practitioners to respond to messages.

3.3.4 Bi-Directional Communication

The second part of bi-directional messaging is having a transport layer to send and receive messages. For SMS the transport layer is usually a web based REST API connecting to either a third party service provider or a local Android phone running an app that polls a web hook waiting for SMS messages to send [194]. Using an Android app is a quick way to setup and works even where no mobile operator provides a gateway [153]; however, using a third party operator can scale when needed and does not require constant maintenance.

In the mWACH platform we have used the Nairobi based third-party provider Africa's Talking to access a five digit shortcode on Safaricom. By setting this shortcode to be reverse billed we made messaging free for all participants. The flow of messages from system to participant and back to the clinic is illustrated in Figure 3.5. A message from the study

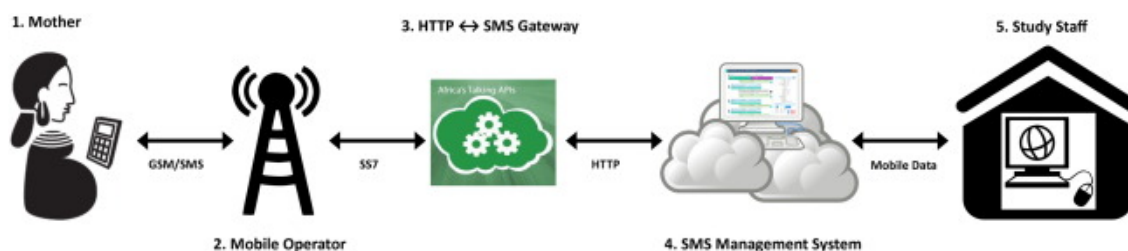


Figure 3.5: Bi-directional flow of SMS messages through the mWACH platform.

staff (5) is sent through the mWACH platform (4) to Africa’s Talking (3) which passes the message on to Safaricom (2) which uses telephony protocols to delivery it to the participants phone.

The Africa’s Talking gateway also provides delivery reports for SMS messages. These reports are sent to a call-back URL on the mWACH platform and report the status of a message once it either reaches or fails to reach the intended mobile number. Having delivery reports for messages is import for real time monitoring of system outages. Delivery reports are also useful for knowing which messages have been received and if a contact has changed numbers.

The transport layer is not limited to SMS – IVR, USSD, WhatsApp, or even a service like Slack could also be used – as long as there is an API for automated sending and receiving messages. In 2018 WhatsApp released a business API for sending automated messages through approved third-parties. The maternal health SMS project MomConnect [176] in South Africa has started testing WhatsApp for mothers who can use it. The primary advantage of WhatsApp over SMS or IVR is cost when sending large quantities of messages per week. However, since any project would still have to support non-WhatsApp users a shortcode for SMS would still be needed.

Chapter 4

ENCOURAGEMENT: MWACH V1

4.1 Overview

This chapter presents results from the first deployment of the mWACH platform – referred to as mWACH V1 – with a focus on understanding the user experience and SMS messaging patterns. This chapter contributes to the overall thesis by examining how a hybrid computer-human SMS system can engage pregnant women in Kenya in health-related communication. As the first deployment of mWACH the findings described in this chapter influenced the design and implementation of future projects in this thesis.

Data for this chapter focuses only on the 100 mothers in the two-way messaging group as part of a larger 12-month RCT at Mathare North Health Clinic on the outskirts of Nairobi [193]. This findings show that the mWACH approach is capable of engaging the majority of participants in health-related conversations. The computer automates the bulk-sending of personalized messages to patients, allowing the human to read patients’ replies and respond to those in need of attention. We show that receiving messages from the system triggers participant communication and the amount of communication increases as participants approach their expected due date. In addition, analysis of participants’ messages shows that they often contain sensitive health information conveyed through a complex mixture of languages and ‘txting’ abbreviations, all of which highlight the benefits of including a human in the workflow. Our findings are relevant for HCI researchers and practitioners interested in understanding or engaging underserved populations.

4.2 Introduction

The proliferation of mobile devices throughout the world, and particularly in low-income countries such as Kenya, is providing opportunities to create mobile applications that deliver health and information services to people living in poverty. Many of these applications focus on SMS, a low-cost communication channel that is supported on even the most basic mobile phones. However, the majority of existing SMS-based health applications use one-way ‘push’ messaging, in which generic messages are sent to large numbers of users who have no way to respond [53, 106] or they enable two-way communication between users and a server by requiring that users submit highly-structured messages that can be automatically parsed by the server [45, 130]. Although many SMS-based services would prefer to provide patients with personal attention that takes into account their specific context, they usually employ a small number of health workers who have limited amounts of time to engage in direct communication with individual users. This suggests a need for new approaches that increase the capacity of health workers and allow them to communicate effectively with a larger number of patients.

This chapter makes a novel contribution to HCI4D by investigating if a hybrid computer-human SMS system can engage pregnant women in health-related conversations. Unlike prior research, our system leverages the different strengths of *both* the human and the computer and we argue that each plays a critical role in communicating with and engaging patients. The computer automates the bulk sending of personalized, time-specific messages to a large number of patients, which allows the human to spend time reading patients’ replies and responding to those that have specific questions or concerns. This approach allows the human to effectively manage a much larger number of patients than would be possible if s/he was required to send all of the messages manually. Moreover, by allowing participants to send unstructured messages to the system in any format or language that they choose, we aim to allow users who may have little experience with SMS to participate and engage with the system.

We evaluate our hybrid approach through a 12-month deployment with 100 pregnant women in Kenya. Pregnant women in Kenya often have to wait for several hours at the clinic to see a nurse [142], which discourages them from attending their antenatal clinic visits and decreases the amount of advice and support that they receive during pregnancy. We use our hybrid SMS system to send pregnant women timely information that is tailored to their specific context and connect them directly to a nurse who is able to respond to their questions or concerns. This research was part of a larger RCT that aimed to assess if two-way SMS communication can influence antenatal clinic visits, infant immunization rates and adoption of family planning practices. However, discussing the long-term medical outcomes of the trial is beyond the scope of this chapter. Instead, we focus on understanding how low-income women use our computer-human SMS system to engage in health-related communication and we analyze the conversations that are enabled by our approach.

Key findings from the design and evaluation of our hybrid computer-human SMS system show that, despite most participants having little prior experience with SMS, our approach successfully engages the majority of participants in meaningful health-related conversations. We show that receiving automated messages from the system triggers participant communication and that the amount of communication increases as participants approach their delivery date. Finally, although prior work has presented large-scale quantitative analyses of SMS usage data [67], to the best of our knowledge we contribute the first in-depth analysis of unstructured SMS conversations with low-income women in Kenya. We show that participants used a complex mixture of languages and ‘txting’ abbreviations to communicate highly personal health information over SMS. Taken together, these findings suggest that our hybrid computer-human SMS system has the potential to engage users in meaningful health-related communication.

4.3 A Hybrid Computer-Human SMS System

The goal of our work is to use a hybrid computer-human SMS system to engage pregnant women in Kenya in health-related communication. Designing an SMS messaging system that

is appropriate for maternal health settings requires that we take into account a complex and nuanced set of design criteria. For example, the date at which a woman is predicted to give birth and the date at which she actually gives birth can vary by several weeks. As a result, it is often impossible to accurately predict when a woman will give birth and automatically switch her from antenatal to postpartum messaging. To overcome challenges like this, our hybrid SMS system has been designed to incorporate input from *both* a human nurse and automated software, with each playing a critical role in the resulting communication.

4.3.1 Design

At a high-level, our approach allows automated SMS scheduling software to handle the majority of outgoing messages. This greatly reduces the total number of messages that the nurse is required to manage, allowing her to instead spend time reading and responding to unstructured messages submitted by participants. In this way, the nurse is able to effectively communicate with a larger number of patients than would be possible if she had to send all of the messages manually. Keeping in mind the challenges of designing for maternal health settings, we identified the following criteria that influenced the design of our system:

Personalize messages

As much as possible, we personalize the content of the messages to each individual woman's specific context. We send messages that take into account gestational age and provide tailored message tracks for groups of participants with special circumstances, such as first time mothers, adolescent mothers, or women who have had previous cesarean sections. Moreover, we allow participants to choose the language (English or Swahili) that they receive messages in, and the day of the week and time of day that the messages arrive.

Encourage engagement

Our approach encourages women to engage with the system in several ways. First, each message includes a simple question relevant to the informational content of the message. Answering the question provides a natural method for engagement. Second, messages that women send to the system do not have to be structured or specially formatted in any way. Instead, women can submit questions, concerns or updates in any format and in whatever language they prefer.

Provide professional health care

Since we cannot predict the nature or content of the messages that women send to the system, it is important that responses come from qualified health professional familiar with the local culture and language. Moreover, the information in the automated messages was carefully crafted by maternal and child health experts to ensure that it was both relevant and medically accurate. Connecting underserved women directly to a health professional was one of our major goals.

Increase the nurse's capacity

To allow the nurse to easily communicate with a large number of women we designed a laptop-based interface that provides a portal through which the nurse can view and reply to messages (see Figure 4.1). The center of the interface depicts messages that have been sent and received by the selected participant as a conversation. Messages from the system, the nurse and the participant are displayed in different colors so that it is easy to interpret the participant's communication history. At the top of the interface is an input box that the nurse can use to send messages to the participant. The left side of the interface provides a searchable list of participants with new messages filtered to the top of the list. This allows the nurse to quickly locate and view unread messages that she has not yet responded to. Finally, the right side of the screen displays the participant's visit history and health status, such

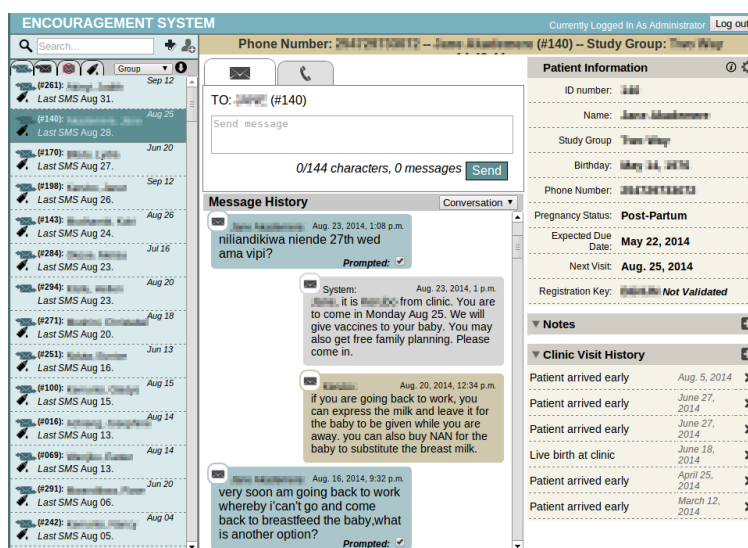


Figure 4.1: The laptop-based interface that the nurse uses to manage all participant communication. The system has a list of searchable contacts on the left, the communication history of the currently selected participant in the center, and mechanisms for tracking the participant's visit history and health status on the right

as her next scheduled clinic visit or when she is expected to give birth. Using this interface, the nurse can quickly understand a patient's prior communication, visit history and medical context and personalize her responses accordingly.

Adapt to participant circumstances

The unpredictable nature of pregnancy means that the sequence of outgoing messages may need to be dynamically adjusted based on a participant's changing circumstances. For example, if a woman gives birth earlier than expected, we need to be able to change the messages that she is receiving accordingly. Thus, rather than defining a fixed set of messages to be sent at sign-up, the messages sent to participants are instead dynamically selected according to each woman's current status at the time of sending. If the nurse changes the status of a woman in the system, the messages that she receives will automatically adapt to her new status.

Allow participants to respond only when they want to

Several prior studies [53, 195] used escalating reminder systems in which participants received a telephone call when they did not respond. Our original design emulated this model. Initially, if a participant did not respond to a message, the message would be repeated every 48 hours for four days until a response was received. If no response had been received at that point, the system would prompt the nurse to call the participant. However, early feedback from focus groups with health workers and pregnant women revealed that this interaction model was overly demanding, and we changed our design to allow participants to respond only when it made sense.

4.3.2 Message Content and Timing

The content of the SMS messages sent to participants was carefully crafted by maternal and child health experts using Kenya's antenatal care guidelines. In addition, we conducted focus groups with 22 pregnant women and 10 health workers prior to the start of the study to obtain feedback and iterate on the message content. Findings from these focus groups suggested that participants really wanted to feel as though someone cared about them. In addition, they wanted the messages to be directed specifically to them and take into account their personal situation. To achieve these goals, we designed the message content to include the name of the participant, with each message appearing as if it came from the study nurse (see Table 4.1). The messages were then further tailored to each participant based on her specific expected delivery date and covered a range of health and pregnancy related topics, including birth preparedness, factual information on what to expect, family planning, breastfeeding and immunization. The messages were then further personalized based on the participant's current pregnancy status, such as being a first time mother or adolescent. Finally, as shown in Table 4.1, each message that was sent by the system included a specific question designed to encourage the participant to respond.

Based on each participant's specific criteria, the system automatically selected and

dispatched the relevant pre-programmed message at the appropriate time. In several cases, the final automated message was longer than the maximum size of a single SMS (140 characters) and was broken into two messages that were sent at the same time (see Table 4.1(b) for an example). In addition to health related information, participants also received automated appointment reminders two days before each scheduled clinic appointment (see Table 4.1(c)). Messages were also sent on the day of enrollment, on the day a clinic visit was recorded, and when a delivery was recorded.

Finally, women’s delivery dates in Kenya are typically calculated based on the date of their last menstrual period. However, this method is frequently inaccurate and, in reality, they give birth anywhere from several weeks before to several weeks after their predicted date of delivery. Thus, if a participant had not given birth by the time all of the scheduled antenatal messages had been sent, the final antenatal message was repeated for an additional two weeks or until a birth was recorded in the system, at which point they began receiving tailored postpartum messages. In total, each participant that completed the study received approximately 35 automated messages in the seven to eight months between enrollment and completion, with at most two system messages per week.

4.3.3 Implementation

Our system is a python-based Django web application that connects to the Kenyan telecommunication system through Shujaa Solutions¹, a now defunct third party SMS gateway that provides an API that our server can use to post outgoing messages. Incoming messages from participants are routed to our server via a callback URL. In addition, by establishing a reverse billing five-digit short code, we were able to allow participants to send and receive messages free of charge.

¹Shujaa Solutions: <http://shujaa.co.ke>

<p>(a) Automated message at 27 weeks pregnant [Name], this is [Nurse] from clinic. Iron helps carry nutrition to your baby. If it is low, you feel tired. Are you taking iron or do you need tablets?</p>
<p>(b) Automated message 8 weeks after giving birth (1) [Name] this is [Nurse] from clinic. Are you having trouble breastfeeding? Giving the baby other fluids can cause illness and weakness. (2) Please continue to only give breast milk to your baby for at least six months. How long do you plan to breastfeed?</p>
<p>(c) Automated clinic appointment reminder [Name], it is [Nurse] from clinic. You are to come in Friday Aug 08. We will give vaccines to your baby. You may also get free family planning. Please come in.</p>

Table 4.1: Examples of automated messages sent by the system at (a) 27 weeks pregnant; (b) 8 weeks after birth; and (c) 2 days before a scheduled clinic visit. Note that, as is the case for (b), any messages longer than 140 characters were sent in two parts.

4.4 Evaluation

To understand women’s engagement with our hybrid SMS system, we analyzed messages sent during a year long deployment with 100 pregnant women in Kenya. Our work is part of Mobile WACH, a larger IRB-approved clinical trial investigating maternal mobile health initiatives and behavior change. The deployment was based at Mathare North Health Center, a city council clinic located approximately 10km from Nairobi, the capital city of Kenya. Mathare provides basic health services to patients, including normal antenatal services for pregnant women, but refers all complicated cases to a higher level health facility.

4.4.1 Study Nurse

A dedicated nurse was hired for the duration of the study to enroll participants, respond to SMSs, keep track of anticipated and completed patient visits, monitor when participants gave birth, and conduct post study surveys. The nurse was female, computer literate and

had experience reading and responding to email and participating in Skype conversations. We provided the nurse with a laptop and access to the system management portal so that she could communicate with participants. Prior to the start of the study, the nurse received roughly two hours of training on how to use the software.

4.4.2 Participant Enrollment

Starting in August 2013, pregnant women visiting the clinic for their first or second antenatal visit were invited to participate in the study. Volunteers needed to have personal or shared access to a mobile phone, be able to read and respond to messages in English or Swahili (the two national languages of Kenya) and be planning to reside in the area for the duration of their pregnancy. In addition, due to technical limitations of the SMS gateway, participants were required to possess a SIM card for the telecommunications provider Safaricom. Fortunately, the majority of Kenyans already use Safaricom. During the entire study only two people failed to meet the inclusion criteria, and this was because they did not possess a Safaricom SIM card. In total, we enrolled 100 women to receive messages from our hybrid SMS system. All participants sent and received messages using their own personal or shared mobile phones. During enrollment the nurse explained that participants would not be charged for any messages sent to the five-digit study number. All participants received a small monetary compensation for their participation.

4.4.3 Procedure

After screening volunteers and obtaining informed consent, the study nurse conducted a structured interview with each participant in which she collected the participant's demographic information, medical history, and current usage of and prior experience with technology. She also collected their mobile phone number and explained that they would receive one to two messages per week, starting at 20 weeks of pregnancy and continuing until 10 weeks postpartum. She explained that the messages would contain questions about their health and that they should respond, although she also made it clear that they should not use the service

for emergencies. Each participant also received a business card that contained the study's five-digit short-code. After enrollment, the study nurse did not meet with participants again until several weeks after they had given birth, although participants did see other clinic nurses when they attended their normal antenatal appointments. All participant communication with the study nurse was through participants' personal mobile phones.

4.4.4 Data Collection and Analysis

The goal of our paper is to investigate if we are able to engage pregnant women using our hybrid computer-human SMS system and to characterize the resulting conversations that take place. To achieve this goal, we analyzed data from 100 structured interviews conducted with participants at enrollment, 3661 automated messages sent by the system, 423 messages sent by the nurse and 944 messages sent by participants, and data from two 45-minute semi-structured interviews with the study nurse that were conducted over Skype.

We began by examining the data from interviews conducted at enrollment and characterized the population of women that participated in our study, organizing them according to age, income, language, education level, employment, and their current usage of and prior experience with technology. Following this, we conducted an iterative, in-depth analysis of messages that participants sent to the system. We began by noting the language that participants sent messages in (English, Swahili or a mixture) and whether each message contained abbreviations or 'texting'. One of the authors, who speaks Swahili, translated any messages that were in Swahili into English. Then, we coded any messages that (1) answered a question posed by the system; (2) answered a question posed by the study nurse; (3) asked a question related to information sent by the system; (4) asked a new question that was not related to anything sent by the system; and (5) provided the nurse with an unprompted update or information. We further identified any messages that formed a 'conversation', which we define as a back-and-forth sequence of at least two related messages. For each identified conversation, we then further analyzed the messages that made up the conversation to identify who initiated the conversation (the system or the participant) and any themes, events or

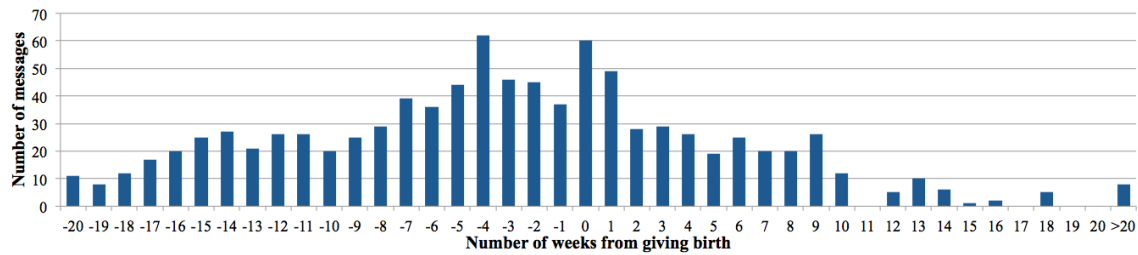


Figure 4.2: Total number of messages sent during each week before and after delivery. Week 0 corresponds to the week in which participants were recorded as giving birth. Week -5 depicts five weeks before giving birth while 5 shows five weeks after giving birth

topics that caused participants to engage in conversation.

4.5 Findings

4.5.1 Understanding Our Participants

Before analyzing participant engagement with the system, it is important to understand the characteristics of the women in our study and their prior experience with technology.

Demographic Characteristics

Our participants were all pregnant women aged between 16 and 37 years ($M=24.1$, $SD=4.6$). 38 participants reported that they had attended primary school, 51 had started (but not necessarily finished) secondary school, while 11 had some form of post-secondary education. The majority of Kenyans in Nairobi speak a complex combination of Swahili, English, slang words and local languages known as Sheng. However, since the language of education and government in Kenya is English, most of our participants prefer to receive formal communication in English as opposed to Swahili, and 97 participants chose to receive SMS messages in English. 86 participants reported that they live in single-room houses, 94 within 10km of the clinic, and 88 walked to the clinic for their appointment. 85 participants said that they were self-employed or housewives with family sizes ranging from 2 to 6 people

($M=3.1$, $SD=1.3$) with a mean monthly household income of 6500 ($SD=3728$) Kenyan shillings (roughly US\$73). These characteristics suggest that, despite living close to Nairobi, Kenya's capital city, many of our participants are living in extreme poverty [167].

4.5.2 Measuring Participant Engagement

We use the number of messages that participants sent to the system as our measure of engagement. Thus, women who communicate more by sending more messages have a higher level of engagement. The total number of weeks that participants received messages from the system varied according to when they enrolled in the study and when they gave birth. To account for this variance, we divided the total number of messages that each participant sent by the number of weeks that they were receiving messages to obtain our final engagement metric, which we report as the average number of messages that each participant sent per week of participation.

The majority of participants engaged with the system

Between August 23, 2013 and August 31, 2014 we received a total of 944 SMS messages from participants. During the same time period, the system sent 3661 automated messages and the study nurse sent 423 messages to participants. On average, each participant sent 0.46 ($M=0.26$ $SD=0.54$) messages per week or two messages per month for the 7-8 months that they participated in the study, with 73 of our 100 participants sending at least one message. In addition, many participants communicated on a regular basis throughout the study (see Figure 4.2) with the highest amounts of engagement recorded during weeks close to participants' delivery dates. This high level of participant engagement is encouraging, particularly since the majority of our participants did not have much experience with SMS and prior work has struggled to engage these kinds of users with SMS communication.

Participants were more engaged close to their delivery date

To analyze how participants' communication varied throughout pregnancy, we calculated the total number of messages that participants sent in relation to when they gave birth. Our findings are shown in Figure 4.2. Week 0 in the graph shows the week that participants were recorded as giving birth, with values before 0 showing weeks before giving birth and values after 0 showing weeks after giving birth.

In general, participants' engagement with the system increased as they approached their delivery date. This suggests that participants were more motivated to communicate the closer that they were to giving birth. In addition, the questions that the system sent to participants in the third and fourth weeks prior to their expected delivery date specifically asked, “*Are you worried?*” and “*Do you have any questions about where to go in labor and how to get there?*” Similar open-ended questions were asked in other automated messages, however, they elicited the most responses in the weeks just prior to delivery, which indicates that this was an opportune time to provide participants with someone to whom they could direct their questions or concerns about the process of giving birth. After delivery, the number of messages that participants sent gradually decreased again, with a notable drop in messages submitted more than 10 weeks after birth. Since participants stopped receiving messages from the system at 10 weeks postpartum, it makes sense that, for the most part, they also stopped sending messages to the system.

Receiving messages triggered engagement

The vast majority of messages sent by participants were on the same calendar day as they received a message from the system (705 out of a total of 944). A large number of these were direct responses to questions that the system had asked. However, many also triggered additional questions unrelated to those posed by the system, with 39 different users each submitting at least one new question. For example, one participant received the following automated message:

"[Name]² this is [Nurse] from Mathare clinic. Regular, strong stomach pains are a sign of labor. If you feel strong tightening of your belly, leaking of fluid or any bleeding go to the clinic. Have you had any labor pains? How often do you feel them? Are you worried?"

40 minutes after receiving this message, the participant sent³,

"Hi,am not as such worried,but i have a little fear.i have never had any labor pains."

Then, ten minutes after this initial response, she sent,

"Hi.madam [nurse],kindly will u state for me items that are needed when i will attend clinic during labor pains e.g clothing for the baby,basin,.and what els....."

to which the nurse then responded,

"Hi [name], you need baby clothes, shawl, lesa, basin, spirit, cotton wool. that is mainly it. the rest you can inquire from the nurses when you reach the hospital."

Exchanges like this suggest that receiving messages from the system reminded participants that they had someone to whom they could direct their questions or concerns and highlight the potential for the automated messages to trigger engagement.

Some participants did not respond

Twenty-seven participants did not engage in any communication during the study. We tried to contact these participants to understand some of the reasons for their lack of response and were able to successfully reach eight participants. Three explained that they shared

²All names have been removed to preserve anonymity.

³All messages quoted in the paper show participants' original spelling and grammar.

a phone with their husband, which made it difficult for them to read and respond to the messages. In one case, a participant said that her husband deleted the messages before she had a chance to read them. In several other cases, the person who answered the call told us that the participant had changed their phone number during the study. Moreover, several of the phone numbers that participants had provided to us appeared to be incorrect, with the number either not being a valid phone number or with the person who answered the call claiming to not know the participant.

4.5.3 Understanding Participants' Use of Language

Kenya has two national languages, English and Swahili. Our findings show that participants' messages frequently included a complex mix of English, Swahili and Sheng (an informal mixture of the two), 'txt' abbreviations, and numerous spelling and grammatical errors, all of which highlight the benefit of having a human interpret and respond to messages.

Participants sent messages in a mixture of languages

Since 97 participants chose to receive messages in English, it makes sense that the majority of messages sent to the system (784 out of 944) were also in English. However, since only three participants chose to receive messages in Swahili, we were surprised to find that 41 participants sent messages that contained at least some Swahili, with 13 communicating entirely in Swahili. Follow-up conversations with the nurse showed that although some participants were more comfortable with Swahili, they feared that the messages would contain formal Swahili rather than the colloquial language that they typically use. A total of 73 participant messages were entirely in Swahili and a further 87 contained a mixture of Swahili and English. Our analysis of these messages suggests that participants mix Swahili and English freely, choosing whichever word comes to mind first. For example,

"my headache is ve ly pain and kizuguzugu"

The word 'kizuguzugu' (a misspelling of 'kizunguzungu') means 'dizzy' and, although we

cannot be sure, we hypothesize that this participant was either unsure of the English word for dizzy, or simply more familiar with the word in Swahili. Other participants more blatantly used Sheng, mixing English and Swahili whenever was convenient and incorporating new words not pulled from either language:

"My nxt visit will be on 29th may.na nina sweat sana siku hizi xana xana usiku.can u elaborate on this pliz"

The Swahili part of this message means ‘I am sweating a lot these days a lot a lot at night’. The first time the participant writes ‘a lot’ in Swahili, but the second and third time it is written in the slang version commonly used in Sheng. Although this complex mixing of languages may be understood by a bilingual Kenyan, it would be extremely difficult for any language processing software to interpret these messages.

Participants frequently used ‘txt’ abbreviations

‘Txting’ refers to the use of abbreviations and other techniques to craft SMS messages, often with the aim of reducing the overall length of the message [24]. Analysis of our participants’ messages shows that 30% or 287 messages used ‘txting’ that often made the messages difficult to interpret:

"the 1st thing iwould like 2say hi 2u hpng dat ur fine.my qstn is dat when ivisted clnc last mnth itryd 2expose my complication 2the dctr dat the lab told 2hav some medicine coz of infctn bt they gav out 2 me some amoxil capsols 4me 2use 2x3 pa day bt whn iused them 4the 1st 4day iwas flng stomach strng pain den ijust chsn my self 2surender tkng thm did iwrng or icld cntnu usng them bcz its how they use 2cure"

In addition, several participants appeared to not know how to input certain characters into their phone. For example, one participant responded to a question about heartburn with,

"yesiamhavingheartburninaweekitocurstwice"

These findings suggest that many participants would find it difficult to compose messages that required specific formatting and it is an advantage that our system allows participants to send unstructured messages. In addition, the mix of language, 'txtng', and errors highlights the importance of considering the cultural and linguistic context for automated messages and human replies.

4.5.4 *Understanding the SMS Conversations*

To understand the conversations that arose from participants' engagement with the system, we labeled any sequence of two or more related messages as a 'conversation' and analyzed the length and content of these conversations. In total, we recorded 703 conversations that took place during the study, ranging in length from two to ten messages (see Figure 4.3). 529 conversations were initiated by automated system messages, while the other 124 were initiated by participants.

Conversations with two messages: question and answer

The majority of two-message conversations fall into two categories: (1) the automated message asks a question and the participant responds, and (2) the participant asks a question and the nurse responds. Participant responses to system questions ranged in length and complexity from simple 'yes' and 'no' answers to longer, in-depth updates, such as,

"Am not having heartburn but am having alot of pain at the side of stormach and at the back.sometime it enter in the legs that i cant even make a single step till the pain disapear.then there is something coming up from stomach such as am unable to breath."

There were 73 additional questions posed by participants that the nurse was able to answer with a single message. Many of these questions demonstrate the value of connecting participants to a health professional. One message asked,

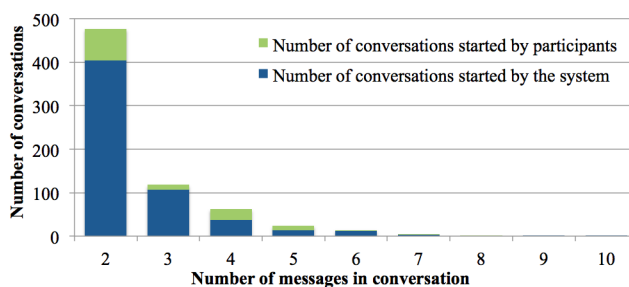


Figure 4.3: Graph showing the number and length of conversations (two or more related messages) started by the system and by participants.

"I have a question that is disturbing me, can a pregnant woman be dewarmed?"

while another inquired,

"Hey [nurse] i ga ve birth on satudy bt my bby is cryng of stomach is it advaisable i giv ha medicine?"

It is unlikely that these participants would be willing to wait in line for hours at the clinic just to ask these kinds of questions, which suggests that our system successfully lowers the barrier between our participants and health-related information.

Conversations with three or more messages

The majority of conversations longer than two messages were initiated by the system and then continued by the participant and the nurse. In many cases, participants would respond to a system question with a related question. For example,

System: *"[Name] this is [Nurse] from Mathare clinic. When you come in for labor, we will give you a birth notification for the baby. Delivery at the hospital could save your baby's life. Do you have any questions about where to go in labor and how to get there?"*

Participant: *"My worry is how 2 get in the hosp. Becoz of insecurity in our*

area incase labor starts at midnight. Another question is when i get at the hosp. 4 example at mathare at midnight will they able 2 open 4 me?"

Nurse: *"Hi [name], This is [nurse]. The maternity unit at Mathare operates for 24 hrs a day, so if you go there at midnight they will definitely attend to you. Thank you."*

This highlights the value of including carefully crafted questions in the automated messages that encourage participants to share their questions and concerns. Other conversations took place solely between a participant and the nurse. In these cases, a participant usually contacted the nurse seeking additional information or advice. For example,

Participant: *"how about if i'give the baby other liquids,like cow milk what will happen?"*

Nurse: *"Hi [name]. How old is your baby? If the baby is below 6 months, we advise you give only breast milk because the baby's digestive system is not yet fully developed to digest other foods. If i may ask. Why do you want to give cows milk?"*

Participant: *"she is almost three months. very soon am going back to work whereby i'can't go and come back to breastfeed the baby,what is another option?"*

Nurse: *"if you are going back to work, you can express the milk and leave it for the baby to be given while you are away. you can also buy NAN for the baby to substitute the breast milk."*

The longest conversation that we identified consisted of 10 messages in which the participant asked four different questions that were answered by the nurse. After receiving her answers, the participant concluded the conversation with,

"Thanks alot. Those are some of the questions that were really bothering me."

Participants communicated sensitive personal information

Analysis of participants' communication revealed that many were willing to send personal health information over SMS. Several inquired about sexually transmitted diseases:

"i'was tested 4 siphilis on monday.is it posible someone can stil have other std?"

while others asked if it was safe for them to have sex:

"Hi [nurse],am gud since we were 2getha.just wanted 2 ask u whetha a baby can b affected by sex while in the womb apart from sexual transmitted diseases."

Interestingly, some participants communicated sensitive information through updates without asking a specific question:

"Thanks a lot.I don't have pains nor bleeding. If I can cough or sneeze I notice little urine come out. Am satisfied in early stage is okey. Am grateful for your concern."

Participants' willingness to communicate this personal information over SMS has several potential implications. For example, since participants may be hesitant to talk about embarrassing or sensitive topics in person [95] the SMS system could provide them with a more comfortable communication channel for airing their concerns. Finally, many participants expressed gratitude for the support that they received:

"Dank U 2, 4 Being Consern About My Health. Am Not Realy Bad, But We R Stil Struggling With Life & Pray 2 Survive It. Coz 4 Me It's Had."

These messages are particularly rewarding because they were unsolicited, which gives us reason to believe that participants were genuinely grateful and that it made a difference to them to feel as if someone cared about them.

Some participant questions went unanswered

For our hybrid computer-human SMS system to be successful, it is crucial that the nurse is also engaged and sends timely responses to participants (although we made it clear during enrollment that participants should *not* use the system as an emergency service). Our analysis of participant messages revealed a total of 50 questions that were never answered. In addition, in some cases the nurse responded several days after the message was received. Although the nurse's lack of response to messages did not have any observable effect on participant engagement, we plan to change the design of the interface to help the nurse do a better job of responding to messages. First, although the software made it easy to see unread messages, the nurse would often read a message but not respond immediately. However, after reading messages, it was difficult to keep track of messages that she had read but not responded to. To overcome this challenge, we plan to change the design of the system management portal so that, in addition to highlighting unread messages, it also flags messages that have not yet received a response. The nurse also described that she feared the laptop would get stolen if she carried it on public transit or left it at the clinic overnight. Thus, she frequently left the laptop at home when she went to the clinic for the day and was then unable to read and respond to messages until she got home in the evening. To overcome this issue, we plan to design a mobile version of the software that the nurse can use on her personal Android phone.

4.5.5 Themes that motivated communication

We identified several common themes that motivated women to reach out to the nurse. In general, we observed that a large number of participants were willing to send extremely detailed and personal health information over SMS. Understanding participants' opinions regarding the security and privacy of their messages will be interesting future research.

Concern

A large number of participants' messages were motivated by concern, both for their baby and for themselves. Many of these messages included phrases like "*is it safe*", "*is it harmful*" or "*is it normal.*" Participants frequently inquired about foods that were safe to eat or feed to their baby:

"does eating avocados makes the baby to grow big that he/she may not come out easily?am confused about eating some foods."

After giving birth, participants' concern was primarily for the welfare of their baby, and they would frequently reach out to ask for advice about what to do,

"My child is hot, coughing, has rashes, flu, loss of appetite, what do i do?"

In some cases the nurse was able to provide answers that saved the participant an unnecessary visit the clinic. However, for any potentially dangerous scenarios, she always advised participants to visit the clinic as soon as possible,

"please take your baby to the doctor tomorrow so that she can receive medical attention."

Professional Advice

Many participants submitted questions that asked for clarifying information about rumors or other beliefs. These messages often included the phrases "*is it true*" or "*some people say*" For example,

"is it true that drinking chocolate in pregnancy triggers bleeding when giving birth?"

These kinds of messages are highly encouraging because they provide an opportunity to give participants medically accurate advice and dispel rumors that may be untrue or even harmful. In addition, follow-up messages suggest that women appreciate and heed the advice provided by the nurse,

"Thank you very much for your advice ill make sure i follow it."

A variety of other messages asked the nurse for factual information, such as if the clinic operated 24 hours a day, what to bring with them when they go into labor, or where to obtain a birth certificate for their child. Several also forgot the date of their next clinic appointment and used the system to find out,

"Plz idont no the date of going to clinic will u help me?"

These messages suggest that it was beneficial for participants to have a connection to a health professional who was able to provide them with the information that they sought.

Gratitude

Finally, many participants expressed gratitude for the support and encouragement provided by the system. One sent,

"Hi.madam [nurse],thanks for ua informations.ua messages are really encouraging,"

while another communicated,

"Dank U 2, 4 Being Consern About My Health. Am Not Realy Bad, But We R Stil Struggling With Life & Pray 2 Survive It. Coz 4 Me It's Had."

These messages are particularly rewarding because they were unsolicited, which gives us reason to believe that these participants were genuinely grateful for the information and advice that they received and that it made a difference to them to feel as if someone cared about their welfare.

4.6 Discussion

Our findings clearly demonstrate the potential for hybrid computer-human SMS systems like ours to engage women in health-related communication during pregnancy and postpartum. We believe the success of our approach relies on the contributions of *both* the computer and the human. Since participants only met the nurse once at enrollment, we argue that it is not only the personal face-to-face connection to the nurse that causes them to engage with the system. Rather, it is the combination of carefully crafted personalized messages, the perceived relationship with the ‘nurse’ created by the automated messages *and* the messages sent by the nurse, each of which contribute to participants’ motivation to communicate. The streamlined, tailored system messages provided a foundation on which further communication was built, and participants’ responses to these messages suggest that they perceived the automated messages to come directly from the nurse rather than an automated system. With the system automating the bulk of the sending, the human was able to fulfill the higher order demands for advice and person-specific care. In addition, the varied topics and questions messaged by participants and the unpredictable nature of pregnancy motivate the need for a human to dynamically and adeptly control the messages that are sent. Participants often asked sensitive medical questions that require a trained professional to answer or triage to higher-level in-person care.

Taken together, these findings suggest that our system is able to supplement currently overstretched health care systems in a meaningful way. Not only do hybrid systems like ours allow for health education, appointment reminders and information dissemination, they may engage women in their own care, creating more informed patients and allowing for triage of both necessary and unnecessary in-person clinic visits. Moreover, enabling this communication requires little in terms of direct or indirect costs to the participant (*e.g.*, no travel time, transportation, waiting to see a nurse) and may decrease unnecessary visits to the clinic, which could help to reduce clinic waiting times for all patients.

The use of a hybrid messaging system does raise an interesting ethical debate: participants

may think that the automated messages are being personally composed by the nurse. Our hybrid design is a compromise between an entirely automated system and a manual one. We use technology to efficiently accomplish what a human could do manually without significantly changing the process or workflow. Moreover, in our case, the outgoing automated messages have been carefully crafted by global health experts and tailored to each participant's health status. In addition, all participant messages and questions are reviewed and responded to by a qualified nurse. We believe that this approach is ethically acceptable, allowing underserved women to receive personal advice from a trained health professional, which would not be feasible if the nurse had to send all of the automated messages manually.

Our study also exposed several interesting design opportunities that we plan to explore. For example, we would like to make it easier for nurses to keep track of questions sent by women. One potential approach could be to model the web-based interface as a 'ticket management system' or email client instead of using an SMS-focused paradigm. Another interesting addition would be to allow the nurse to create and save template responses to frequently asked questions. For example, several participants asked about hospital operating hours, the safety of eating particular foods, vaccine schedules *etc.* Having template answers to these questions may help the nurse to respond more quickly and efficiently.

As we consider how to scale our hybrid system to larger deployments, we need to carefully consider how best to preserve the privacy of patients' personal health data. In our study, we explained that participants would receive messages on their own phones and we allowed them to choose the time of day that messages would arrive in part to allow them to choose a time when they would likely be alone. To avoid accidentally revealing a woman's pregnancy to her husband or family, we only began sending messages at 20 weeks of pregnancy. At the clinic, access to the messages was limited to the study nurse who logged in with a password. We hope to learn more about patients' opinions on privacy through follow-up interviews at the conclusion of the ongoing medical study.

Finally, an important consideration in any HCI4D project is the sustainability of the system. Since women were able to send messages to the system for free, we are unable to

tell how engagement would change if women had to pay for messages. The cost of each incoming or outgoing SMS was US\$0.03. We paid for a total of 5028 messages, costing a total of US\$150.84. Our biggest cost was a US\$300 monthly fee to keep the SMS gateway open. However, this is a fixed fee that will not increase as the system is scaled up which means that (at least for the foreseeable future) the cost per participant will decrease as the number of participants increases. We also anticipate that additional challenges will arise as we scale the system to a large number of clinics and health workers.

4.7 Conclusion

This chapter describes a hybrid computer-human SMS system that we designed to encourage pregnant women in Kenya to engage in health-related communication. Findings from a 12-month deployment of the system involving 100 pregnant women show that the system successfully engages the majority of participants by leveraging the different strengths of *both* the computer and the human. Moreover, analysis of the resulting conversations that took place between the system, the nurse and participants suggest that our hybrid approach provides a valuable communication channel that connects pregnant women directly to a qualified health professional. Taken together, our findings contribute a valuable and novel dimension to the growing body of HCI4D research that aims to understand and engage underserved populations.

Chapter 5

MWACH XY: MALE PARTNER ENGAGEMENT

5.1 Overview

Maternal health outreach and engagement is a common goal of mobile health (mHealth) projects for development. Male partners of pregnant and postpartum women are known to be important influences on their health behavior. This chapter presents a novel extension to the mWACH platform presented in Chapter 4 by exploring how to engage male partners in SMS-based family planning conversations. The two primary contributions of this chapter to the overall thesis are: 1) the design and implementation of 3-way messaging for inclusion of male partners in a semi-automated bidirectional SMS platform and 2) the deployment of mWACH for family planning focused decision making and support.

A 12-month randomized controlled trial of a family planning counseling SMS program was conducted in western Kenya using our system. A total of 260 pregnant women and 101 of their male partners were enrolled in the system. We analyze enrollment and usage data from this trial to compare baseline technology use and demographics of mothers and their male partners. Our findings demonstrate significant technology gender divides exist within the study population. Finally, we explore how both the mothers and their male partners interacted with the SMS system through an analysis of over 11,500 messages. We conclude that it is feasible to include rural Kenyan men in an SMS-based family planning discussion program and that their inclusion does not dramatically affect the mothers' engagement.

5.2 Introduction

According to the 2014 Kenya Demographic and Health Survey (KDHS) [140], 40% of Kenyan women using modern contraceptive methods were not informed about the potential side

effects. Moreover, 31% of women reported discontinuing contraceptive use within 12 months, and of this group, 29% cited health concerns or side effects as a primary reason for stopping. Studies have also highlighted multiple provider-level barriers to distributing information about and accessing family planning [191]. This results in 18% of married Kenyan women expressing an unmet need for family planning – an unmet need that is to some extent driven by a gap in knowledge and understanding.

The use of SMS for information access and dissemination has been well documented in HCI4D and Global Health fields [31]. However, most of the work on maternal health has focused on reaching pregnant women at clinics or at home through community health workers [53]. Systems that reach out directly to mothers are increasingly common [150, 171]. Studies suggest that reaching out to male partners during prenatal care and about family planning decision-making may improve outcomes [211], but few interventions have included them. In this chapter we examine the feasibility of including male partners in a maternal health SMS intervention and how this affects women’s participation and engagement with the intervention. This work presents three main contributions to the HCI4D community:

1. A description of the design considerations and system modifications necessary to incorporate male partners into sensitive maternal health conversations over a toll-free SMS gateway.
2. A characterization of the extent of the gender-based technology divide within the study population.
3. An evaluation of the participation and engagement of mothers and male partners in a family planning SMS intervention.

Not surprisingly, we show that on most measures of technology access and usage women have less access than men. Our contribution in this regard is to show that this gender divide exists within couple dyads. Despite this gender divide mothers engage significantly more than their male counterparts with the same family planning and maternal health SMS content.

However, when men do engage, they do so in much the same way as their female counterparts. They respond quickly, talk about both family planning and maternal health, and show a willingness to ask personal questions.

5.3 Related Work

In HCI4D projects for community outreach, the need to design universally available services constrains the design space to simple text and voice based applications. These are accessible on even the most basic mobile phone [151]. From IVR and live operator systems, to SMS and data-enabled systems, this design space has been well documented [121, 138]. In Sub-Saharan Africa, automated and bidirectional SMS has been used for projects as diverse as vaccine cold chain management [14], financial reimbursements [51], and medication reminders [113]. SMS systems, such as UNICEF's U-Report, have grown recently to large multi-country SMS platforms with millions of users [2]. Sub-Saharan Africa is ideal for these projects due to the low cost and universality of SMS, use of Latin-based scripts, and the prevalence of SMS gateways.

A primary focus of mHealth SMS interventions has been on maternal health [11, 150, 171]. One of the largest projects in this area is MomConnect, which runs nationally in South Africa and has connected over half a million mothers to health information via SMS [176]. Another maternal health project to have reached scale is m4RH, which is an opt-in reproductive health SMS service in Kenya and Tanzania [95]. Through a series of SMS-based menus, m4RH allows people to request information on different family planning methods. This differs from the semi-automated approach on which our study is based [150], where weekly SMS prompts are sent to all participants and a trained clinician answers individual questions [150].

One concern of mHealth interventions in HCI4D is that they can amplify existing gender inequalities [206]. Surveys of initial m4RH users showed that while a majority of users are women, a substantial number of men queried the database [109]. A systematic review of gender relations in mHealth projects revealed that many projects only reinforce existing gender inequalities. [93]. To date there has been very little analysis of how men participate in

Partner Status	Control	Intervention	Total
	130	130 (n_i)	260 (n_a)
Not Invited	52	49	101
Invited	29	29	58
Invited + Enrolled	49	52	101
Enrolled + Phone Shared	12	9	21

Table 5.1: Male partner status of enrolled female participants. Total study enrollment was 260 n_a with 130 participants in the intervention arm n_i . Women’s enrollment is further broken down by male partner status: not invited, invited (but not enrolled), both invited and enrolled, and finally enrolled but sharing a phone.

SMS messaging projects for maternal health and family planning. Therefore, it is important to understand how the inclusion of male partners in an SMS intervention affects the overall engagement with the system.

5.4 Methodology

While the focus of this work is not on clinical outcomes, the background of the larger randomized control trial (RCT) is necessary to fully understand the design constraints and context. This work represents the collaboration between the University of Washington (UW) Department of Computer Science and Engineering, the UW Department of Global Health, and the Kenyatta National Hospital/University of Nairobi in Nairobi, Kenya. The first author built the entire SMS platform in close collaboration with public health colleagues who designed and implemented the RCT. A study team of nine Kenyans with a diverse background in nursing, data management, and field work was hired for the duration of the study.

The project started in July 2016 and ended in September 2017, enrolling a total of 260 women from Ahero sub-County Hospital and Bondo County Hospital in the Nyanza region of western Kenya. The study was approved by the institutional review boards at the University of Washington and the University of Nairobi and Kenyatta National Hospital. Before recruitment

began, small focus groups of mothers, partners, and providers were conducted. The second author, a family planning fellow at the University of Washington, developed 164 templates for maternal health and family planning messages.

5.4.1 Eligibility and Enrollment

Women were recruited to the study at two public maternal-child health clinics in the Nyanza region of western Kenya. Eligibility criteria for the study included: being at least 14 years old; being pregnant and within 12 weeks of estimated due date (EDD); having daily access to a mobile phone with a Safaricom SIM card; being able to read and respond to SMS in English, Swahili, or Luo; and having the intention of remaining in the study area for six months. The requirement for Safaricom SIM cards was due to the use of a shortcode – which in Kenya is a dedicated five digit phone number which can send and receive both SMS and voice calls. Setting up a shortcode is a time-consuming and bureaucratic process. For our study we used the third-party gateway, Africa’s Talking¹, which simplified setup and provided simple backend API’s for connecting to our shortocde. Through Africa’s Talking, a shortcode costs about \$150 USD per month and enables us to setup toll-free SMS so participants can send replies to the system at no cost while we pay about \$0.007 USD per outgoing SMS. Limiting study participation to Safaricom subscribers was not a significant barrier since Safaricom dominates mobile services with 65% of the market [139]. In our study only 12% of ineligible participants did not have a Safaricom SIM.

There were many reasons for exclusion from the study, but the primary reason was gestational age. Of 648 women screened for enrollment 337 (56.4%) were ineligible. The non-mutually exclusive reasons for ineligibility were: being more than 12 weeks from EDD 155 (46%) - these women could be enrolled at a future ANC visit; not having daily access to a mobile phone 95 (28%); receiving antenatal care at a different hospital 68 (18%); not having a Safaricom SIM 40 (12%); not knowing their phone number 15 (4%); unwilling to

¹<https://africastalking.com/>

receive SMS 4 (1%); and participating in another research study 31 (9%); Additionally, 51 eligible participants declined or desired to enroll on a different day.

During recruitment, women were randomized (1:1) into a control arm receiving standard care or an intervention arm receiving bi-directional SMS. All participants who reported having a partner were allowed to invite him to participate as well. Partner inclusion was not randomized, for two reasons: first, our study design needed to be generalizable to the larger population since any scale-up or deployment of the intervention would include both partnered and unpartnered women; second, we felt that it was important to understand how both partnered and unpartnered women used the system. For all partners invited into the study, a home visit by a male Kenyan staff member was attempted in order to enroll the partner and complete enrollment data collection.

5.4.2 Informed Consent

Female participants provided written informed consent at the time of enrollment. Male participants were consented at the time of the home visit. This was done via a home visit rather than a phone call because focus groups revealed that it was much more appropriate to talk with men about maternal health issues in person. Consent counseling was conducted, and forms provided in, either English, Swahili or Luo based on participant preference. The consent process made it clear that the study was about family planning, and that participants would be sent medical information via SMS and would have the option to respond to messages. If at any time they felt uncomfortable or did not wish to continue with the study they could SMS *STOP* to the study shortcode and would be removed from future SMS messaging.

5.4.3 Total Enrollment

Table 5.1 shows the breakdown of all mothers based on their partner's enrollment status. Of the 260 total participants, 198 (76.2%) were partnered, and of these, 159 (80.3%) invited their partner and 39 (19.7%) did not. This left 101 (43.9%) participants who either did not have a partner or elected not to invite their partner. Of the 159 partners invited, 101

	Not Invited	Invited	Enrolled	Phone Shared
n_i	49	29	43	9
Automated System Messages:				5176 (45.1%)
Mother	1893	1095	127	347
Partner	0	0	41	0
Both	0	0	1673	0
Participant Responses:				3575 (31.2%)
Mother	1345	643	1027	173
Partner	0	0	387	0
Study Staff Replies:				2718 (23.7%)
Mother	994	466	354	139
Partner	0	0	187	0
Both	0	0	578	0

Table 5.2: Distribution of 11469 SMSes sent or received in the study. All messages from a dyad sharing a SIM card or phone are counted as coming from the mother.

(43.9%) were enrolled in the study and were sent SMS. In nine cases (15%) this was because the male partner did not want to participate, however, more commonly it was because we were unable to contact the partner (20%) or the partner was living outside of the study area (37%) making a home visit for informed consent impossible. Five partners were recruited but were not successfully enrolled in the SMS system due to technical issues. and 58 mothers (25.2%) invited their partner but their partner was not enrolled. Since randomization occurred independently of partner enrollment status, the breakdown between each group and control or intervention arm was not completely equal. If a partner was enrolled, both members of the couple received the same automated SMS messages. Twenty-one enrolled dyads shared a mobile phone in which case there was only one number to send messages to (last column of Table 5.1).

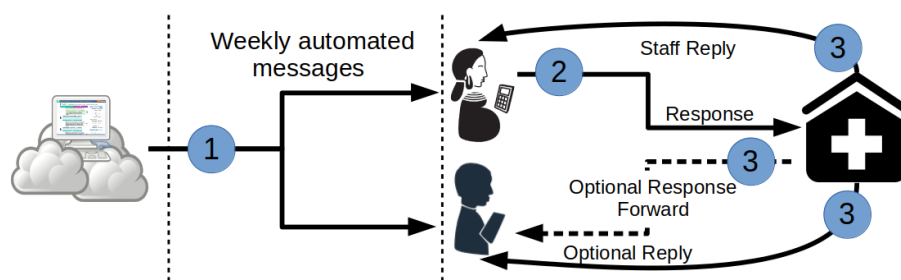


Figure 5.1: System design for three-way SMS conversation in enrolled partner dyads. (1) Automated messages from the system go to both partners. (2) Replies from either partner go only to the system. (3) Replies from study staff by default go only to the original sender, but message and response can optionally be forwarded to the other partner. This design allows study staff to have confidential conversations with either partner.

5.4.4 Data Sources

In this paper, three primary sources of data were used for analysis: 1) self-reported demographic and technology use data taken from the enrollment questionnaires; 2) the content and meta-data associated with the SMS messages stored in the system; and 3) notes from field visits by the first and second authors as well as Skype calls and email logs between the research team and Kenyan study staff.

During the study, a total of 11,469 SMS were sent or received (Table 5.2). The automated weekly messages from the system, which were sent from the week of enrollment to 6 months after delivery, accounted for 5176 of these messages (45.1%), participants sent 3575 (31.2%) messages into the system, and the study staff sent back 2718 (23.7%) replies. For all of these counts, duplicate outgoing messages sent to both the mother and father are counted as one SMS. All mothers in the intervention group ($n_i=130$) sent 3188 messages to the system. There were 43 participants in the intervention arm who had their partner enrolled and did not share a phone with him. These 43 women sent 1027 SMS messages to the system while their male partners sent a total of 387.

5.5 System Design

For mWACH XY the platform was adapted based on the version used in mWACH X. Two major changes needed to be made to this existing system. First, the automated messaging had to be modified to include family planning messages and messaging needed to change based on any family planning methods initiated. Second, the system UI and back-end messaging system needed to handle sending and receiving messages from both the mother and the male partner.

5.5.1 Semi-Automated Bidirectional Messages

There were 164 messages in the SMS bank, each of which was translated into English, Swahili, and Luo (the dominant languages of the Nyanza region). Messages were adapted from previously created SMS banks to focus on family planning concerns, misconceptions, and needs. Six focus groups with women and men from the community surrounding the study clinics found that both women and men felt that SMS conversations with a trained nurse would help them talk as a couple about family planning.² Informed by feedback from these focus groups, previous SMS banks and the Theory of Planned Behavior [74, 7], specific family planning messages were developed. These automated messages fell into three categories:

ANC Messages tailored to the expected due date (EDD) of each mother. These messages started 14 weeks before EDD and included a mix of maternal health information – such as emphasizing facility delivery or information on important danger signs – and general family planning messages. For example, the message sent four weeks before EDD mentioned multiple family planning methods:

{name}, this is {nurse} from {clinic}. There are many effective options for family planning after you deliver, including condoms, pills, injection, IUCD (coil), and implant. It is your choice – you can ask me or a nurse at the clinic about your options. Is your partner supportive of family planning?

²A manuscript describing findings from the focus groups is in preparation.

All messages were modified for participants with enrolled partners to address the couple. For the above message ended with “*Have you talked about family planning as a couple?*” for the partner enrolled group.

General Postpartum Messages based on the actual delivery date and addressing general maternal health issues. The first few weeks after delivery these messages asked about any challenges the parents may be having. By the fourth week, messages began emphasizing family planning methods. Each week a different family planning method would be featured with information about how it works and how often it would require clinic visits. Messages in this phase also addressed specific misconceptions about different methods such as this message sent nine weeks after delivery:

{name}, this is {nurse} from {clinic}. There are some myths in the community about family planning. The injection does not cause infertility, so it is safe to get the injection if you are planning more children in the future. What have you heard about the injection?

Family Planning Specific Messages based on the date a couple started a family planning method. These messages were developed for this study and required major changes to the system so it could handle sending messages based on both the delivery date and family planning initiation date. Six different family planning tracks were created based clinic offerings – Pills, Injection, Implant, IUCD, Condoms, and Tubal Ligation. Once study staff were notified via SMS that a participant began a particular method they needed to update the system with her chosen method. At that point, tailored messages would start, beginning with a congratulatory message. The first four weeks of each family planning track focused on concerns and challenges associated with that method. For example the second week after a mother received a contraceptive injection (a common family planning method that lasts up to 12 weeks) both partners in the enrolled track received this message:

{name}, this is {nurse} from {clinic}. Did you both know that irregular menstrual bleeding or no bleeding at all is normal and healthy with the injection?

	Partner not invited		Enrolled women			Partner enrolled		Enrolled men		Enrolled vs partners	All women vs partners
	n		n		n		n		(paired)*	(unpaired)**	
Age	101	21 (19-25)	58	24 (21-29)	101	23 (20-27)	101	30 (26-35)	<0.0001	<0.0001	
Education	101		58		101		101		0.01	0.03	
No secondary		29 (28.7)		19 (32.8)		40 (39.6)		37 (36.6)			
Some secondary		18 (17.8)		15 (25.9)		20 (19.8)		9 (8.9)			
Secondary complete		33 (32.7)		14 (24.1)		24 (23.8)		26 (25.7)			
Above secondary		21 (20.8)		10 (17.2)		17 (16.8)		29 (28.7)			
Income											
Household	43	70 (25-150)	36	100 (52-175)	56	100 (50-150)	98	80 (50 – 200)	0.85	0.91	
Individual	81	0 (0-40)	53	1 (0-50)	87	10 (0-50)	97	80 (40-150)	<0.0001	<0.0001	
ICT Access and Use											
Airtime per week (USD)	100	100 (50-250)	100	100 (50-150)	58	100 (50-200)	99	200 (100-350)	<0.0001	<0.0001	
SMS sent per week	95	21 (7-50)	94	5 (3-20)	57	5 (2-10)	90	10 (3-30)	0.20	0.30	
Used internet	101	40 (39.6)	101	34 (33.7)	58	23 (39.7)	101	68 (67.3)	<0.0001	<0.0001	
Used WhatsApp/Facebook	101	31 (30.7)	101	21 (20.8)	58	17 (29.3)	101	46 (45.5)	<0.0001	0.001	
Touchscreen	101	33 (32.7)	101	23 (22.8)	58	15 (25.9)	101	36 (35.6)	0.02	0.12	

Table 5.3: Demographic characteristics and technology access by gender. For continuous variables the median and interquartile range (IQR) are shown, for categorical the count and percentage are shown. * paired t-test for continuous, McNemar test for binomial, McNemar-Bowker test for multinomial. ** t-test for continuous, χ^2 for categorical

After the first four weeks the frequency of family planning specific messaging was reduced and participants received general maternal health message every other week. The automated messages for all participants continued until 24 weeks (six months) after delivery. This automated messaging schedule required the system to track delivery dates and family planning start dates for all participants. Study staff at each clinic kept system data updated by abstracting paper-based clinic records, talking with participants during study visits, and SMS inquiries for participants in the intervention arm.

5.5.2 Adding Partner Messaging

A significant component of system development was to enable male partners participation in the conversation. The process of modifying existing open-source SMS systems for three-person group messaging in a health context is one contribution of this work. Group messaging over SMS is common but it uses the Multimedia Messaging Service (MMS) extension which is not available on third party SMS gateways in Kenya. This meant that we had to simulate multi-party conversations by duplicating outgoing messages over a single shortcode on the

SMS gateway.

The initial system designs intended the platform to replicate SMS or WhatsApp group messaging as closely as possible. We explored the possibility of all messages from either partner automatically forwarding to the other partner. This is very similar to Tangaza [137] and other systems for efficient and cheap group messaging over basic phones. However, the personal nature of health conversations and interpersonal relationship dynamics of a maternal health messaging system require different assumptions than social or friend based group messaging. Prior work has shown that when engaging pregnant women in open-ended SMS conversations, it is not a question of if but rather when sensitive information will be shared over SMS [150]. If women knew every message they sent would automatically be sent to their partner it could deter some questions and constrain the conversation. Also, instantaneously forwarding messages between partners could open the system up for abuse where a couple could start using the toll-free shortcode for free SMSes between them. Although observing this hypothetical emergent behavior would have been interesting, it is not within scope of the study or the intended use of the system.

Our final design is best described as semi-automated group messaging. Figure 5.1 illustrates the messaging flow for an enrolled couple when both partners have separate SIM cards. All automated system messages were sent to both partners at exactly the same time each week. From the participant's perspective, the only indication that both members received the SMS was the inclusive wording of messages. Replies from either partner went only to the system and were displayed on the web interface in a single conversation thread – much like group messaging on a smartphone. Study staff would then review, translate, and if necessary reply to the message – deciding if the reply should go to only the original sender or both partners. It was important to have the default be the lowest risk option and so by default the reply would only go to the partner who sent the message being responded to. However, if appropriate and if it would facilitate the conversation, the incoming message and outgoing reply could be forwarded to the other partner. The web interface for the messaging system was designed to seamlessly include the male partners. Incoming messages were color-coded by sender and

icons next to the message clearly indicated if it had come from the mother or partner. When study staff replied, toggle buttons indicated if an outgoing message would be sent to the mother, partner, or both. The default action was to reply just to the sender of an original message but with minimal extra work messages could be forwarded to both partners. This feature was used quite often. Table 5.2 shows that study staff sent a total of 1119 messages to participants with an enrolled partner. Of these messages 578 (51.7%) were sent to both partners simultaneously while 354 (31.6%) were sent to the mother only and 187 (16.7%) were sent to the partner only.

5.6 Participant characteristics

In this section we use responses from the extensive enrollment questionnaire to better understand the demographic characteristics of participants as well as their prior experiences with technology. The inclusion of 101 male partners offers an opportunity to compare baseline technology use between women and men from the same community as well as within the same household. This analysis helps to better understand the context of the intervention and to inform the selection of technology for future projects. Previous work has indicated that mobile phones and other technology amplify already existing gender differences [206]. It is, therefore, important to understand who the participants are and their existing technology use so that we can design mHealth solutions for greater equity and inclusion.

5.6.1 Basic Demographics

The basic demographic data in Table 5.3 reveals that within couple dyads and the population as a whole, women and men had significant differences in terms of sociodemographic and technology use characteristics.

First, the women were much younger than the men. The median age of women with a partner enrolled in the study was 23, and the median age of enrolled men is 30. The educational background of the women and men in the study also differed. While the percentage of women and men who reported having *no secondary* education (33.8% and 36.6%) and

having *completed secondary* school (27.3% and 25.2%) were both similar, only 8.7% of the men reported having *some secondary* while 20.4% of the women did. This higher secondary school dropout rate among women is a trend seen throughout Sub-Saharan Africa [172]. Median household income was approximately 100 USD per month with no major differences reported between the men and women. The median individual monthly income was 0 USD for the women and 80 USD for the men. With a mean rural Kenyan household size of 4.4 [140], this places our study population in the 1.3 billion people globally living on less than \$1.25(USD) per day.

The enrollment questionnaire assessed baseline technology use. We wanted to gain a sense of what type of phone the study participants had, and how they used it. The bottom half of Table 5.3 summarizes the technology access and use data. We see that the men were much more likely to report having used the Internet as well as having used WhatsApp or Facebook. As a proxy for determining if participants had a smartphone, we asked if their phone had a touch screen or not and 27.3% of women said their phone had a touch screen compared to 35.9% of men. These numbers are slightly higher than other work exploring women's technology use in Kenya [205], which may indicate how quickly cheap touchscreen phones are being adopted. The last two columns of Table 5.3 display the outcome of paired statistical tests comparing the mother and partner in each dyad as well as unpaired tests on the larger population of all women to all enrolled men. Interestingly, while men reported spending about twice as much per week on airtime, both men and women said they sent about the same number of SMS messages per week. The only two variables that are not significantly different between women and men are self-reported household income and number of SMS messages sent per week. Taken together these demographic data indicate that data-based services still haven't reached the pockets of enough rural Kenyan women to be universally available and SMS remains the most ubiquitous messaging platform for our study population.

5.6.2 Phone Features and Use

We asked participants what features their phones had and which, if any, they used. In Table 5.4 we show the percent of mothers whose phones had each feature. It is important to note that numbers reported for features are a lower bound since participants might not be aware of the feature. For all features a paired t-test shows significant difference between the mothers and their partners. The only feature fewer men report having is a torch (flashlight) on their phone – a surprising finding since this is a feature built into even the most basic phones. This may be partially explained by the fact that more men reported having touchscreen phones and the use of the camera flash as a torch is not as intuitive or discoverable as the torch on more basic phones.

Having a phone with certain capabilities does not mean they are actually used. In our data, the significant gender differences in features that participant phones had was amplified when taking in to account the features that are actually used. In Figure 5.2 the percent of participants who used (green/bottom), had but did not use (yellow/middle), and did not have (blue/top) is shown. The only truly universal feature is SMS, with 100% of both women and men saying their phone has SMS and around 90% of all groups saying they use SMS.

n_a	Enrolled women			Partner
	Not invited n = 101	Invited n = 58	Enrolled n = 101	n=101
Torch	92.1	98.3	90.1*	79.2
Radio	80.2	89.7	83.2**	95.1
Play Music	70.3	74.1	72.3*	84.2
Take Pictures	59.4	46.6	60.4*	73.3
Show Videos	42.6	20.7	37.6**	62.4
Multiple SIMs	36.6	39.7	34.7**	51.5

Table 5.4: Percentage of participants self-reporting phone feature use at enrollment based on partner enrollment status. For enrolled participants significant p-values for paired t-test comparing enrolled mothers and partners are indicated at **p<0 .01, *p<0.05.

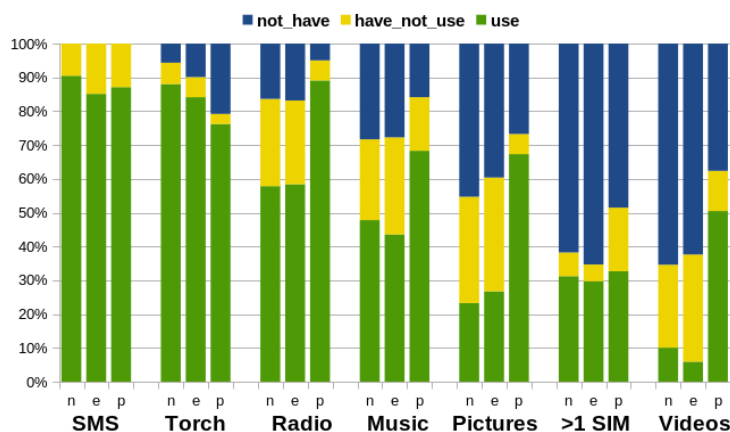


Figure 5.2: Phone feature use among participants. Percent of participants who used (bottom/green), had but did not use (middle/yellow), and did not have (top/blue) each phone feature. Participant categories: (n) mothers with partner not enrolled, (e) mothers with enrolled partners, and (p) partners.

Large gender divides in both access and use existed for the other features. For example around 90% of participant phones support FM radio; however, only about 60% of women reported using radio on their phones while almost 90% of the men did. The large fraction of phones capable of playing FM radio demonstrates an important non-data enabled multimedia function mobile phones served for our study population. Large gender differences in actual use also existed for playing music, taking pictures, and watching videos. Lastly, the number of participants who reported using multiple SIM cards, for both women and men, was greater than the number of participants who said they had a dual SIM phone. This suggests that many people were switching SIM cards to get the best deals on phone calls. It should be expected that at least a quarter of the participants will have multiple phone numbers and the system should account for this.

5.7 Measuring Male Engagement

Two primary research questions were (1) how men’s use and involvement in family planning conversations would be similar to or different from mothers’; and (2) if inclusion of male

partners affected the participation of mothers in any way. In this section we attempt to answer these two questions. First, we analyze statistics of system use to see how responsive partners were and if their inclusion changed the overall participation of the mothers. Second, we examine the content of the messages and highlight interesting and novel interactions with the system. The major takeaway from this analysis is that, as expected, male partners engage less with the system than women. However, when men do engage, it is in almost exactly the same manner as the women.

5.7.1 *Quantifying Engagement*

Of the 52 men enrolled in the intervention group, nine shared a phone with their partner, and so did not receive individual messages. The remaining 43 male partners sent a total of 387 messages with a median of 7 , IQR(3 - 13). This is drastically less than 130 women in the two-way group who sent a median of 20 IQR(10 - 38) messages.

The cumulative frequency plot in Figure 5.3 shows the number of messages each percentile of male partners and mothers sent. A larger proportion of male partners sent fewer messages, with 55% sending less than five messages. The maximum number of messages sent by a male partner was 43 and the graph shows that 10% of men sent more than 20 messages while 60% of women sent more than 20 messages. The number of messages sent by mothers with a partner in the study and mothers without a partner in the study did not differ.

While it is clear that the men engaged less than the women, an important follow-up question is: which men engage? Are the men who engage the most the partners of mothers who send the most SMS messages? In only eight instances does the male partner send more messages than his female partner. We found that there was no correlation between the number of messages sent by the female and male partner in each couple ($r=0.14$, $p=0.38$).

It was also possible that including male partner would decrease the mother's engagement. To understand this we compared the total number of messages sent per mother based on the partner's enrollment status (not invited, invited, and enrolled) Table 5.5 shows the proportion of each group that responded at least once, as well as the median number of SMS sent for

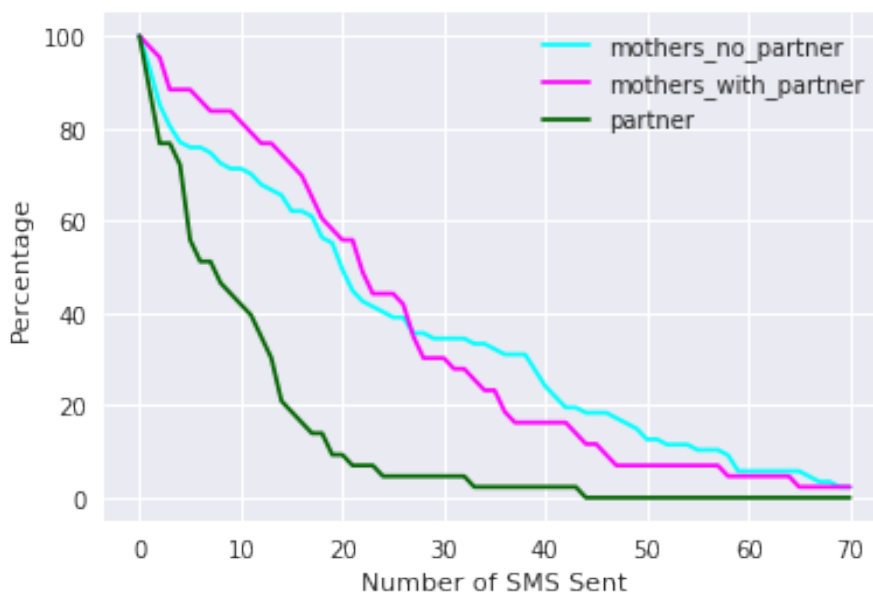


Figure 5.3: Cumulative frequency of SMS messages sent by mothers (grouped by partner status) and the enrolled men. Only 10% of men send more than 20 messages while 50% of women do.

the whole study period. A majority of both men and women engaged at least once with the system. Of all mothers, 123 (94.6%) sent at least one message into the system and 38 (88.4%) men sent at least one SMS (second column of Table 5.5). More importantly, if we look at each couple as a whole, then 100% of dyads sent at least one message to the system. Thus, including the male partner engaged more households than would otherwise be included in the conversation. The unequal variance t-test between the number of messages sent by mothers in the enrolled group and others was calculated. There was no significant difference between mothers with an enrolled partner and the other groups of mothers. Including the male partner in the conversation, therefore, does not change the frequency with which any mothers engage. However, comparing mothers with an enrolled partner and their partners, there is a highly significant difference in engagement, with men sending 37.7% as many messages as the women.

Another way to analyze engagement is from the time interval between when the system

	n_i	Active (n / %)	Total SMS median (IQR)	P
All Mothers	130	123 (94.6)	20 (9.25 - 38)	0.84
Not Invited	49	47 (95.9)	23 (14 - 40)	0.38
Invited	38	34 (89.5)	15.5 (1.3 - 37.3)	0.58
Enrolled	43	42 (97.7)	21 (13.5 - 32.5)	ref
Partners	43	38 (88.4)	7 (3 - 13)	<0.0001
Couples	43	43 (100)	33 (18 - 41)	ignore

Table 5.5: SMS engagement by individuals and couple dyads. Active (column 2) is a count of participants who sent at least one message at anytime during the study. In the Total SMS column the median and inter-quartile range are given.

or study staff sent an outgoing message and the system receiving an incoming message. Of the 5176 outgoing system messages, 2040 (39.4%) were replied to within seven days by either the mother or partner. Figure 5.4 plots, on a log scale for time, the cumulative distribution function (CDF) of response times for all messages from the mothers and partners. Unexpectedly, the curves for both women and men are exactly the same. Within one hour of sending messages out about 60% of replies were received. The fact that all of these curves are identical means that 1) even though the women have less ICT access they can be thought of as being *online* with regards to SMS and 2) even though the men do not engage with the system as much as women, when they do, it is within the same timeframe as the women.

5.7.2 Participant Conversations

As we have shown, even though the male partners participated less in family planning SMS conversations, when they did participate it was in much the same way as the women. The next question asks with what content did the men engage? Throughout the study, every incoming message was processed by study staff and categorized by topic. The topic of a message cannot be determined from the SMS content alone, but requires knowing the context of the messages that preceded it. For example in the following message exchange:

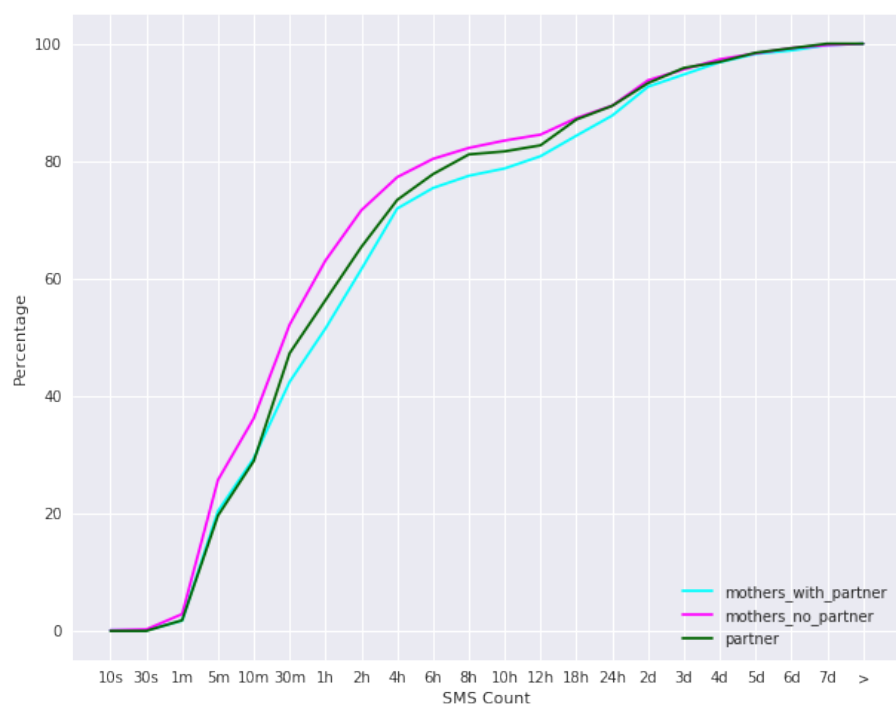


Figure 5.4: Response times between an outgoing message and an incoming message on a logarithmic time scale. Results are broken down by partners, mothers with enrolled partner, and mothers without an enrolled partner.

System Both 06-01 8:01 – *Spacing pregnancies by 2 years promotes health for mothers and babies. Have you talked about when to have another child as a couple?*

Partner 06-01 8:24 – *yes*

Mother 06-01 12:14 – *Yes, not less than 5 years*

The single answer reply from the partner would be classified as family planning since it engages with a family planning related system message. Figure 5.5 shows the topics of messages sent by partners and mothers. It is apparent that the men and women engage in very similar topics and at basically the same frequency. About 70% of messages from both mothers and partners are evenly split between family planning and maternal health. While not surprising (more than half of the messages we set out were about family planning), this was by no means guaranteed. This shows that the primary use of the system was centered around the two intended topics. The fact that all participants engaged in the primary topic the system was intended for is encouraging. Messages in the ‘other’ category tended to be about the study or unrelated health concerns.

What exactly did the participants talk about? Were the only responses to the automated questions short Yes/No answers, or did the men engage in substantive family planning discussions and address real concerns? One, admittedly rough, estimation of engagement is message length. We found that over 77.3% of messages were ten characters or longer, which indicates more engagement with the system than Yes/No replies. For the mothers 76.7% of their messages were longer than ten characters, whereas for the male partners, 81.6% of messages are longer than ten characters. We now present a few examples of longer messages and interesting SMS conversations involving male participation in the study. This first example demonstrates how the system enabled a multi-party conversation between the system, study nurse, and both partners. This conversation took place in Luo and was translated by study staff into English. It starts with a generic infant care question three weeks after delivery:

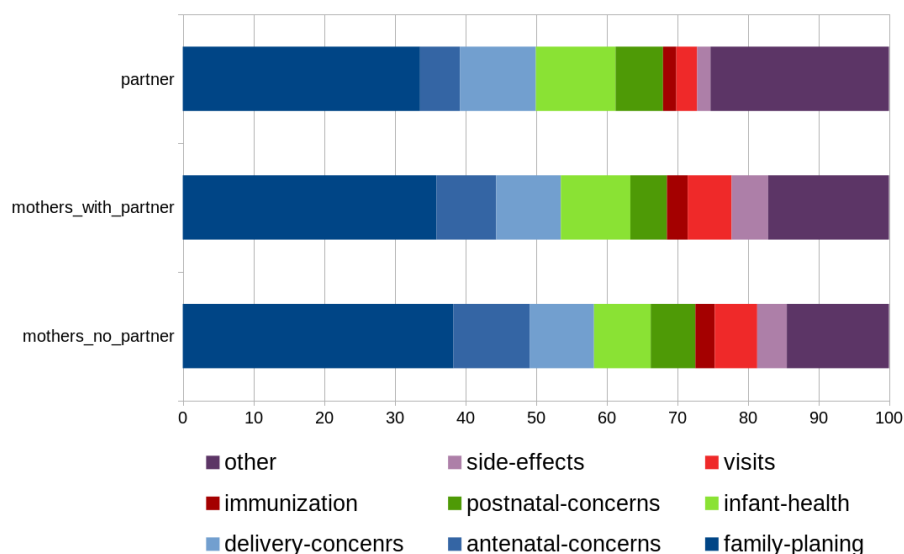


Figure 5.5: The topic of incoming messages by group. All mothers and partners engage with the SMS system around the same topics and at the same frequency.

System Both 11-07 14:01 – *Please ask us if you need any information. We are here for you and your baby. Are you having any challenges?*

Later that evening the partner replies:

Partner 11-07 18:23 – *Mama {name} and I are doing fine, although for the baby he has difficulty when passing stool, he can pass stool once in a week, so I don't know where the problem is!*

And then the next morning the mother asks two family planning related questions.

Mother 11-08 10:45 – *if i use the coil are there any problems?*

Mother 11-08 10:55 – *if you want to insert the coil is it a must you come with your husband*

It is interesting that the mother brings up contraceptive use on her own even though the most recent system message did not talk about family planning. Participants often did not reply directly to the content from the automated system and clearly saw the SMS exchange as a general maternal health resource.

The next day the study staff arrive at the hospital and begin to respond to messages. The important question from the partner is forwarded to the mother so she is also aware of the

conversation. Next the study staff answer both questions.

Nurse 11-08 11:34 Mother – *Hi {name}, if you want to be inserted for the coil it is not a must you come with your husband, but if he has time he can accompany you. you will experience cramps during menses and during the first few days.*

Nurse 11-08 11:37 Both – *Hi {name}, have you started giving the baby other food and fluids apart from breast milk?*

The first answer about not needing to come with your husband to get an IUCD is sent just to the mother. The second question asking about other foods is a response to the partner's question and sent to both mother and partner. Within 20 minutes the mother replies to the nurses question about having other foods.

Mother 11-08 11:52 – *until 6 months is when he will start taking other fluids*

This example illustrates how complex the multi-conversation messaging can get. Including the forwarded message, the mother received three messages from the system in less than five minutes and they were out of order for a natural conversation. However, this did not stop her from understanding and replying.

This next example demonstrates why it is important for the system to not automatically forward messages between the partners. In replying to an automated message about family planning, one mother asked if it was possible to do family planning without your husband knowing.

Mother – *can you do family planning without your husband knowing?*

Nurse – *yes you can, if he does not approve of family planning and you really want to use it.*

Mother – *i will soon come alone*

The study nurse also used the ability to have a conversation with just one partner to talk with the men about sensitive issues. For example in the following exchange a partner asked about family planning methods that are applicable to men only. This exchange took place in English and with the exception of names is represented exactly.

Partner – *Hi can you tell me a method of family planning that is applicable to men only*

Nurse – *hello {name},for now men can use condoms and vasectomy, although this is an irreversible method.it can be used when you feel you have had enough children and you need a permanent family planning method.*

They then go on to exchange three more messages about the details and side effects of a vasectomy.

When analyzing the messages sent with the male partners' phones, we observed 12 messages that were clearly sent by the mother such as this one:

Partner – *i feel like my tummy is heavy and the baby is not moving,blood builder medicines that i was prescribed for i have not yet bought,that is where i need help.*

For many of the messages it is impossible to know who is actually sending the message, but the use of first person to describe the pregnancy clearly indicates that this message from the partner's phone was written by the mother. This sharing of the phone may also have happened in the other direction as well (i.e., the partner was responding on the mother's phone number) but coding all 3188 messages individually was beyond the scope of this analysis so it is unclear how commonly this occurred.

5.8 Discussion

A central paradox of all ICTD behavior change work is that the people that most need outreach are also those with the least access to ICTs. It is important to choose the correct technology and medium based on the target population and demographics. Like most of Sub-Saharan Africa, the women in our study had limited access to data-enabled services and smartphones. It might be tempting to create a WhatsApp or Facebook messaging application since both channels offer the developer more features. However, only 27% of women in our study said they had used either service. Slightly more men (45%) reported using these data-enabled services – but that still means we would be reaching less than 50% of the population. Our data show that every single phone is capable of sending and receiving SMS

messages and, more importantly, the weekly use of SMS between genders was basically the same prior to the study. By choosing to use toll-free SMS as the communication channel for this intervention we were able to reach and engage the mothers. Any M4D project that plans on reaching end users needs to have an SMS or IVR fall back even if it uses data enabled channels for some users. This makes the intervention scalable across the technology requirements and needs of the whole population.

The use of toll-free SMS messaging means that the cost of messaging must be paid by the provider – in this case the research team, but if a similar service were to be scaled up an NGO or the Ministry of Health would need cover the messaging cost. Previous work with IVR has shown that as soon as an intervention stops using a toll-free number usage significantly drops off [197]. No one has replicated this study in SMS, but requiring participants in our study to pay for SMS messages would likely reduce engagement. However, with SMS, the toll-free costs are nowhere near as large as in an IVR system. The monthly \$150 USD cost for the short code was fixed regardless of how many messages we sent and at \$0.007 USD per message \$100 USD would send over fourteen thousand messages. With SMS the communication cost is not a prohibitive factor to scale.

One aspect of the system that is difficult to scale is the need to the synchronize internal variables that drive messaging with each participant’s current status. This is a major challenge in the design and implementation of automated messaging platforms in M4D. For a small RCT such as ours, this was manageable with a dedicated staff – but if a maternal health platform scales to an entire district or nation, then simple and easy data collection methods need to be deployed. For example, MomConnect in South Africa has community health workers register pregnant women over basic USSD menu systems [176]; however, there has been little work analyzing the accuracy of this approach and how well it captures future events such as deliveries.

Since male engagement was not randomized it will be impossible to know for certain if there are any causal effects of male partner inclusion in the family planning conversation. However, our findings indicate that at the very least involving male partners does not negatively impact

the conversation held by women and actually engages more households. The inclusion of men in family planning messaging raises several interesting questions. First, how do we know if the limited engagement we did see from the men is typical, or should we expect more engagement? Male partners in Kenya typically do not attend prenatal clinic sessions. With few points of contact with the health system it becomes more difficult to enroll male partners in the SMS platform. In our study we conducted home visits to reach the partners but if this were to scale up to the district or national level home visits quickly become impractical. A similar question for the women is how does a health system engage with them before they come in pregnant? Family planning knowledge and agency might be more useful *before* a pregnancy than after. One solution to both of these problems is to have an opt-in SMS platform such as m4RH. This raises questions of equity and access, since those who learn about and contact the system are often less likely to need the information.

Maternal health outreach and engagement is a common goal of mobile health (mHealth) projects for development. Male partners of pregnant and postpartum women are known to be important influences on their health behavior. This chapter presented a novel extension to previous HCI4D research by exploring how to engage male partners in SMS-based family planning conversations. First, we explore design considerations for inclusion of male partners in an existing semi-automated bidirectional SMS platform. A 12-month randomized controlled trial of a family planning counseling SMS program was conducted in western Kenya using our system. A total of 260 pregnant women and 101 of their male partners were enrolled in the system. We analyze enrollment and usage data from this trial to compare baseline technology use and demographics of mothers and their male partners. Our findings demonstrate significant technology gender divides in the study population. Finally, we explore how both the mothers and their male partners interacted with the SMS system through an analysis of over 11,500 messages. We conclude that it is feasible to include rural Kenyan men in an SMS-based family planning discussion program and that their inclusion does not dramatically affect the mothers' engagement.

Chapter 6

LONGITUDINAL ANALYSIS OF ENGAGEMENT IN MWACH X

The ICTD and Global Health space suffers from an acute case of pilotitis [73] where projects are continuously started but never transitioned to scale. Even the most rigorous of these short-lived projects will only report summary statistics of system utilization and engagement. For example Figure 3 in the original WelTel paper [113] shows the proportion of weekly SMS responses that are *Sawa*, *Shida* or *no response*.

While useful for understanding overall system usage, such graphs do little to inform the design and implementation of future projects. For example are the sample participants always failing to respond? If so this would indicate that maybe the service should be changed for that subset of the population. Or how many people stop receiving both SMS messages and phone calls – indicating their SIM card has changed.

In this chapter I present an in-depth analysis of participant engagement and system performance during the mWACH X study. This was our longest study with over two years of postpartum messaging enabling a look at messaging trends over time. The goal of this analysis is to understand the complexity of engagement with the system and use of mobile phones by the study population. The data includes all messaging activity from the start of the study in November 2015 through the end of May 2019. With over two years of messaging for participants, this longitudinal analysis offers important insights into the sustainability, utilization, and effectiveness of SMS projects.

6.0.1 Data Used For Analysis

There are two primary data sources for this analysis. First, baseline demographic data were obtained during an extensive enrollment questionnaire. This enrollment form had over 900 questions, although the use of skip-logic meant that many were only asked if certain criteria

where met. The technology subsection of the enrollment form included 38 questions on current and past technology use which forms the core of this analysis. An enrollment form exists for all 824 participants in the study and there were only minor differences between the three study arms at base line (in employment and previous number of pregnancies).

Second, the mWACH platform provides a rich source of system level meta-data. The system records the delivery status for each message sent. Starting in November 2016 these delivery reports were collected in real-time so the time to deliver a message could be calculated. Delivery reports only record if a message successfully reached the intended mobile phone, not if the message has been read. This is equivalent to the two gray checks on WhatsApp – there is no equivalent of the blue checks showing the read status. The system can also report the response time and response rates for each participant. Aggregating these statistics over the entire study for each participant reveals longitudinal trends that shorter pilot studies fail to explore.

6.0.2 Measuring Participant Engagement

In order to understand who does and does not utilize a platform like mWACH, we must first define a measure of engagement. The naive approach to this would be to simply use the response rate, like so:

$$response_rate = \frac{messages_incoming}{messages_outgoing}$$

However, should all outgoing messages be used or just the system messages (i.e. skip messages from the study nurse)? Similarly, should all incoming messages be counted or just those sent within two days of a system message? Should we also remove the visit messages which are sent outside of the usual message day? Additionally, as defined, the response rate can be over 100% if a participant sends in more messages than they receive.

A better measure of engagement with the system is the percent of weeks active. This is the number of weeks a participant has sent at least one message divided by the total number

of weeks in the study:

$$percent_active = \frac{weeks_active}{weeks_in_study}$$

This measure of engagement represents how frequently an individual interacts with the system. However, now someone who sends a single message to the system on three weeks and someone who sends five messages to the system on three weeks has the same level of engagement. As we will see, the vast majority of interactions with the system involve just one participant message (though a larger fraction of system messages result in no interaction at all). Still, a measure of interaction intensity is useful. For this we will use messages per week:

$$messages_per_week = \frac{messages_incoming}{weeks_in_study}$$

The last measure of system engagement is the percent of successfully delivered messages.

$$percent_derived = \frac{successfully_messages}{total_system_messages}$$

I refer to this metric as system engagement since receiving SMS messages successfully requires either keeping the same SIM card or updating the system with a new phone number. However, message delivery is a passive form of engagement where as weeks active requires an action on the part of the participant. Messages from the study nurse are excluded because they are only sent in reply to incoming messages and as is shown later in this chapter have a high success rate.

6.1 Baseline Demographics

Before analyzing how participants used the mWACH platform in this study it is important to understand who the participants are. A total of 824 pregnant women were enrolled in the system randomized 1:1:1 into the control, one-way, and two-way arms. The analysis in this section includes participants from all three arms since no interaction with the system is required. Using the total participant pool rather than just the 276 in the two-way arm makes this analysis more representative. This subsection breaks the baseline demographic data down by region within Kenya and by whether or not the participant has a touchscreen

phone (used as a proxy for a 'smartphone').

The mX project recruited participants from four sites in Western Kenya (Ahero, Bondo, Rachuonyo, Siaya) and two sites in outskirts of Nairobi (Mathare and Riruta). Figure 6.1 shows the location of all sites within Kenya. In general the peri-urban location of the Nairobi sites means there is more access to technology and health services. Initially, mX started at only three sites: Ahero and Bondo in Western Kenya, and Mathera near Nairobi. Due to slow enrollment the other three sites were added after one year. The original three sites account for 68.3% of participants and are the ones I am most familiar with.

The complete breakdown of total participants and two-way participants per site is shown in Table 6.1. Basic summary statistics of incoming and outgoing SMS counts for the two-way group are also shown. A total of 39,993 outgoing system messages were sent with a median of 155 (IQR = [135 - 165]). The system received roughly a quarter as many incoming messages from participants and there was much more individual variance in participant messaging. In total 10,122 incoming messages were received with a median of 23 (IQR = [4,53]). The two Nairobi sites have a total response rate of 35.6% which is higher than the Western sites response rate of 25.3% ($p=0.0004$ ind ttest).

Interestingly, study nurses reply to a higher percentage of participant messages in the Eastern sites (70.1%) than at Nairobi Sites (55.3%). There are many hypotheses for this difference in nurse replies from the two areas. First, it could be that since the Nairobi sites have a higher participant response rate, more incoming messages there are trivial replies not requiring a response such as "*Thanks*" or "*I am doing well*". Alternatively this difference could just be due to variations in how individual nurses choose to reply or not.

The enrollment interview questions also showed some regional differences in basic demographic data and technology access. Table 6.2 breaks down the age, educational level, and monthly income. Overall, the median age at enrollment was 27 (IQR 23–31) and the median number of years spent in school was 9 (IQR 8–12). The proportion of participants who answered questions about income was much smaller than for age and education. Of the 527 (64.0%) participants who gave their total household monthly income the median was

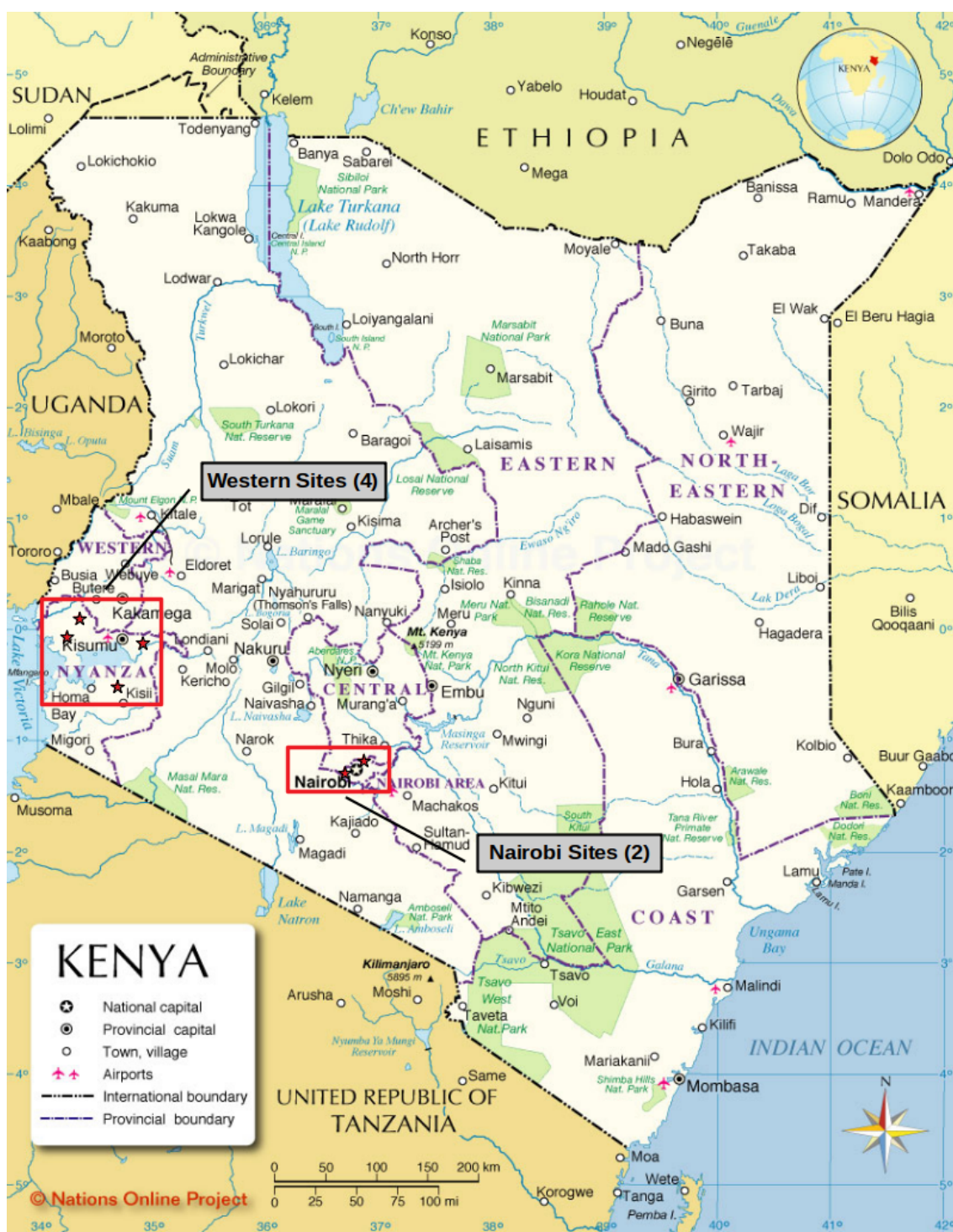


Figure 6.1: Map of Kenya with the six study sites started. Two sites are on the outskirts of Nairobi and the other four in Western Kenya near Kisumu. Map from Nations Online Project

	Location	n	Two-Way	System	Participant	% Response	Nurse	% Response
Ahero	Western	164	55	8530	1399	16.40	1138	81.34
Bondo	Western	171	57	8666	2229	25.72	1398	62.72
Rachuonyo	Western	120	39	5931	912	15.38	704	77.19
Siaya	Western	63	21	2894	614	21.22	372	60.59
Mathare	Nairobi	228	76	10184	3566	35.02	2198	61.64
Riruta	Nairobi	78	28	3788	1402	37.01	551	39.30
Total	—	824	276	39993	10122	25.31	6361	62.84

Table 6.1: Breakdown of enrollment (n) and total messaging numbers by study site. The total number of messages System, Participant and Nurse messages are only for Two-Way participants at each site. Response rates are simple fractions based on messaging totals.

\$80 (USD) (IQR 40–150). More participants – 636 (77.2%) – gave their individual monthly income and this with a median of \$20 (IQR 0–50) was significantly lower than household monthly income.

All demographic data in Table 6.2 is further broken down into sub-groups based on the region (Western or Nairobi) and whether or not the participant had a touchscreen phone at enrollment. Touchscreen ownership acts as a proxy for smartphone ownership – since asking if your phone has a touchscreen has an unambiguous meaning and maps almost perfectly to smartphone ownership. It was chosen as a metric to split the participants because having a smartphone dramatically changes what your phone can do. For each statistic and variable a t-test was conducted to determine if there were significant differences between the groups. Participants in the Nairobi do not differ from the Western area in age, but do have slightly more years of education and a higher monthly household income with median of \$100 USD vs \$60 USD ($p = 0.001$). However, the cost of living is most likely higher in Nairobi as well so what this difference means in standard of living is unknown. Splitting the demographic data based on touchscreen ownership also shows a significant difference based on years of education and household income but not age.

Besides basic demographic data, the enrollment questionnaire also included baseline technology use which is summarized in Table 6.3. Overall the median number of SMS messages sent and received each week is 5 while the median amount spent on airtime each

	n	All	n	Nairobi	n	Western	p	n	Touchscreen	n	No Touchscreen	p
Age (years)	824	27 (23-31)	306	27 (23-31)	518	27 (23-31)	0.51	207	27 (23-30)	617	27 (23-31)	0.95
Education (years)	822	9 (8-12)	305	10 (8-12)	517	9 (8-12)	0.03	207	12 (10-13)	615	8 (8-11)	<0.00001
Income (USD) Household	527	80 (40-150)	199	100 (50-160)	328	60 (30-133)	0.001	130	120 (70-200)	397	60 (30-120)	<0.00001
Income (USD) Individual	636	20 (0-50)	272	15 (0-62.5)	364	20 (82-50)	0.15	165	40 (0-90)	471	15 (0-50)	<0.00001

Table 6.2: Basic demographic baseline data for all participants split based on location (Nairobi or Western sites) and having/not having a touchscreen. For each data point the p-value from a t-test is given.

week is \$0.50 (USD). The most surprising finding from our baseline data is that just 5.8% of participants report having used an IVR system. We knew that participants would be unfamiliar with the term *IVR* so the question tried to explain what an IVR system was. The English version of this question was: *"Have you ever used an interactive voice (IVR) system, such as using your phone to give information to banks or companies without having to speak to a person or responding to a survey on your phone by pushing buttons?"* The extraordinary low number reporting IVR use is surprising in the context of the many IVR based ICT4D projects [98, 197, 122] and large NGO projects such as Viamo's 3-2-1 service []. However, in Kenya where literacy rates are high, the language uses a Latin based script, and SMS is cheap most companies will use text based communication channels such as USSD or SMS.

Only 22.0% of participants report having used the internet and 25.1% have a touchscreen phone. Again each metric is split into groups based on region within Kenya and touchscreen ownership. Comparing the Nairobi sites to Western sites there is no observed difference in how much airtime per week participants use or the number of SMS messages per week they send. However, significantly more participants at the Nairobi sites report having used the internet (32.7% vs 15.6%), used Facebook or WhatsApp (35.0% vs 13.1%) and a touchscreen (43.5% vs 14.3%).

The right most columns of Table 6.3 split the participants technology baseline data by touchscreen ownership. For all metrics there is a significant difference between smartphone

	n	All	n	Nairobi	n	Western	p	n	Touchscreen	n	No Touchscreen	p
Airtime pwk (USD)	796	0.50 (0.50-1.00)	305	0.50 (0.50-1.00)	491	0.50 (0.50-1.00)	0.62	199	1.00 (0.50-2.00)	597	0.50 (0.50-1.00)	<0.00001
SMS Sent pwk	659	5 (2-10)	278	5 (2-15)	381	4 (2-10)	0.17	175	10 (3-20)	484	4 (2-10)	0.001
SMS Received pwk	653	5 (3-10)	279	5 (3-15)	374	5 (3-10)	0.67	177	10 (4-20)	476	5 (3-10)	0.01
Had Phone (yrs)	792	5 (3-7)	304	5 (3-7)	488	5 (3-7)	0.05	203	6 (4-9)	589	5 (3-7)	<0.00001
Used Internet (yrs)	824	181 (22.0%)	306	100 (32.7%)	518	81 (15.6%)	<0.00001	207	109 (52.7%)	617	72 (11.7%)	<0.00001
Used WhatsApp/FB	824	175 (21.2%)	306	107 (35.0%)	518	68 (13.1%)	<0.00001	207	126 (60.9%)	617	49 (7.9%)	<0.00001
Used IVR	824	48 (5.8%)	306	7 (2.3%)	518	41 (7.9%)	<0.00001	207	20 (9.7%)	617	28 (4.5%)	<0.00001
Touchscreen	824	207 (25.1%)	306	133 (43.5%)	518	74 (14.3%)	<0.00001	207	207 (100.0%)	617	0 (0.0%)	—

Table 6.3: Baseline technology use data for all participants split based on location (Nairobi or Western sites) and having/not having a touchscreen. For each data point the p-value from a t-test is given.

owners and feature phone owners. The median amount of airtime spent per week by smartphone owners is \$1.00 USD, which is double what non-smartphone owners spend. Similarity the median number of SMS sent per week by smartphone users is 10 vs 4 for non-smartphone owners. Internet use is 52.7% vs 11.7% in favor of smartphone users and Facebook or WhatsApp use is 60.9% vs 7.9%. Having a smartphone is a clear indicator of more technology use. The question is how this prior experience with SMS and the internet effects engagement with the system during the study.

One important finding from the baseline technology questions is on what sharing a phone means in this context. Previous research has shown that phone sharing is an important part of technology access in low resource settings [173]. For this reason we asked participants if they shared their phone and if so, who had it for most of the day. Figure 6.2 shows that the majority (70.3%) do not share a phone. The surprising finding is that half of those who do share a phone report that they have access to the phone all day and that very few shared phones are left at home but rather act as someones personal phone. The implications of this are that someone who shares a phone, but is the primary user would engage with the system in a similar manor to someone who doesn't share a phone.

6.1.1 Comparing Mobile Services

There are four main ways of creating a mobile based service: USSD/SIM App, SMS, IVR, and data/Internet. For these purposes I am grouping USSD and SIM App's together because

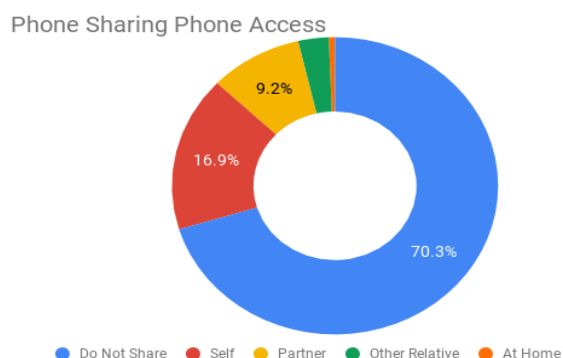


Figure 6.2: Phone sharing breakdown. At baseline 70.3% of participants do not share a phone. Of those that share their phone 16.9% say that they primarily have access to the phone all day.

from a user perspective they offer similar text based menu features, although SIM Apps can integrate more closely with the phone. In Kenya the most popular SIM App is the Safaricom mobile money system mPesa which has near universal adoption. mPesa is the only service that has the same adoption rates at the Nairobi sites and Western sites, as seen in Figure 6.3. At the Nairobi sites 87.9% of participants report sending at least one SMS a week while at the Western sites only 69.4% do. Internet use is similarly skewed with 32.7% of Nairobi participants having used it and only 15.6% of Western participants saying the same. IVR use is also significantly different but in the opposite direction with 2.2% of Nairobi participants saying they have used IVR and 7.9% of Western participants reporting they have used IVR. One explanation for this difference is that a common reason to use IVR is when someone has SIM card issues or is locked out of their mPesa account. In Nairobi this can be fixed by a visit to the Safaricom shop, but in Western Kenya a call is easier.

Figure 6.4 looks at these same four services by touchscreen ownership. Here the universality of mPesa, SMS, and IVR stands out as there is no significant difference in use between smartphone owners and feature phone owners. However, five times as many smartphone owners report using the internet (52.6% vs 11.7%). Taken together the low overall internet use and high SMS use indicates that SMS is an appropriate medium for reaching the target

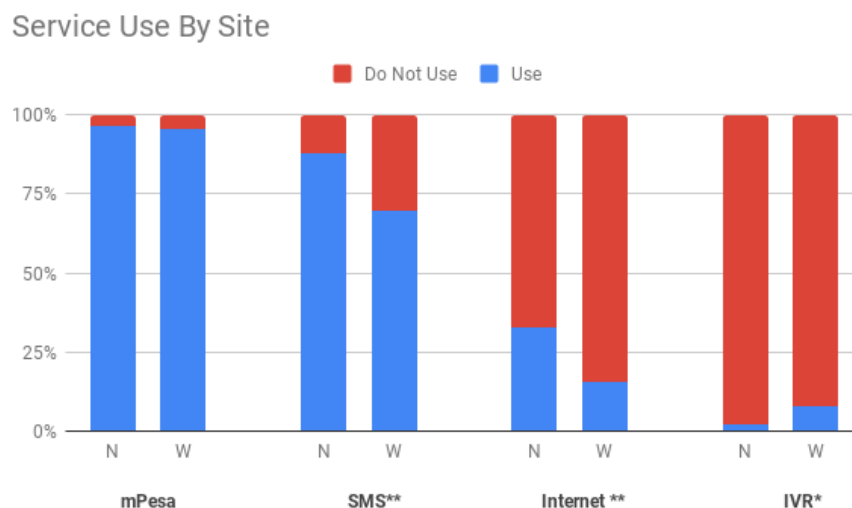


Figure 6.3: Comparison of baseline data for mPesa, SMS, Internet, and IVR use between participants at Nairobi and Western sites. (* $p < 0.01$ ** $p < 0.00001$)

audience of pregnant women in Kenya.

6.1.2 Comparing Mobile Phone Features

Looking beyond just phone services we also have data on the specific features participants phone have and whether or not they use them. Figure 6.5 and Figure 6.4 show the breakdown of these features by region and touchscreen ownership. The only feature that is not significantly different in either case is radio usage with around 70% of participants reporting using their phone for radio regardless of region or type of phone. In light of these findings and similar results from MXY in Chapter 5, there is a surprising lack of ICTD research published on radio use on mobile phones in sub-Saharan Africa.

The regional phone feature comparison in Figure 6.5 shows videos, games, and Facebook/WhatsApp are all used by less than 50% of participants with the Western sites having significantly lower usage. Around 50% of participants use their phones for photos and music again with the Western sites having significantly lower usage. Over 75% of participants use their phone as a flashlight with more participants at the Western sites reporting this use.

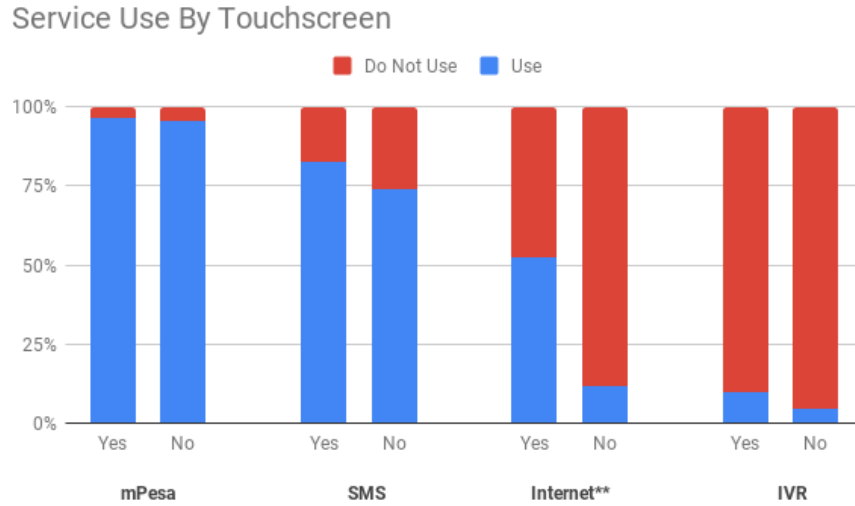


Figure 6.4: Comparison of baseline data for mPesa, SMS, Internet, and IVR use between participants with and without a touchscreen. (* $p < 0.01$ ** $p < 0.00001$)

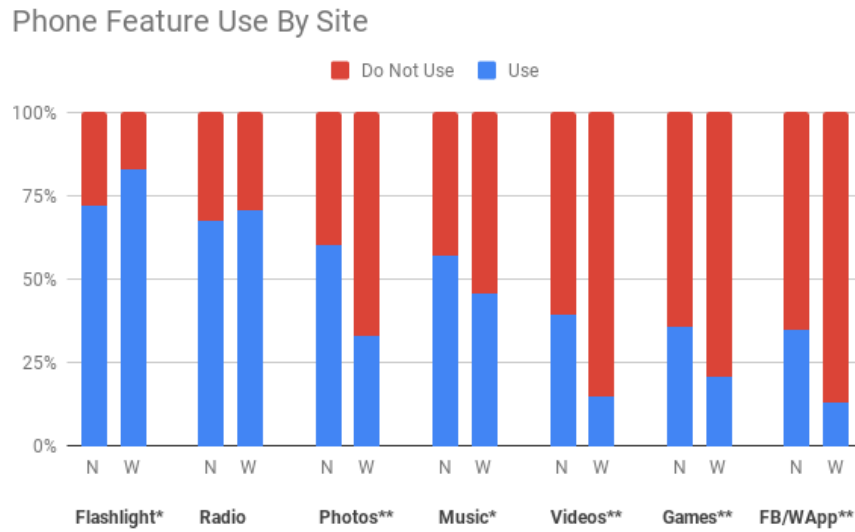


Figure 6.5: Comparison of baseline use of specific phone features between participants at Nairobi and Western sites. Non-use of a feature does not mean the phone does not have that ability. (* $p < 0.01$ ** $p < 0.00001$)

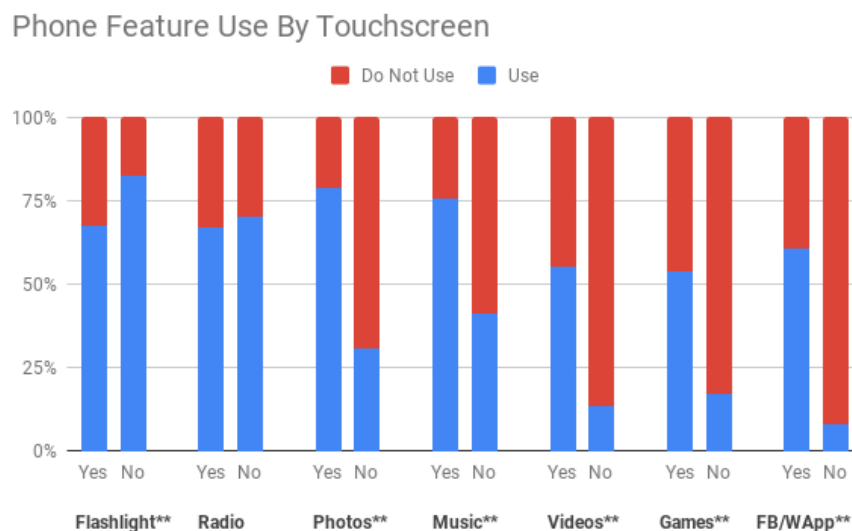


Figure 6.6: Comparison of baseline use of specific phone features between participants with and without a touchscreen. Non-use of a feature does not mean the phone does not have that ability. (* $p < 0.01$ ** $p < 0.00001$)

The touchscreen phone feature comparison in Figure 6.6 shows a large difference in feature use between touchscreen owners and non-touchscreen owners. For example, 78.7% of touchscreen owners say they use their phone for photos but only 31.0% of non-touchscreen owners say they do. This is a clear example of the increased functions that a smartphone provides around additional apps and data. It is possible to take basic photos, listen to music, or watch small videos on a feature phone – but all are much easier on a smartphone. Again, a major exception of this is SMS since the touchscreen service use breakdown showed little difference in SMS use.

6.2 System Statistics and Analysis

6.2.1 Messaging Overview

This section presents a longitudinal analysis of system performance and engagement for mWACH X. The two-year duration of messaging provides a unique opportunity to observe

how interactions change over time. All data for this section comes from information the system has recorded about each message sent or received from the start of the study through May 2019. In the previous section Table 6.1 showed overall message totals by site, however there is a large variation in how participants interact with the system. This variation can be seen in the CDFs in Figure 6.7 which show the distribution of total number of system, nurse, and participant message sent or received by each mother in the study. The automated system messages account for most of the messaging with 97% of participants receiving more than 100 messages. In contrast the median number of messages sent by participants was 23 and 26 participants (9.4%) sent zero messages. However, the top 10% of participants send more than 93 messages with the most active participant sending 236 messages.

The key insight that the CDF in Figure 6.7 shows is that stark difference between the bottom 10% of participants who sent almost no messages and the top 10% who send almost 100 messages each. The goal of the following analysis is to better understand these different usage patterns. Instead of using total number of messages, a more useful metric is *weeks_active* since participants can be in the study for a variable amount of time.

Using *weeks_active* treats someone who sends multiple messages a week the same as someone who sends one message a week. This is a reasonable approximation since most engagements with the system involve just one or two participant messages. Table 6.4 shows the number of system messages based on participant responses. Of the 39,993 system messages 32,794 (82.0%) receive no response and 5,546 (13.87%) receive just one response. Half of the remaining 4.13% of messages generate two participant messages with a long tail of longer conversations. From Table 6.4 it is clear that longer participant conversations are the exception, not the rule.

Besides *weeks_active* the second metric of engagement I will use is *percent_delivered* which is a measure of how many messages successfully reached the participant's phone. This data comes from SMS delivery reports which can be enabled through the Africa's Talking API. A delivery report means that the SMS has reached the mobile phone – not that the SMS has been read.

Participant	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	total
Count	32794	5546	1037	340	138	56	21	25	14	7	5	4	2	1	2	1	39993
Percent	82.00	13.87	2.59	0.85	0.35	0.14	0.05	0.06	0.04	0.02	0.01	0.01	0.01	0.00	0.01	0.00	1

Table 6.4: Breakdown of number of participant messages in response to each of the 39993 outgoing system messages sent. 82.0% of system messages have no participant response while 13.87% have a single participant response and 2.59% have two messages from a participant. There is a long tail of longer conversations.

Figure 6.8 shows a CDF of both metrics of engagement. The blue line is *percent_delivered*. A flat line at 1.0 would represent 100% of messages being delivered. Overall 73.4% of messages were successfully delivered, but there was a large variance in individual success rates. From the curve we can see that 20% of participants received less than 50% of their messages while 50% of participants receive at least 83% of messages.

The lower red line in Figure 6.8 shows the CDF of percent weeks active. No participant is active for 100% of weeks they were in the study, but three participants were active for over 90% of weeks – meaning they sent messages into the study for over 100 weeks. These participants were extreme outliers the 90th percentile of weeks active is 48% and the median is 13%. With over two years of messaging being active for 13% of the time represents around 14 weeks of interacting with the system.

A natural next question is: what, if any relationship exists between *weeks_active* and *percent_delivered*. Figure 6.9 plots *weeks_active* vs *percent_delivered* and shows no clear trend. However, *percent_delivered* is a clear limiting factor for *weeks_active*. Participants with a low *percent_delivered* also have a low *weeks_active*. The reverse is not always true – there are participants who received all of their messages but are never active. The main conclusion from Figure 6.9 is that message delivery is necessary but not sufficient for actively engaging with the system.

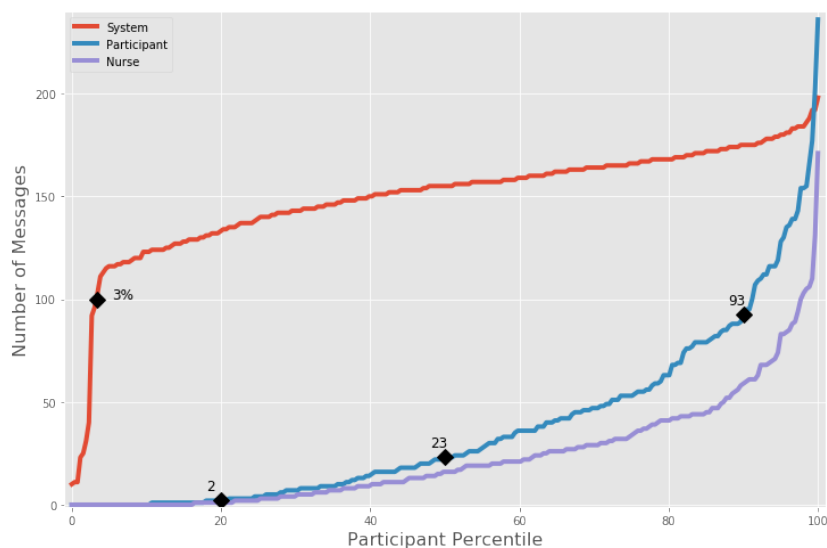


Figure 6.7: CDF of SMS totals for the two-way group. Only 3% of participants received fewer than 100 system messages. The median number of participant messages is 23 and the top 10% sent more than 93. A total of 26 (9.4%) participants sent zero messages.

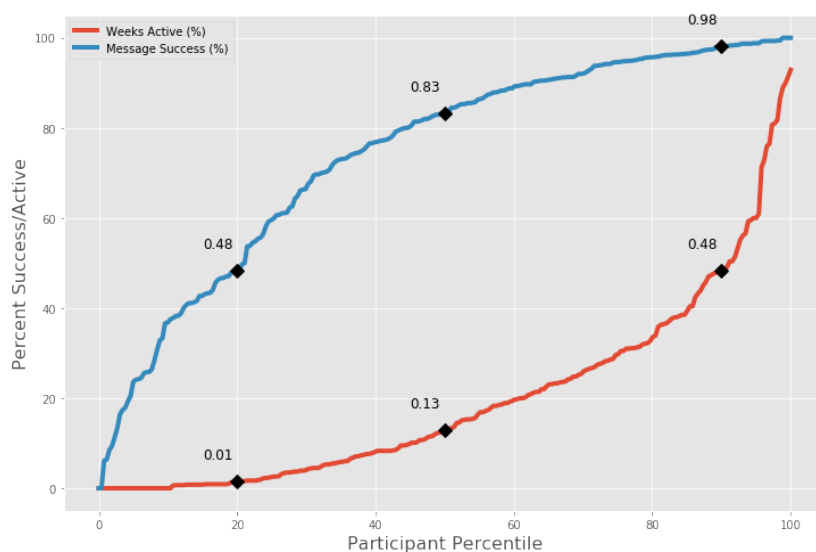


Figure 6.8: CDF engagement calculated as both message success and weeks active for two-way group. A flat blue line at 1.0 would represent 100% of messages delivered. The median of message delivery is 0.83 and the median percent of weeks active is 0.13. For context a weeks active ratio of 0.5 represents being active for over 52 weeks.

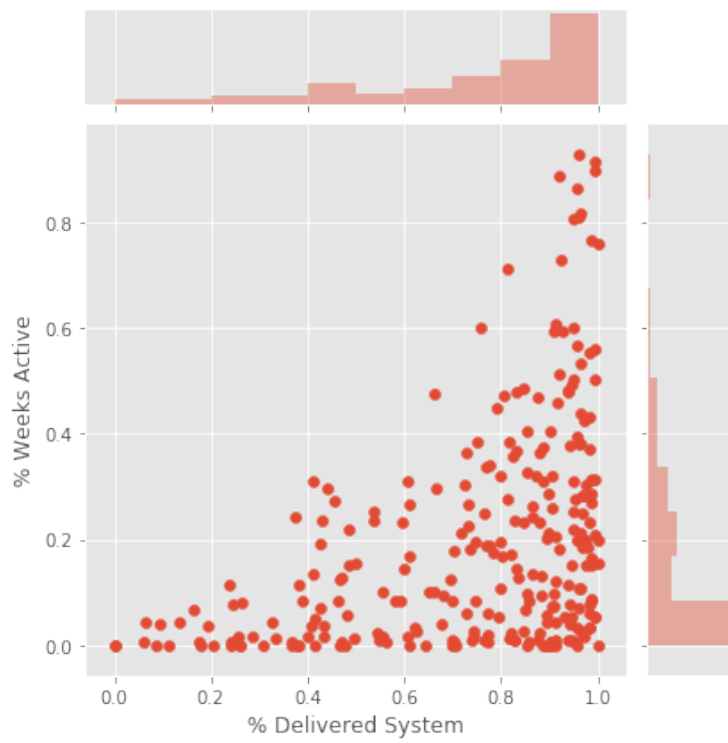


Figure 6.9: Relationship between percent of system messages delivered and percent of weeks active. No participant is active for more weeks than they successfully receive system messages.

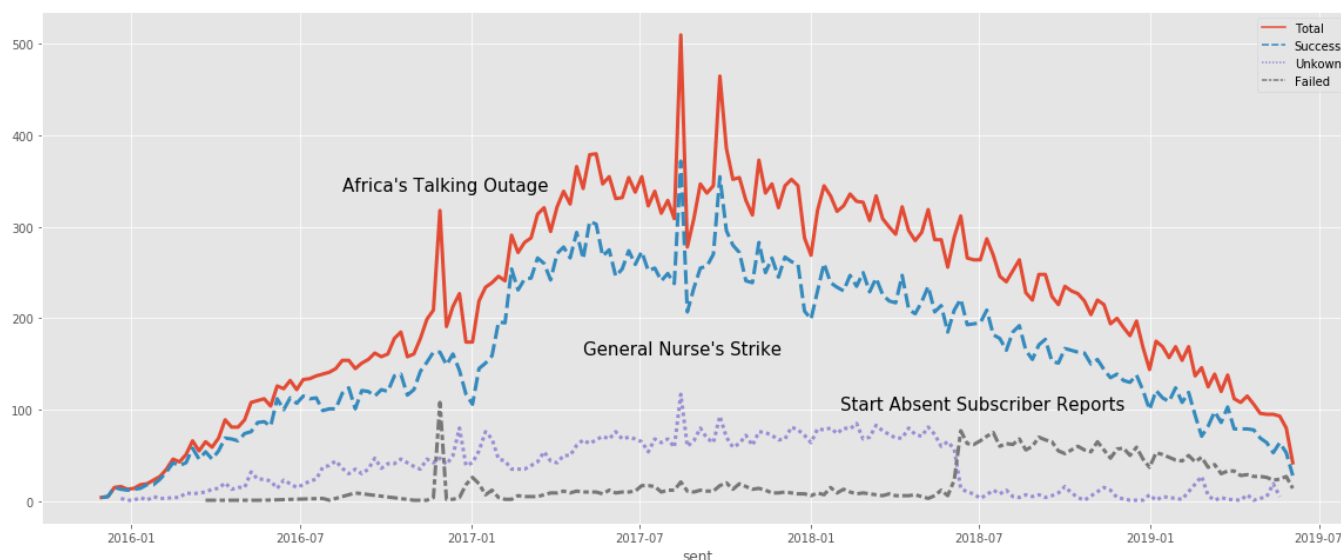


Figure 6.10: Total number of messages per week from the start of the study (Nov 2015) through the end of May 2019. This count includes all outgoing messages (system + nurse).

6.2.2 Message Error Rates

An SMS sent over Africa’s Talking can end up in one of three states: a *Success* when the message is successfully delivered, a *Failure* when a message fails to get to a phone, and *Unknown* when a delivery report is not returned for a message. Figure 6.10 shows the total number of SMS sent per week in red and splits this by message status in the other colors. This number includes both messages sent by the system and nurse from the beginning of the study through the end of data collection.

The first thing that stands out in Figure 6.10 are the multiple spikes in messaging. These occur for multiple reasons. The large spike in November 2016 is because of a three day outage with Africa’s Talking. Once outgoing messages were restored, all failed messages were resent, this resulted in a spike in both total number of messages and failed messages. In the three years the study has been in operation this was the only significant downtime. Another source of spikes in the summer of 2017 was because of an ongoing nurses’ strike that we sent messages for. During these spikes the total number of messages and successful messages both

go up. Lastly, there were dips in the total number of messages sent at the end of each year which represents a two week break for study staff and thus no outgoing nurse messages were sent.

The second thing that stands out in Figure 6.10 is how the unknown message count and failed message count swap in June 2018. This is the result of a change in the Africa's Talking delivery report API to report an *AbsentSubscriber* 24 hours after a message is sent. Before June 2018, the longest time between sending a message and receiving a delivery report was just under 24 hours. In this period 23.8% of messages were unsuccessful delivered and for 20.8% we received no delivery report. After June 2018, when *AbsentSubscriber* errors were being returned 29.7% of all messages were unsuccessful with only 3.0% having no delivery report. The *AbsentSubscriber* failure status is for SIM cards that are not currently on the network but still active – for example a phone is off or a SIM card switched out. Taken together this is 1) strong evidence for a 24 hour cutoff in SMS delivery and 2) a good indication that the vast majority of messages with a missing delivery report are for *AbsentSubscribers*.

The time of receiving a delivery reports gives an upper bound for the SMS delivery delta – which is defined as the time between sending system message and the message being received by the phone. Figure 6.11 plots a CDF of these delivery deltas with the x-axis on a log scale in hours. This graph excludes the 16.5% of SMS messages with no delivery report. Around 50% of messages are received within 14 seconds of sending, 75% of messages received 4 minutes after sending, 80% of messages received 35 minutes after sending, and 90% of messages received 11 hours after sending. At exactly 24h the remaining 6% of received delivery reports arrive with an error status of *AbsentSubscriber* – all other delivery reports with an error status are received within seconds of sending. The main takeaway of Figure 6.11 is that most messages get to a phone within 30 minutes and after that there is a low probability of a message succeeding.

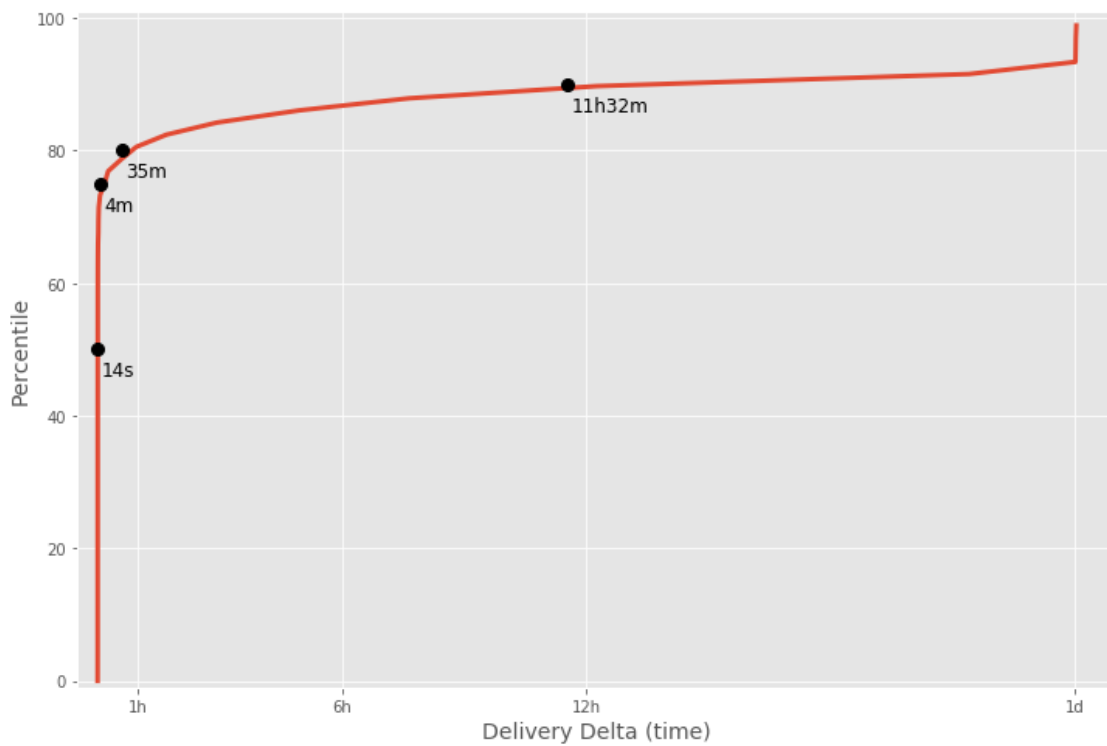


Figure 6.11: CDF of SMS delivery delta (the time between sending a message and receiving a delivery report). The x-axis is log scale in hours and the data is only for SMS messages with a delivery report – thus excluding messages with an unknown status.

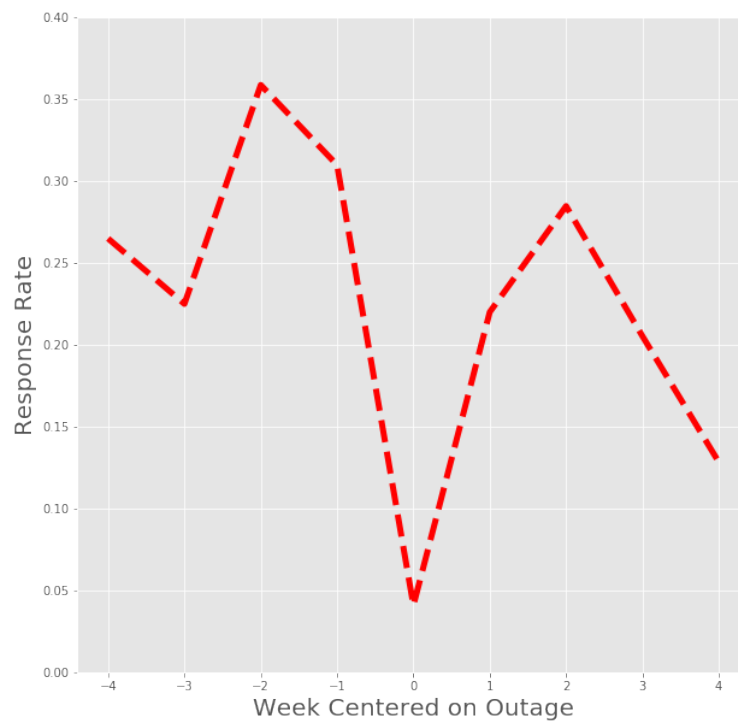


Figure 6.12: Percent of participants who responded per week for the four weeks before and for weeks after the Africa's Talking system outage Tuesday November 22, 2016.

6.2.3 A Natural Experiment

The outgoing SMS outage in November of 2016 created a small natural experiment to test how important the automated system messages are for creating engagement with the system. From 4pm on Tuesday, Nov 22, 2016 until 10am on Friday, Nov 25, 2016 all outgoing messages failed to send. However, during this time the system was still able to receive SMS messages. A total of 206 messages failed to get delivered during this system outage. Using this three day interruption of messages as a natural experiment we can observe what happens when the weekly prompts are not sent. How does not receiving message affect the response rate? Does missing one week affect the response rate in future weeks?

Figure 6.12 shows the total response rate (incoming messages divided by outgoing messages) for all participants effected by the outage for the four preceding and following weeks. The average response rate in the four weeks before and after the outage was 0.29 and 0.21 respectively. During the three days of the outage the system only received six messages from participants representing a response rate of 0.041. Additionally, none of these messages were from participants receiving messages during this period. In the third and fourth week after the outage the response rate drops which is due to the fact that this coincided with the Christmas holidays and many participants were traveling during this time. Previous findings from Chapter 4 and Chapter 5 have shown that most responses are received the same day as outgoing System messages and therefore act as check-ins, prompting engagement. This natural experiment confirms this by showing what happens to the incoming message rate if the automated prompts are accidentally skipped for a week.

6.2.4 Categorizing SMS Failures

Of the 39,993 messages sent by the system to the two-way group, 74.8% were successfully received. Of those that did not succeed, for 16.5% there was no delivery report and the status was technically unknown – though most likely these messages were all *AbsentSubscriber* failures. For 3,710 messages we received a delivery report of a failure or the message never successfully got to Africa’s Talking. These numbers are summarized in Table 6.5 with the

Status	Number	Percent
Total Sent	39993	—
Success	29902	74.8
Unknown	6591	16.5
Failed	3710	9.3
Error Type	Number	Percent
AbsentSubscriber	2345	63.2
UserDoesNotExist	679	18.3
DeliveryFailure	176	4.7
UserNotProvisioned	51	1.4
Other	459	1.4

Table 6.5: Breakdown of two-way system message status and frequency of common delivery errors. The vast majority unknown status messages are *AbsentSubscriber* errors from before June 2018.

breakdown of failed message by error type. A majority (69.5%) of failed delivery reports are for *AbsentSubscriber* errors. The next most common error is *UserDoesNotExist* at 18.3%. This error means that the phone number is not currently registered to Safaricom. The remaining errors are system failures of some type and account for just over 10% of errors.

Table 6.5 only counts system messages to the two-way group. We also have delivery reports for the one-way group but there is no difference in delivery success between the two arms. However, there is a large difference between the delivery success rate of system messages vs nurse messages. Figure 6.13 shows a violin plot of these three distributions. The median SMS delivery success rate for nurse messages was 94.6% but for system messages was 74.2% in the two-way group and 71.6% in the one-way group. This makes sense since the study nurse is always responding to incoming messages and someone who recently interacted with the system is more likely to still have their phone on. Since there is a variable amount of nurse messages per participant, for all of this analysis we only use the system messages.

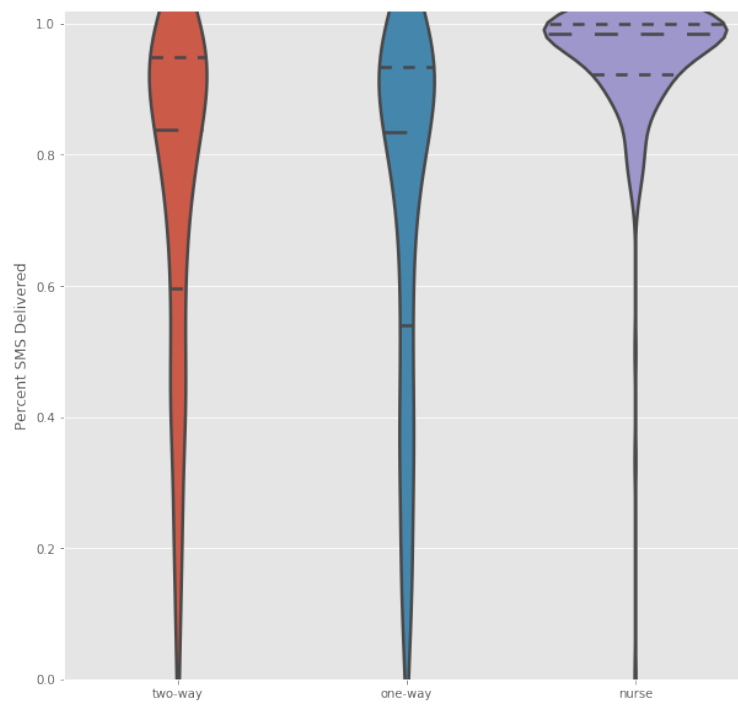


Figure 6.13: Violin plot of SMS delivery success rates comparing system messages to the two-way and one-way group and nurse messages to the two-way group.

Visualizing All System Messages

The main goal of this longitudinal analysis is to understand how participants engage with the system over time. To really see the complexity of engagement I developed single graph modeled of of the GitHub activity timeline to show the status and participant responses for every single system message. The heat map in Figure 6.14 is the result of many iterations attempting this goal. Each row of this heat map has a single square for each system message sent to an individual participant. The first message sent is on the left and the last message sent is on the right with gray squares meaning that participant received fewer system messages.

The three shades of green represent success with no response (light green), success with one response (mid-green) and success with more than one response (dark green). The two shades of orange represent messages with a missing delivery report (light orange) and AbsentSubscriber error (dark orange). These two statuses are linked because missing delivery reports most likely mean an AbsentSubscriber error. The UserDoesNotExist errors are in purple with all other errors in brown.

The participants are sorted by three metrics. First they are grouped into five bins based on successfully received messages. The large group at the top are all participants who received more than 80% of messages. The small group at the bottom are all participants who received less than 20% of messages. Participants are grouped again into four bins based on weeks active. So the dark green group at the very top are all participants who successfully received more than 80% of their messages and were active for more than 75% of the time. The large light green group at the bottom of the top group are all participants who received more than 80% of their messages but were active for less than 25% of the time.

There are several interesting groups. Participants at the top of the 20 - 40 bucket who responded at the beginning but then fail to get messages. The large light green section represents participants with little engagement but who are receiving the messages. Curiously, the purple UserDoesNotExist errors usually stop. Also, one can find unusual patterns like the dark green streak.

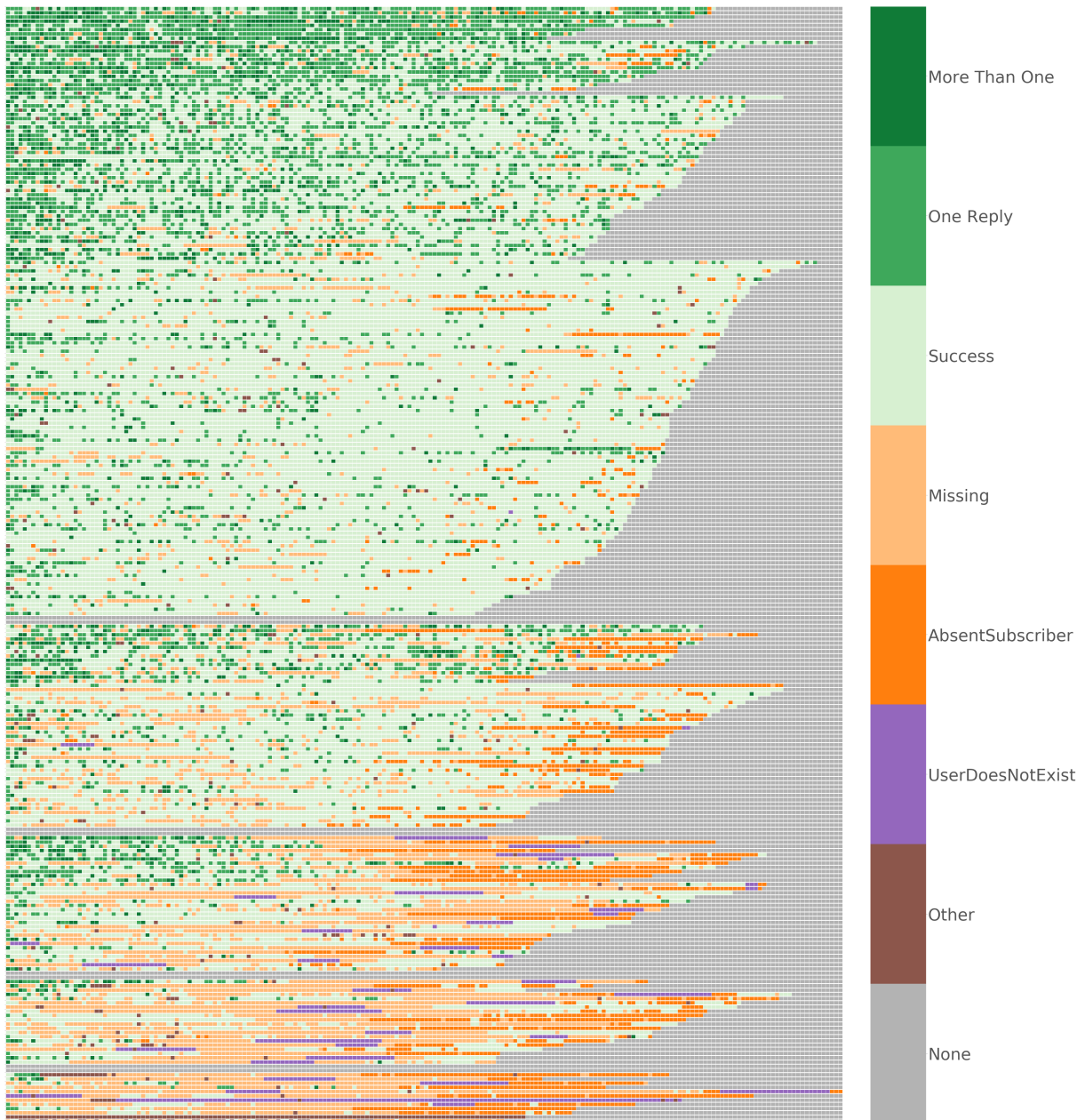


Figure 6.14: Heat map of all system messages sent to two-way participants in the study for longer than 100 weeks. Each row is all messages sent to an individual participant and each square is one message color coded by status. Participants are sorted on percent of successful messages, weeks active, and then total messages. The large group at the top are all participants who received more than 80% of messages. The small group at the bottom are all participants who received less than 20% of messages.

Another thing that stands out in the heat map of all system messages (Figure 6.14) is that many participants begin receiving messages again after long streaks of failed messages. In MomConnect – the large scale South African SMS project [210] – participants were removed from messaging after five consecutive messages failed. For MomConnect this meant that 10% of their participants were dropped out of the service. The rate of missed messages is much higher in MX and using a threshold of five consecutive failed messages 66% of participants would have been dropped. Figure 6.15 shows a CDF of the longest consecutive miss streak for MX participants. Half of participants have a miss streak longer than 7 messages and 25% longer than 29 messages.

In MX we never stopped sending messages when participants were missing them. This was done almost to a fault, since one participant failed to receive every single one of her 160 messages. However, this does give us an opportunity to see what happens to phone numbers who have had long missing streaks. The plot on the right in Figure 6.15 shows the future success rate of messages based on the current longest miss streak for a participant. For example, with a miss streak of zero, 74.7% of all future messages will succeed (our overall SMS delivery success rate). For participants who have a longest miss streak of at least 5 messages, 43.2% of future messages succeed. By a miss streak of 25, almost a third of all future messages still succeed. This shows that there are few participants who miss future messages, rather they go through periods of receiving messages and periods where the mobile phone can not be reached. Another, possible reason for this high rate of future SMS delivery after long miss streaks is that in MX, unlike in MomConnect, participants are able to register multiple SIM cards.

6.2.5 Messaging Changes Over Time

The heat map in Figure 6.14 is great for showing individual variation and giving a sense of the overall usage patterns. However, there is too much going on to really quantify engagement.

In FigureAggregating each metric per quarter (every 13 weeks) makes it easier to see the overall trends.

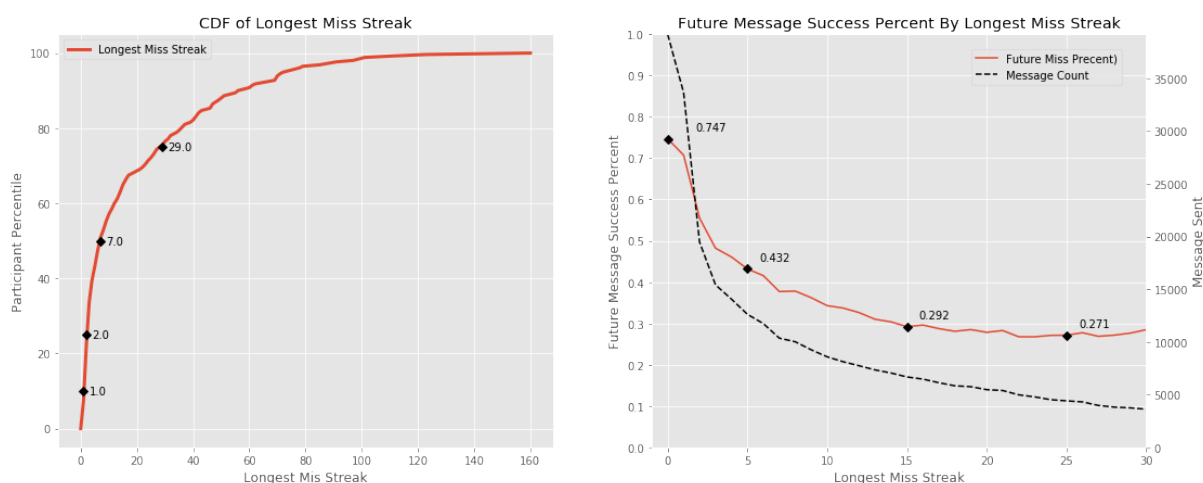


Figure 6.15: CDF of longest miss streak (left) and future SMS delivery success by current longest miss streak (right).

In Figure 6.16 each metric is aggregated to make it easier to see overall trends. On the left the average percent of successful message delivery, messages per week, and percent weeks active is shown. To capture change over time, all messages are grouped based on quarter since enrollment – with a quarter defined as 13 weeks. The ten quarters represent 130 weeks of system messages, the median number of weeks in the study was 120 so after ten quarters there are still a lot of participants receiving messages. Message delivery (red line) starts at 80% of messages and drops to 70% by quarter ten. The average percent weeks active drops by over 30% starting at 40% and drops to 7.6% by the end of the study. A similar drops can be seen with percent weeks active (purple line).

A complicating factor in understanding engagement after enrollment is that participants deliver their baby at different times relative to the start of the study. Delivery is a significant event in the mother's life and a focus of the early SMS content. A question that naturally arises is how mother's use the system just before and after they give birth. The graph on the right of Figure 6.16 shows SMS delivery success, weeks active, and messages per week centered on delivery instead of enrollment. The SMS delivery rate in the 8 weeks before and after delivery is high – this is because it is the start of the study for all mothers. The percent

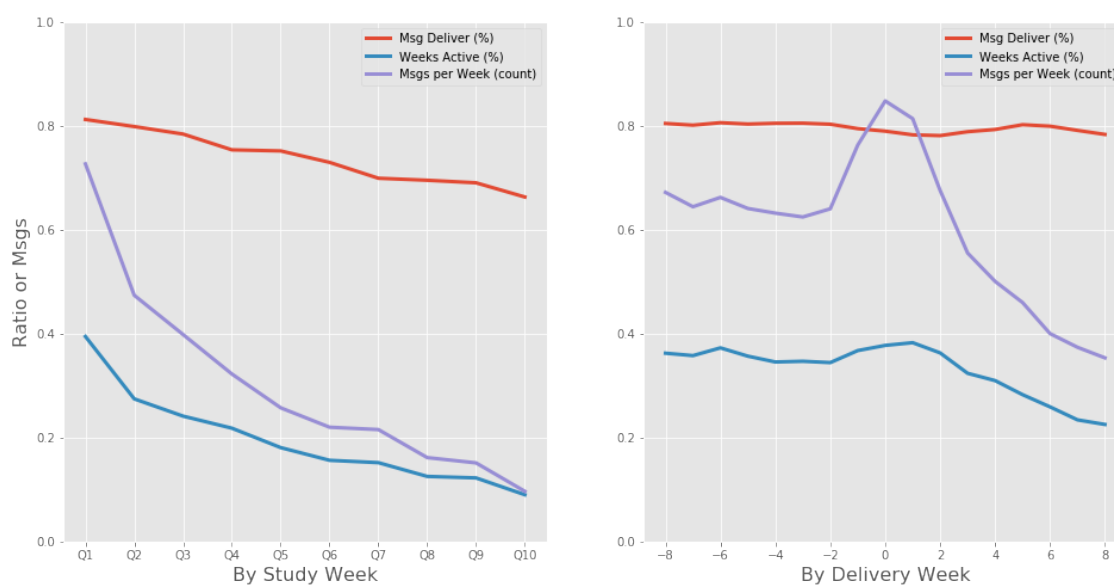


Figure 6.16: Message delivery (red), weeks active (blue), and messages per week (purple) by quarter since enrollment for quarter in study (left) and centered on delivery week (right).

of mothers who are active remains level at just under 40% until delivery when it starts to drop. This is partially explained by the fact that mothers are eligible to enroll until delivery as well as the fact that the ANC period generates a lot of questions. While there is no big difference in weeks active right around delivery, the total number of messages does go up – as shown by the bump in messages per week at zero. This means that while number of mothers who interact with the system right at delivery does not change, those who do interact send more messages than in the weeks before or after delivery.

The aggregated averages overtime in Figure 6.16 are useful for showing the overall trend; however, just like the graph from the WelTel study they do not show the full complexity. If instead of plotting the average we plot percentiles we can see how the most engagement and least engagement changes overtime. We know from the heat map in Figure 6.14 that there are clusters of participants that frequently respond, others that almost never respond, and still others that fail to receive most of their messages.

In the first graph on the left in Figure 6.17 we see that it is actually only the bottom

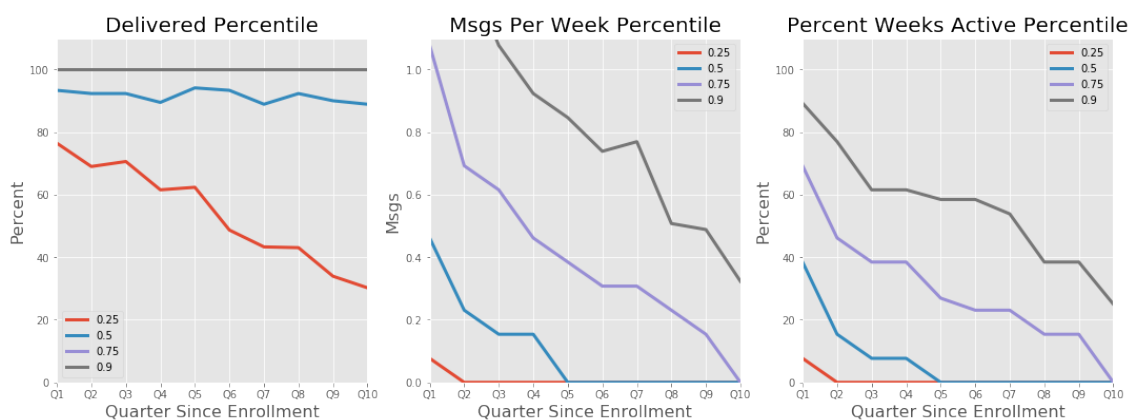


Figure 6.17: Percentiles for Message delivery (left), messages per week (middle), and weeks active (right). Percentile lines at 25% (red), 50% (blue), 75% (purple) and 90% (black).

quartile of participants who stop receiving messages. The 50th percentile stays at 90% for the entire study and the 75th is close to 100% – meaning they receive all messages the entire time. Meanwhile the bottom 25th starts at 78% and drops to 32%.

The right most plot in Figure 6.17 shows that the bottom quartile of participants are really one active in the first quarter and even then it is for less than 10% of the weeks. The 50th percentile is active for the first four quarters but not after that and the 75th percentile is active for nine quarters. The top 10% of responders are active for the entire study and while their percent of weeks active drops from 90% in the first quarter to 25% in the last this is still more active than the median participant for almost the entire 130 weeks. The percentile of messages per week (middle) is just a slightly scaled version of weeks active.

The three percentile graphs help tell a more concise story than the heat map. We see that there are a group of highly responsive participants as well as a group of participants that stop receiving their messages. At least 50% of the participants are active through at least one year but only the top 25% are active through two years. This means there is a large fraction of participants who are receiving messages in the last year of the study but not responding.

Another measure of engagement is the length of the messages participants send. The character limit for a single SMS message is 160 characters. Only 3.9% of participant messages

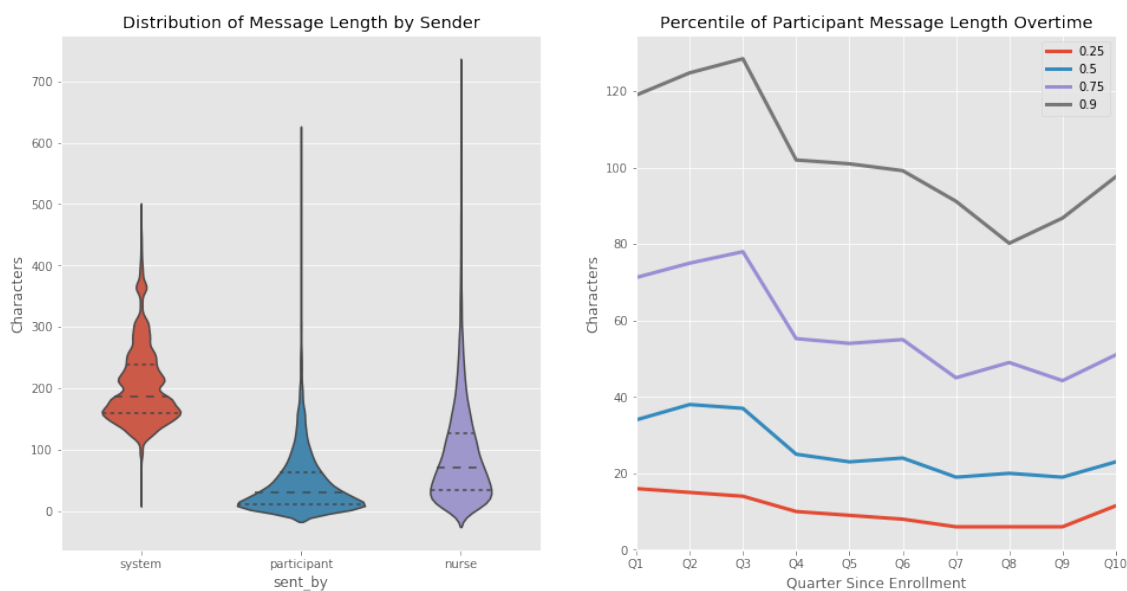


Figure 6.18: Percentile of participant message length by quarter since enrollment.

are over the 160 limit with the median SMS length being 30 characters and the IQR of [12 - 63]. (Surprising the longest participant message was 609 characters.) For nurse messages 15.8% were longer than 160 characters and the median length was 70 with an IQR of [34 - 126] (again the longest nurse message is an extreme outlier at 709 characters). The system messages are very different from the participant or nurse messages. A total of 70.1% of system messages are longer than 160 characters with the median length being 187 characters (already two messages) and IQR of [155 - 229] This is all summarized in the left plot of Figure 6.18.

The right side of Figure 6.18 shows the 25th, 50th, 75th and 90th percentile of participants incoming message length by quarter from enrollment. Rather than seeing participants message get smaller over time, they mostly stay the same. The bottom quartile of messages drops from 18 characters to 10 characters and the median goes from 40 characters to 30 characters. This shows that even though participants engage with the system less over time when they do engage with roughly the same amount of content.

6.3 From The Platform's Perspective

When doing a longitudinal SMS intervention the system must be synchronized with the real world. In the case of MX, the system needs to know each participants expected delivery date (EDD) and next visit date. Once a delivery happens the system needs to be updated with the new data so that ANC messages can be switched to postpartum messages. In this section we explore how well the MX system did at capture delivery dates and switching to postpartum messages. (The word delivery is overload here since we have SMS delivery as well as the delivery of the baby.)

First, Figure 6.19 shows the distribution of weeks between sign-up and EDD (on the left) and the difference between EDD and actual delivery date (on the right). This figure includes all arms of the study – but excludes participants who we never received a delivery notification for. On average, participants are 16 weeks from their EDD at enrollment and deliver two weeks before their EDD. The large spike in participants enrolled 12 weeks before their EDD corresponds to a scheduled ANC visit. Almost 25% of participants delivery on the week of their EDD, but for all other participants, even if the system learns about the delivery immediately, the messaging will be off from the intended schedule. Participants who deliver before their EDD will miss the last few weeks of ANC messaging and participants who deliver after their EDD will get repeated messages asking if they have delivered.

Knowing when a delivery occurs is critical for determining which message to send when. None of the clinics had a digital record system that can be queried to find this data. Rather, ANC records exist in paper logbooks such as the one from Mathare pictured in Figure 6.20) from which medical records must be abstracted and entered into the system. Abstracting delivery reports isn't even an option for mothers who go to a different clinic to deliver.

With this physical gap between the real world data and the system, the version of reality the system believes to be true drifts away from what actually happened. The distribution of weeks between actual delivery and when the system is notified of the delivery on the left of Figure 6.21. Again, this figure includes participants from all three arms of the study. For

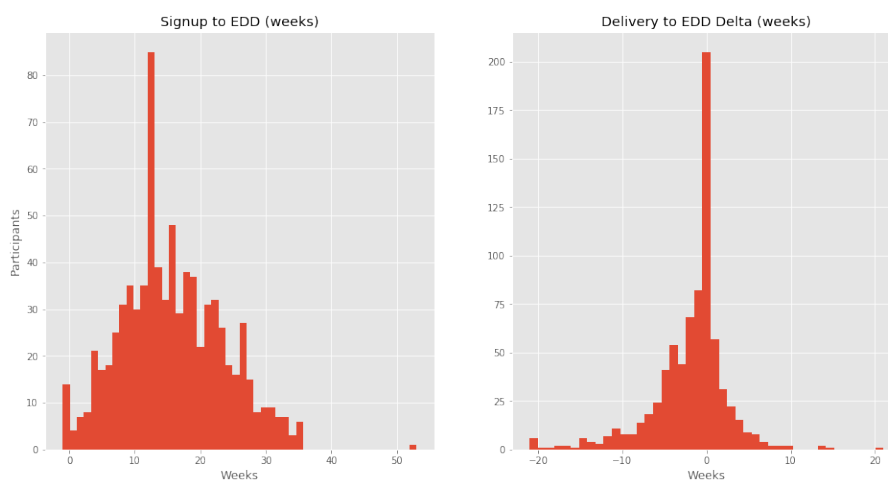


Figure 6.19: Distribution of expected ANC period (signup date to EDD) and difference between EDD and actual delivery.

30.4% of participants, the system learns of their delivery the same week it happened. There is a spike in notifications at 6 weeks after delivery when a clinic and study visit occur. There are 41 (5.4%) participants for whom the system learns of their delivery 13 or more weeks late. On the right side of Figure 6.21 we see the distribution of EDD to system delivery notification. This is a true representation of how far the system was off from reality since until information is updated the EDD at enrollment is the base for sending out messages.

The obvious question about the system notification delta for a participant's delivery is how it is influenced by the SMS messages. The system records how the study learned a delivery occurred through three main sources: an SMS message, a phone call, or a clinic visit. Figure 6.22 shows the same histogram from Figure 6.21 broken down by notification type and study group. What stands out is that almost 40% of the participants who notify the study of their delivery the same week it occurred did so through SMS. We also see that the large spike in notifications 6 weeks after delivery is all from clinic visits. On the right side of Figure 6.21 the notification delta is broken down by study group and the two-way messaging group accounts for almost 50% of all notifications during week zero.

The CDF of notification delta in Figure 6.23 shows that one week from delivery the study

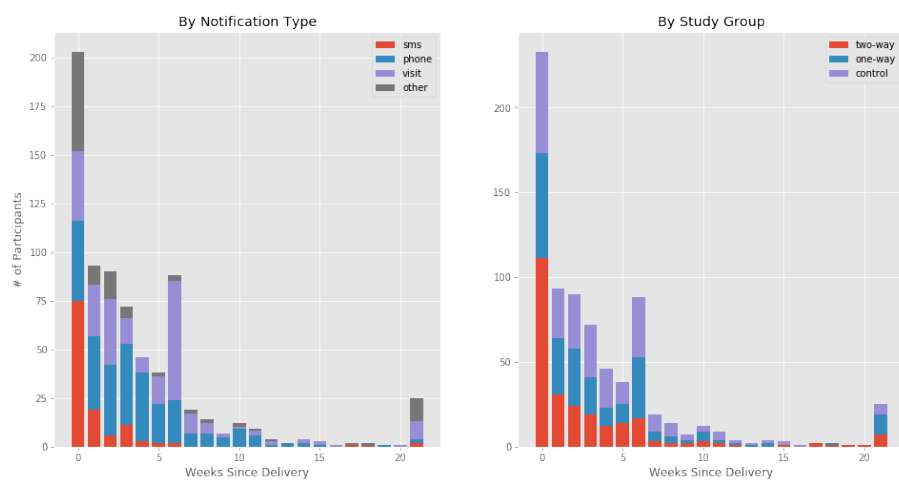


Figure 6.22: The delivery to system notification delta grouped by notification type (sms, phone, visit, other) on the left and by study arm on the right.

has learned about the delivery for 40% of the two-way but only 20% from the one-way and control. After the week 6 clinic visit there is no difference between the study arms in how many deliveries the study knows about. The fact that average notification times for the two-way group are so much shorter indicates that mothers are using SMS to communicate with study staff during the critical period right before and after delivery.

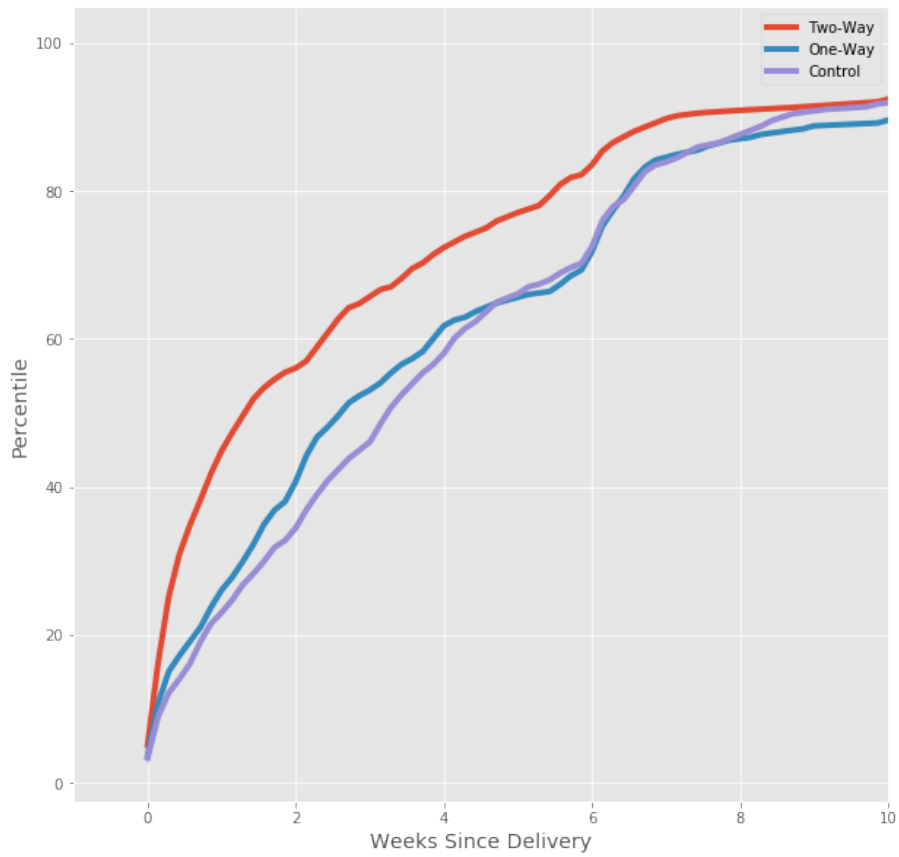


Figure 6.23: CDF of weeks between delivery and system notification (notification delta) split based on study arm.

Chapter 7

USSD: THE THIRD UNIVERSAL APP

7.1 Overview

This chapter takes a step back from SMS to examine the larger universal mobile service design space. These services allow researchers and practitioners to create interventions accessible on any mobile phone. Because of the prevalence of low-end feature phones this has traditionally meant using IVR or SMS based services that do not require a smartphone or data connection. The ability to limit participation in mWACH to only a basic phone was a primary reason for choosing SMS over an Android App that provided more features and a cohesive user experience.

There is a long history within the M4D community of using basic voice and text based services to reach and connect end users on the devices they own. In this chapter I argue for the use of Unstructured Supplementary Service Data (USSD) as a platform for universal cell phone applications. Looking through over a decade of ICT4D research this chapter describes how USSD can extend and complement current uses of IVR and SMS for data collection, messaging, information access, social networking and complex user initiated transactions.

The two key contribution of this chapter to the overall thesis are: 1) to situate bi-directional SMS platforms such as mWACH within the IVR, SMS, and USSD continuum and 2) explore the USSD design space and how an mWACH style platform could be ported to a USSD shortcode. Based on these findings we identify situations where a mobile based project should consider using USSD with increasingly common third party gateways over other mediums. This analysis also motivates the design and implementation of an open source library for rapid development of USSD applications. Finally, I explore three USSD use cases, demonstrating how USSD opens up a design space not available with IVR or SMS.

7.2 Introduction

Mobile phone applications have tremendous impact in global development across the domains of health, agriculture, education, and finance. The field, referred to as Mobiles for Development (*M4D*), builds upon the ubiquity of mobile phones and standard digital telephony. Two fundamentally different approaches are used for deploying services on mobile phones. One approach is to build services using generic features provided by the mobile operator, such as voice or text, ensuring the service is available on every single mobile phone handset. The other approach is to take advantage of capabilities of specific types of handsets, including installing and executing applications on individuals' mobile phones. Both of these approaches are important, but in this paper we focus solely on the first approach, targeting applications which are intended to have broad reach and high coverage across a population. One strength of this generic approach is that it can be used to reach people owning basic mobile phones. Even though smartphones have a growing market share, a very large number of people will continue to use basic or feature phones over the next decade. Being able to reach owners of basic phones is particularly important for development initiatives reaching the poorest people.

Mobile applications designed to work with carrier provided services, and not utilize features of specific handsets have generally focused on voice through Interactive Voice Response (IVR) and text through Short Message Service (SMS). IVR and SMS are *Universal Apps* in that they make services uniformly available on every mobile handset. There is a third option for a Universal App, which is less frequently used in M4D: Unstructured Supplementary Service Data (USSD). USSD is a protocol defined inside the Global System for Mobile Communications (GSM) standard.¹ USSD is a session based protocol, like IVR, that supports the exchange of text data, like SMS, thus filling a gap in the M4D design space. The most common uses of USSD are for customers interacting with a carrier's services, such as querying

¹USSD is not supported on CDMA networks, however GSM covers 90% of the world and is particularly ubiquitous in those regions with the lowest smartphone penetration.

HTTP Only Gateways	SMS	IVR	USSD	Coverage
Twilio [†] , Nexmo [^]	✓	✓		World Wide
txtNation [±]	✓	✓	✓	World Wide
Infobip	✓	✓	✓	9 different countries
Africa's Talking [‡]	✓	✓	✓	Kenya
Ideamart [#] , SMSGH [%]	✓		✓	Sri Lanka, Ghana
Panacea Mobile ^{\$}	✓		✓	South Africa

[†] http://twilio.com/ [^] http://nexmo.com/ [±] http://txtnation.com [*] http://infobip.com [‡] http://africastalking.com [#] http://ideamart.lk [%] http://developers.smsgh.com ^{\$} http://panaceamobile.com				
SaaS Enabled Gateways	SMS	IVR	USSD	Coverage
VumiGo ^{\$}	✓		✓	11 countries across Africa
Voto Mobile [¶]	✓	✓		Ghana, Tanzania, South Africa
TextIt [♣]	✓	✓		Android, Twilio
Telerivet [◇] Frontline Cloud [♠]	✓			Android, Twilio, Nexmo

^{\$} http://vumi.org [¶] http://votomobile.org [♣] http://textit.in/ [◇] https://telerivet.com [♠] http://frontlinesms.com				

Table 7.1: Third party REST gateway providers and SaaS companies.

airtime or subscribing to information services. This is done by sending a *star code* (such as *144# to check an airtime balance with Safaricom in Kenya) and accessing interactive menu systems. USSD has a number of advantages over SMS, such as providing greater privacy, which makes it a candidate for some important M4D applications. However, until recently, it has been very difficult to deploy USSD applications due to requirements of working directly with carriers. As third party companies have begun making USSD services available, there are growing opportunities for M4D to utilize a new Universal App.

In this chapter we explore the application design space for USSD by identifying fundamental considerations and trade offs between IVR, SMS, and USSD. By analyzing existing universal applications we propose methods for USSD to enhance and improve M4D services such as data collection, health messaging, and accessing data. Based on these findings we describe the design and implementation of a library for rapidly building USSD applications. Finally, we develop three detailed use cases to show how USSD applications in different domains provide services that neither SMS or IVR alone can accomplish. -

There is a large body of work showing that universal applications built on top of SMS and voice have a significant part in the ICT4D design space. These projects have shown how universally accessible mobile applications can be used as a catalyst for change and demonstrate the role USSD plays in this design space. One of the first ICT4D projects showing the potential of SMS to simplify interactions and reach end users was Warana Unwired [199] which converted a kiosk based information system into an SMS query system. Since then many more projects have explored the SMS M4D design space which we will analyze in Section 7.3.

Within the field of mHealth there is mounting evidence for the importance of SMS and IVR at engaging patients to improve health outcomes. WelTel [113], one of the first SMS based mHealth studies, showed that sending simple, single word SMS messages significantly decreased HIV viral load among ART patients. Following on this work other projects have shown that IVR can offer a richer user experience for patient engagement [98, 143].

Besides messaging, another major use for IVR and SMS solutions has been data collection in the field. Patnaik et al. [149] did an early study in India comparing forms, SMS, and voice and showed that while voice was the most costly it was also more accurate than SMS prototypes. However, Danis et al. [45] showed that SMS based surveys were successfully answered in Uganda. One reason for these seemingly contradictory results just one year apart is the different demographic and cultural context between Uganda and Gujarat, India.

7.3 Universal GSM Applications

SMS, IVR, and USSD can be considered universal applications since each protocol is outlined in the GSM specification and works on every single handset. Universal mobile applications are well suited for M4D deployments that aim to reach end users. There is no need to customize applications for individual handsets as J2ME deployments require or provide standardized hardware to all participants, both of which hamper the ability of a project to scale and reach all target users [162]. The use of these communication channels represents a fundamental trade-off between universal access and the richness of user interactions due to the limited user

interfaces of universal applications. In the next two sections we explore the design space of universal M4D applications showing how USSD complements existing services. We also layout guidelines that assist organizations starting M4D projects to decide on which technologies suite their design goals and project requirements.

7.3.1 Design Space and User Interactions

We first analyzes applications built on top of SMS and IVR to understand the types of user interactions and platforms in use. Fogg and Allen [65] identified five categories of SMS applications for health. Even though this work did not focus on a developing world context, the universality of SMS means it can be applied across applications in the M4D domain and onto other universal modes of mobile communication.

Projecting Fogg's persuasive SMS computing model onto the existing literature we examine well known projects and identified five specific application domains for M4D services. In each of these domains there is a substantial body of work demonstrating the effectiveness of both SMS and IVR. USSD has not been as widely adopted, but the few examples of USSD applications show that it is well suited for complicated user interactions.

Data Collection

Data collection is an important domain for ICT4D acting as an enabling tool to improve service delivery and evaluation. For projects that can provide smart- or feature-phones, there are powerful data collection applications such as ODK Collect [84] and CommCareHQ. However, when data collection is required across large populations or must be accessible on personal mobile phones, default GSM solutions are the most feasible option. In this case the universality of SMS, IVR, and USSD trumps the improved UI and features of device specific applications.

While, SMS has been a common choice for data collection, the unstructured and asynchronous nature requires special syntax for submitting data. In the SMS for Life project, weekly SMS reports were collected containing stock levels for antimalarial medications and

rapid diagnostic tests in rural Tanzania [14]. Each tracked resource was assigned a single letter code. By combining these codes, all necessary fields were reported in a single SMS. This allowed structured data to be automatically parsed from incoming SMSs. Data showed a high average response rate (95%) and low average error rate (7.5%) indicating that, at least in rural Tanzania, simple SMS reporting is feasible at large scale.

Beyond simple data collection, various methods have been developed to capture large amounts of structured data via SMS. Using a RapidSMS based system, Assiimwe deployed reporting system that captured weekly disease counts and stock levels [11]. With almost 40 data fields, a single SMS per field was cumbersome and expensive, while putting all fields into a single SMS was complicated. Instead, four different SMS messages were created with the assistance of a paper job aid. This eliminated much of the syntax while maintaining an average error rate of 8.8%. Other examples of job aids assisting message creation include the Reporting Wheel from Instedd ² to help translate data into nine digit numbers for easy submission and parsing.

SMS data collection can also be accomplished with a multi-SMS survey form. This type of data collection is exemplified by early work from Text for Change sending HIV related quizzes via SMS in Uganda [45]. Each question is a single SMS and participants are asked to respond with a question specific key word prefixed to their answer. These multi-SMS surveys have been standardized as part the SaaS IVR and SMS services from Table 7.1 as well as services dedicated to data collection such as CommCareHQ.

Voice data collection is a good option for low literacy users [121]. Previous results have shown that voice data collection can be more accurate than both SMS and digital forms [149]. Using an IVR system numeric and multiple choice questions can be recorded using Dual Tone-Multi Frequency (DTMF) touch tones and open ended verbal responses can be recorded for later analysis. Multiple projects have used IVR for data collection [78, 111] and just like with SMS, third-party SaaS services have emerged to provide voice data collection.

²<http://instedd.org/technologies/reporting-wheel/>

By using USSD for data collection, the user experience and work flow can be improved in three ways. First, including all transactions in a single session simplifies the user experience around multi-SMS forms. Second, the menu system acts as a built in job aid eliminating the need for complex syntax to fit as much information as possible into a single SMS. Third, using USSD for data collection allows for immediate data validation during the session flow. These advantages of USSD are most helpful for organizations who can not train all end users or when data collection may happen infrequently and enumerators never have time to become comfortable with complex SMS syntax.

Messaging for Awareness

Messaging for awareness and engagement was a cornerstone of Fogg's persuasive technology SMS framework and has become an integral aspect of ICT4D work. The theory behind messaging campaigns is grounded in behavior change communication and relies on two main assumptions: (1) open communication channels improve uptake of available services and (2) frequent small reminders help with the adoption and maintenance of new practices.

The Kenya WelTel project was among the first controlled studies measuring the effectiveness of messaging for awareness [113]. For 12 months patients initiating antiretroviral therapy (ART) received a weekly SMS with the single word *Mambo* (How are you). They were asked to reply with either *Sawa* (Fine) or *Shida* (Not Fine). Even with such simple messaging content, the WelTel study found a significant decrease in HIV viral load between those receiving and not receiving SMSs.

The low cost, quick content creation, and ease of sending SMSs has meant many other projects use it for awareness and engagement. Across the medical domain, projects focused on both health workers [215, 11, 53] and patients [160, 171, 150] have demonstrated high levels of engagement and changes in behavior. In places with low literacy similar projects have targeted maternal health outcomes using voice such as making automated voice calls from a doctor to pregnant women about iron supplements [143].

A review paper on mHealth found that one reason to prefer IVR over SMS is to ensure

privacy for vulnerable populations [79]. For example, a project working with men who have sex with men decided to use voice calls instead of SMS primarily to ensure messages were not saved on shared phones. Recent work with HIV positive youth in Kampala, Uganda has shown that privacy is paramount when sending SMSs [165]. In this respect USSD has the same level of privacy of voice based systems. By requiring a PIN before accessing medically sensitive information, a USSD system can more directly address individuals needs while at the same time protecting doctor patient confidentiality - especially if the phone is shared among individuals.

Accessing Information

Ubiquitous, on demand access to information is one of the most revolutionary features of mobile computing. Enabling end users at the bottom of the pyramid to find information when they need it has been a major challenge in ICT4D. An early project using mobile phones for information access was Warana Unwired [199] which replaced a PC-based kiosk system for information retrieval with an SMS interface at a sugarcane cooperative in India. This system used a smartphone connected to computer to allow correctly formatted SMS queries to access information previously only accessible via a computer. Another ICTD project enabled SMS querying of taxi arrivals in Kyrgyzstan [10]. One finding from that work was that USSD would be more cost effective and user friendly.

The fact that SMS is always available has been a major driver for query based systems. Google and Yahoo both had SMS interfaces [175] prior to more data intensive and graphical interfaces enabled by smartphones. The limiting nature of the SMS responses has lead to work optimizing for this low bandwidth channel [186]. Chen et al. created an SMS based web search system specifically designed for SMS results on low end phones [28]. Voice channels have not been used for data querying largely due to the fact that automatic speech recognition does not work well for the vast majority of languages in the world. The primary use case for voice channels has been to create call in hot-lines accessing information through a live operator [37] [149].

Creating applications for information access over USSD adds a layer of interactivity to SMS based systems. For example in a series of prompts the Vumi Wikipedia Zero application³ allows the user to narrow down a search before returning the first 180 words of the page and section requested. With the user interface of USSD a sugarcane cooperative could provide a menu driven application for accessing individual information as well as a query interface for information about best practices or new procedures. Like SMS, USSD works best with short text based responses and requires the development of backend processes specifically designed for this limited channel.

Social Networks and Group Messaging

Since the first Usenet newsgroups, ICT has been connecting individuals, friends, family, and the larger community. It is no surprise that ICT4D research has looked at connecting beyond the default capabilities of SMS and voice. Odera created Tangazo, as a group messaging platform. Subscribed users could send either voice or SMS messages to custom made groups, and a special SMS syntax was used to manage group membership [137]. Safaricom in Kenya has a small SMS based group SMS messaging service called Semeni⁴ that allows users to create groups of up to 10 members managed through a USSD interface.

USSD as a user interface is well suited for small group social networking. Using our library it would be possible to easily create digital groups connecting and strengthening existing community ties. Grassroots community groups are a common mechanism for development. For example, community health workers lead mothers groups and peer HIV support groups and peer structures are fundamental to microfinancing. A USSD menu system could be used to help facilitate group interactions outside of meetings as well as connect different groups.

While Tangazo connected small groups together IVR systems such as CGnet Swara [126] and Avaaj Otalo [148] create voice networks linking large groups of people. At scale curating user submitted content for access over menu systems becomes an issue [197]. In a large USSD

³<https://github.com/praezelt/vumi-wikipedia>

⁴<http://www.semeni.co.ke/>

based social network similar concerns would arise since search and discoverability are limited to 180 characters. For this reason, we see the benefit of USSD in the social networking domain primarily to help create small focused community based groups.

Complex User Transactions and Interactions

A paradigm shift occurs once useful functions can efficiently and economically be performed on a mobile device. The mobile is transformed from a one-to-one communication tool into a device for accomplishing tasks and interacting with larger systems. However, it is when complex tasks need to be performed that the fundamental limitations of universal GSM applications make systems overly complicated. For example, findings from an SMS based agriculture trading system in Uganda showed that almost all the messages received could not be automatically parsed [183]. Again, the study authors suggested that USSD would be a viable alternative to the complex SMS structure.

Because of their session based nature, both IVR and USSD are more suitable for performing complex user interactions. Many MOs have IVR systems that allow users to top up via DTMF key presses and perform basic administrative features for prepaid lines. One of the most complex IVR universal applications deployed is TAMA which provides treatment support for HIV+ individuals and was tested in India for 12 months [98]. Protected by a PIN patients receive calls from TAMA with reminders to take medication. The system also has the ability to record adherence. Patients can call the system to inquire about symptoms or listen to 30 second health tip messages.

Another example of complex transaction tasks via GSM applications is a pilot project to book train tickets sponsored by the Indian Ministry of Railways [1]. This services is offered on all three universal application GSM channels. The SMS message syntax is quite complex, requiring one SMS with train and station codes plus a correctly formatted date and a second SMS to confirm and authorize mobile billing that has five unique fields. This is in contrast to the USSD interface consisting of a series of menus to select a station, date, and ticket type. Authorization of mobile billing with PIN code occurs within the same session and

confirmation details that act as an eTicket are sent back via SMS.

7.3.2 Advantages and Limitations of GSM User Interfaces

In this section, we analyze the three universal GSM applications from the view point of organizations wanting to deploy M4D solutions and the users they target. When launching a mobile based project the choice of underlying technology is dependent on many factors including, but not limited to, target demographics, administrative overhead, usability, setup costs, maintainability, and scalability. There is no universal method or technology set for a mobile based project and each deployment must be evaluated individually, based on the target country, demographics and partners.

Setup and Infrastructure

When starting an M4D project the first question an organization must ask is what technology is available. Connecting with the telecommunication system can be difficult with two main pathways. At the do-it-yourself (DIY) level, a commodity phone can be used as a modem. With tools such as IVR Junction [198] and FrontlineSMS, a working system can be setup relatively quickly and with low overhead. However, these solutions do not scale well. A more robust solution is needed if simultaneous voice calls must be handled or a high rate of SMS need to be sent (over 1000 SMS/hour). A second limitation of DIY solutions is the inability to have toll free numbers and short codes. So, while setup up costs might be low, it is expensive for end users. Projects often look for ways to send airtime reimbursements [181]. There is also a reliability concern for DIY solutions, since the phone-modem is a single point of failure, risking running out of power or airtime, or loss through damage or theft. DIY solutions are not available for USSD, which is a major reason why USSD applications have been limited to large organizations.

At the next level of complexity, there exists an ecosystem of third party gateways who have partnered with mobile operators (MOs) to make available HTTP gateways for voice, SMS, and USSD. Table 7.1 lists eight of these gateways in two different categories - large multinational and smaller country specific gateways. The multinational gateways can reach

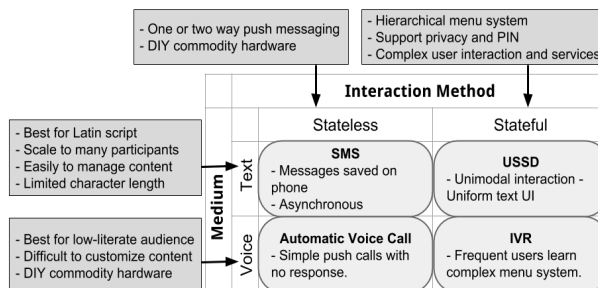


Figure 7.1: Comparison of GSM applications design space split by communication medium and interaction method.

many countries because they send messages from international numbers into countries where they do not have a partnership with an existing MO. This makes prices much higher for places like Kenya and Ghana using Nexmio over local gateways that have direct deals with country MOs. In the last few years the the number of these gateways that offer USSD APIs has been rapidly expanding and this is a significant motivation for this work. We expect more USSD gateways to open up in the future. For example InfoBip claims that USSD services will soon be available in Brazil, Peru, India, Thailand, and Ukraine.

Communication Channel Usability

Each GSM transport layer creates a different user experience and choosing the appropriate one depends on the target audience and service complexity. Figure 7.1 groups each channels based on the mode of communication and the method of interaction. On these axes SMS and USSD are the text based analogs of automated voice calls and IVR. From this perspective USSD can be viewed as a bridge between the session based features of IVR and the text based mode of SMS.

There are four primary considerations for choosing between text and voice channels. The first, and primary reason, is that for most projects creating and collecting text based content will have lower administrative cost. For example three hundred preconfigured text

messages can be modified with the name of sender and recipient when sending an SMS [150], prerecording these as voice messages means content is not as flexible or customizable. This is a major reason to favor text over voice solutions. However, the second consideration is the literacy proficiency of the end users - which may require that voice be used. A data collection program used by trained CHWs can use SMS [11], while, a program targeting pregnant women from a Mumbai slum should use voice [143]. Another reason to potentially prefer voice over text is that non-Latin script support on low end phones is often poor [120]. Lastly, it is important to consider the bandwidth and cost requirements of each channel. While SMS and USSD have extreme character restrictions (160 and 180 chars respectively) they can easily scale to many thousands of messages an hour. Voice, on the other hand, scales well with larger content but establishing multiple simultaneous voice lines raises costs. Such concerns are very dependent on each country, for example voice is cheaper relative to SMS in India while it is the reverse in Kenya.

Stateful and Stateless Messaging

The requirements of an individual project will also require a decision between the session based interaction which maintains state, and stateless automated calls or SMS messages. Stateful applications allow for a broader design space but come with increased setup and maintenance cost. If an intervention simply needs asynchronous messaging there is no need to move beyond SMS or automatic voice calls. The SaaS SMS solutions from Table 7.1 have a backends that simulates stateful connections over SMS in the same way web servers convert HTTP from a stateless to stateful protocol.

However, for certain applications the true session based interaction of USSD and IVR cannot be mimicked. First, session based interactions are not saved on the mobile phone. This is extremely important for guaranteeing privacy from friends and family, particularly when a shared phone is involved and sensitive health topics such as TB, HIV, or family planning are discussed. An additional layer of privacy is gained by entering a PIN before taking any actions. Although not secure against man-in-the-middle or operator snooping,

this is secure against friends and family from learning about a sensitive medical condition. Other reasons to prefer a session based medium are that hierarchical menus are much easier to implement, real time validation is smoother, and interactions happen in real time.

7.4 USSD Library Design

We next describe our library designed to create USSD based applications that work with existing third party USSD gateways. The library is created with the Django web-framework facilitating integrating with Python based message routing platforms such as Vumi and RapidPro. The intention behind this open source reference library is two fold: first to help define the building blocks of generic USSD applications; and second, to provide M4D deployments targeting end users a third option besides SMS and IVR. Currently, the intended user of this library is an organization that has accesses to programming experience, but we are working on methods to create USSD menu systems and questions using Excel worksheets or a graphical flow visualizer. Eventually we hope that third party SaaS messaging providers will incorporate these USSD features into the services they provide.

The library is a set of Python class that can be extended to quickly implement the flow and organization of a USSD menu system. By extending the basic *USSDTransport* class applications with the particular variables required by a third-party USSD HTTP APIs our framework assists with establishing text based sessions. Currently our prototype implementation interacts with the Ideamart USSD simulator⁵ (Figure 7.2) and the Panacea Mobile USSD gateway⁶ via their HTTP API.

Unlike stateless SMS applications a USSD application must maintain the users state throughout each session. Both Panacea Mobile and the Ideamart USSD simulator interact over the stateless HTTP protocol, so an important feature provided by the library is the seamless management of session state. This is done by using the session ID provided by each USSD gateway API as a UUID for a Django HTTP session. In this way each initiated USSD

⁵www.ideamart.lk/idea-pro/ussd

⁶www.panaceamobile.com/gateways/ussd-gateway/

session maintains its own current position and state on the server side and can determine the next USSD screen to send based on the current state and user input.

A USSD application is defined by a linked graph of nodes called *USSDScreens*. The most basic node is a `TextNode` that simply renders a template string to be sent to the user's phone. Each screen is rendered using context variables associated with the current session and user. For example if each contact has a next visit date, preferred language, and age variable linked with them these variables become available in template strings. Two additional nodes extend from `TextNode`. The `MenuNode` links other nodes together attaching numbered options that map to other `USSDScreen` instances. `MenuNodes` provide consistent navigation throughout the library; the numbers one through eight jump to a next node while '0' goes to the last screen, '9' goes to the next screen of a long list, and '#' will go back to the home screen. The characters for all of these actions are easily configurable and can also be disabled on individual screens. We also implement the `QuestionNode`, which validates response and establishes branching logic for the next node based on the response making it possible to create simple forms with a few Python classes. For more specific applications the `TextNode` class can be extended.

The library also has optional screens that can be added before the home screen is shown. This can be used to present the user with the option of initiating the USSD session where they left of last time or requiring that a correct PIN number be entered before the home screen is shown. Because the USSD library is built on top of the Django framework it also makes it easy to have a web accessible admin interface. Administrative tasks include replying to messages left by users in the USSD system, monitoring usage statistics, and exporting collected data. The library also makes it easy to open a web accessible front end that mimics the work-flow of the USSD system. Thus users who have a data enabled phone can access the application over USSD or a web browser. We are also working on an Android application to access the same information over a built in API. Using this library we have built a prototype Maternal Health application for testing the work flow and integration of the libraries building blocks. (Figure 7.2)



Figure 7.2: Left: Example Maternal Health USSD application running on IdeaMart USSD Simulator; Right: USSD home screen for Indian mobile operator BSNL.

7.5 USSD Use Cases and Personas

In this section we explore how the building blocks of our USSD library can create applications that improve the usability and quality of universal GSM applications. These use cases illustrate several advantages of USSD over SMS, practicality when it comes to real time interactions and privacy. Although some parts of these services have been implemented with SMS or IVR, the full USSD implementation is not replicable on any other default GSM medium. They also highlight how the general flow of USSD applications is easily extended beyond the traditional financial services application. However, they also expose limitations of basic GSM applications and why for some cases it might make sense to target phone specific applications utilizing a data channel.

We introduce each use case through a target user's persona and demonstrate how she or he would utilize different USSD applications to accomplish tasks and access services.

7.5.1 *Maternal Health*

Mercy is a 20 year old Ugandan living in the informal settlement of Namuwongo, southeast of Kampala. She is among the ninety percent of Namuwongo residents living on less than one dollar a day. Twenty-two weeks into her first pregnancy she has just attended her first antenatal care visit. There was a two and a half hours wait before she had twenty minutes with a nurse who took vital signs, listened to the babies heart, and drew some blood. Mercy was ushered into a room with twelve other first time mothers where they were told what to expect during pregnancy, where to give birth, and given booklets to record clinic visits and their baby's milestones. She thought she was done after this session, but the nurses sent her to another room - this time with just four other expectant mothers. In this room Mercy found out she was HIV positive.

There was a torrent of information and although the nurses tried to be understanding Mercy found it difficult to focus on this new reality. The nurse asked her for a four digit personal identification number so she chose the last two digits of her and her sisters birth years. They told her that if she called a special number - it looked like a code to top up a phone - she could enter her name and the secret number to access information and ask a nurse questions. Before leaving each of them practiced logging into the system.

Later when Mercy tried the system she liked how the first screen did not mention HIV or being pregnant, rather it displayed a generic health tip with a box to type her name. After hitting send a new health tip displayed with another box, Mercy didn't know what to do with this box so she just hit send again. This time the system told her she had used her two health tips and to call again. That reminded her what the second box was for, her PIN number. After correctly signing in she was presented with four options. Press 1 for New Messages, Press 2 for Starred Messages, Press 3 For Questions, Press 4 for Important information. Going to the new messages Mercy saw the welcome messages, the nurses had told her that this message would change every week. Mercy followed instructions to star the welcome message since it was the only message she had. The screen told her to press # to go

to the beginning and from there she checked the important dates menu which told her how many days until her next visit and when her due date was.

This use case demonstrates how USSD can augment existing SMS based systems that support marginalized individuals living with HIV. By requiring a user name and password to access the USSD system mothers who share a phone with family members can call from any available phone. The application looks like a generic health tip service without proper credentials. We acknowledge that this system still has some security vulnerabilities since USSD communication is not encrypted - but is an improvement on SMS only solutions.

7.5.2 Agriculture

Solomon is a Ghanaian farmer living in Gushie, 50km north of Tamale - the capital of the Northern Region. He and his brothers grow cassava, yams, kola nuts, and bananas for a regional farmers cooperative based in Tamale. The cooperative runs a marketplace to help match produce buyers and sellers. Every time Solomon goes to Tamale, he checks the records for how much his farm has sent to the cooperative, when the number doesn't match his expectations it can take awhile to sort out. The cooperative is also a great place to exchange information with farmers from villages Solomon doesn't visit often. There is a bulletin boards with fliers from the Ministry of Agriculture and NGO's advertising new, and sometimes old, techniques or tools. Solomon doesn't always trust this information and likes to ask what theories his friends have about it before he goes home.

This time when Solomon visited the cooperative they advertised a special star code to call and access information that would be useful to him. He had received SMSs from the cooperative before containing information about when crops would be collected or that the road into Tamale was bad. He also knew there was a way to post crops to the cooperative exchange via SMS. However, he personally didn't know any farmers who had used it and he heard rumors that it didn't always work if you typed information incorrectly.

The next week Solomon received another SMS reminding him about the star code and he decided to call it. The home screen was a menu with the following choices: 1. Weather,

2. Personal Account 3. Exchange Market 4. Community 5. Tips. After sending back 5 the next screen had the title of 6 new tips. The last line told him that 0 would go back and 9 would go to the forward in the list. The first tip was about about using herbicide for weed management - something the cooperative was always pushing and Solomon never had the savings to get. Though he wished to someday to use herbicide. There was an option to give feedback on the tip, so he reported that it wasn't very helpful since it was information he already knew. Solomon pressed 0 to go back and then went to the community forum. This was a message board with six different categories for posts. Solomon went into the section for cassava and found a list of posts by other farmers. He found out that he could make his own post by entering '+' and the '@' symbol would reply to someone. The last thing Solomon did was check the market place where he found a series of menus to help post crops to the exchange and see what the current prices were.

7.5.3 Data Collection

Josephine is a manager working with the National Immunization Program (NIP) in Kigoma, a district capital in Tanzania on Lake Tanganyika. She needs to receive information on a regular basis from health facilities, including weekly reports on vaccine stocks, as well as immediate notifications of stock outs and refrigerator failures. She is excited about the introduction of a USSD based system for data collection from rural facilities, as she has found using USSD to send money to relatives on mPesa both useful and reliable.

As part of the introduction of the new system, Josephine is leading training sessions throughout the district. Health workers from the rural areas have been attending these sessions where the NIP confirms that the current database is current. They use a web based management system to add contact phone numbers for each facility. When a number assigned with a health facility dials the USSD short code they get a simple menu system: 1. Submit Weekly Report, 2. Stock Out, 3. Fridge Status, and 4. Report Histories. The weekly report is a series of questions collecting the stock of six key vaccines as well as case reports for ten diseases. Josephine likes the fact that if the system detects format errors it replies with hints

asking for the data to be resubmitted.

When Josephine calls the USSD number, she gets a different menu system from the health facility version. This admin USSD interface allows her to enter and view data for any facility while on the road. She can also associate new phone numbers with a health facility and manage information on the refrigerators. This is useful when she visits health facilities, because sometimes the data recorded can be inaccurate and, even on her smartphone, she has found it impossible to access the web interface on poor data connections at rural facilities.

7.6 Discussion and Future Work

While we have identified several use cases for USSD that extend the capabilities of GSM universal applications, we acknowledge that USSD is only a piece of the broader ICT4D toolkit. Organizations must survey available options in the context of the requirements of their individual project. If primary audience for a service is non-literate, IVR may be the best solution. If a project needs only simple messaging and privacy is not a concern or if frequent data collection will be done, SMS is a feasible option. And when there are few primary users a custom device specific application may work. However, there are three situations where USSD enables services that otherwise could not be offered as universal GSM applications. (1) when sensitive text based information must be shared, (2) when data collection is infrequent and complex, (3) when complex user interactions must take place.

The feasibility of USSD for large scale applications such as mobile money and interacting with mobile carriers is well documented. The framework presented in this paper shows how USSD can also be used for smaller scale operations similar to IVR and SMS. This opens up a large new design spaces to explore. There is a need to do user studies to understand the limitation of PIN and username based systems, which although the easiest from a technical stand point may not be the most user friendly. Systems should be designed that push the boundaries of what text based USSD systems can do such as creating a digital marketplace or math tutor. The user experience for these systems should be better studied to understand how to scale communities and improve services.

The principal motivation for this work originated in discussions with the Kenyan Ministry of Health (MOH) around SMS applications for maternal health and HIV awareness. Concern was expressed about the safety and ethics of distributing medical information over SMS. USSD seemed to be a solution that offered more privacy, even on the least expensive phones, as well as improved user experience and expanded design space. We were pleasantly surprised to find already existing third party operators offer USSD services. We now are currently in the process of testing the feasibility and design of our maternal health USSD application and working with the MOH to establish guidelines around its use.

7.7 Conclusion

The recent emergence of third party USSD gateways has opened up the design space for M4D applications that can reach users on the most basic phones. Traditionally USSD has only been used for banking and cellular operations, but by examining the existing ICT4D literature we can identify how USSD will complement and extend current uses of SMS and IVR for data collection, messaging, information query, social networking and making complex transactions. We built a prototype USSD application using an new library that provides the basic building blocks for general USSD applications. We have identified when M4D projects should consider using USSD as a solution and our USSD library will well in creating innovative services aimed at marginalized users throughout the world.

Chapter 8

CONCLUSION

Digital technologies exist in a strange duality that is simultaneously intangible and yet very material. There is no way to *'hold'* an SMS message, and unlike a dose of a tuberculosis fighting antibiotic, it is impossible to put an SMS message on truck or observe a patient taking it. This immaterial nature of digital technologies makes them perplexing agents of change — how, one may ask, are pixels on screen or information stored on a hard drive going to *do* anything. However, while digital technologies might not do much, what little they do do, they do extremely well. They are so effective and efficient at storing, transferring, and processing information that they leapfrog other technologies existing on the fringes of electrification, running water, and strong road networks.

The aspiration of ICT4D is that this power of digital technologies can act as a catalyst for actual change. This dissertation is a deep dive into how one of the simplest and most ubiquitous of technologies – the SMS – can be that catalyst. Of course any individual SMS is tiny part of a much larger process aimed at connecting end users to health professionals and domain expertise. Ultimately, this has been the main objective throughout the design and deployment of the mWACH platform. The details of creating this semi-automated bi-directional messaging platform and evolution of projects in the mWACH suite was described in Chapter 3.

8.1 Contributions

In total five different maternal health focused studies were conducted on the platform in peri-urban and semi-rural Kenya. The contributions of this thesis extend beyond the platform development to take a highly quantitative view of mobile messaging in the Kenyan context. In Chapter 4, I present details on system engagement, messaging content and language use,

and semi-automated conversations for the first deployment of mWACH. mWACH v1 showed that not only would mothers respond to automated messages, they do so quickly and find utility in responses from study staff. In Chapter 5 I showed how the mWACH platform could be modified to facilitate a postpartum family planning conversation between the system, study staff, mothers, and their male partners. This work showed that while it is hard to get male partners to engage on family planning via SMS, when they do engage, it is on the same topics as the mothers. Furthermore, by adding the male partners, every couple dyad sent at least one message into the system – something not true of any of the other studies.

A substantial contribution of this work is the longitudinal analysis in Chapter 6. With over two years of messaging for the two-way arm in mWACH X, I was able to take a quantitative view of understanding messaging behavior. This work shows how a small minority of participants are responsible for the bulk of sustained engagement with the system. In addition while the majority of messages are received almost immediately after sending, after a year of messaging the bottom quartile of participants stop receiving all SMS messages. This analysis gives a detailed view of what a large scale SMS intervention should expect in terms of response rate, drop off, and conversation lengths.

Finally in Chapter 7 I describe the larger universal application design space for M4D with stateful and stateless versions of voice and text protocols. I show how the session based text USSD protocol is understudied in the M4D literature and how the emergence of third-party USSD gateways creating has reduced the barrier of entry to USSD projects. This work ends by describing what features could be included if SMS based platforms like mWACH platform were implemented with USSD.

8.2 Final Remarks

This dissertation provides a detailed look at the design, deployment, and evolution of a semi-automated bi-directional messaging platform with the goal of connecting end users to medical staff via basic mobile phones. The mWACH model of open ended questions encouraging conversation between patients and providers is very different from other mHealth

projects with either one-way push messaging or limited two-way responses. Analysis of system engagement and use shows that participants will engage with the system and evidence from our RCTs suggests that bi-directional messaging improves health outcomes over no messaging at all. However, work needs to be done on operationalizing the mWACH model and incorporating it into health systems.

The challenge for future health messaging projects is not understanding what works, but rather figuring how to scale while integrating with standard of care. As systems scale further research is needed into natural language processing techniques that triage and classify incoming messages, decentralized models that utilize community health works as points of contact for participants, and supporting the integration with health system data collection systems. The degrees of freedom for any mobile health project are extraordinarily large and many variables are arbitrary set. Isolating the individual effects like message frequency, duration, or content when evaluating a systematic intervention difficult. The contributions of this thesis will help researches, practitioners, and policy makers decide on the scope and properties when scaling and developing future health messaging platforms.

8.3 Acknowledgments

It is impossible to produce a dissertation of this kind without the the support, encouragement, collaboration and hard work of a large community. I am incredibly grateful to everyone who has helped me, not only along this grad-school journey but also on the path to get here.

First, I would like to thank the staff of UW Kenya and my collaborators at Global Health for making all of this possible. The team in Kenya consists of John Kinuthia, Daniel Matemo, Dyphna Mogaka, Lusi Osborn, Valari Kemuto and many more who have worked one or more of the the projects. Research in ICTD requires getting out of the lab and the work being done everyday in the field to enroll participants, respond to messages, transport blood samples and abstract clinic visit records is what has enabled all of this. At UW Global Health I am indebted to all the work managing projects and keeping the vision of mWACH going from Grace John-Stewart, Alison Drake, Jillian Pintye, and Wenwen Jiang and others.

Second, a big thanks to everyone on my committee for helping me get through to the finish – especially when the last mile of the marathon becomes the hardest. I am forever grateful for the guidance and support of Richard Anderson. From my very first day on UW campus his office has always been open as a source of ICTD and academic knowledge. Very little of mWACH would have happened without the countless hours of work Jennifer Unger and Keseht Ronen who’s passion, care, and energy for health equity will always inspire me. A special thanks to Brian DeRenzi who time and again has gone out his way to bring structure to the chaotic mess of ideas swirling in my brain.

Third, I want to express my thanks to the UW ICTD lab and CSE community for making my time here so special. Waylon Brunette and Nicki Dell, as senior grad students in CSE 606, and Neha Kumar as a post-doc in ICTD passed down wisdom on how to navigate grad school, do ICTD research, and enjoy my time at UW. I have had the pleasure of working on projects with Fahad Pervaiz, Aditya Vashistha, Matt Johnson, Ester Jang, Naveena Karusala, Sam Castle, Samia Ibtasam, Sam Sudar, Galen Weld, and Matt Ziegler. The work being done by everyone at UW ICTD is so inspiring and I look forward to seeing what everyone does in the future.

Lastly, beyond CSE I want to thank my family for cultivating my curiosity challenging me intellectually and teaching me to love and respect the world we live in. My grandparents taught me from an early age that hard work and community are what truly matter in life. My parents raised me in such a way that I not only had the ability to server four years in the Peace Corps but it seemed natural – I know many other people have not had that freedom. Finally, words can not express my appreciation and love for Megan who has always reminded me to step back from the computer and enjoy life – especially when it’s a peanut butter and jelly sandwich on the side of the trail.

BIBLIOGRAPHY

- [1] Now Book Railway Tickets on your Mobile without Internet Connection.
- [2] UNICEF's U-Report Social Platform Hits 1 Million Active Users.
- [3] Amit Agarwal, Bhumika Gupta, Gaurav Bhatt, and Ankush Mittal. Construction of a Semi-Automated Model for FAQ Retrieval via Short Message Service. In *Proceedings of the 7th Forum for Information Retrieval Evaluation, FIRE '15*, pages 35–38. ACM.
- [4] Sheetal K. Agarwal, Anupam Jain, Arun Kumar, and Nitendra Rajput. The World Wide Telecom Web browser. In *Proceedings of the First ACM Symposium on Computing for Development, ACM DEV '10*, pages 4:1–4:9. ACM.
- [5] Sheetal K Agarwal, Nitendra Rajput, John Canny, and Apala Lahiri Chavan. IUI4DR: Intelligent user interfaces for developing regions. In *Proceedings of the 13th International Conference on Intelligent User Interfaces, IUI '08*, pages 437–437. ACM.
- [6] Icek Ajzen. From intentions to actions: A theory of planned behavior. In *Action Control*, pages 11–39. Springer.
- [7] Icek Ajzen. The theory of planned behavior. 50(2):179–211.
- [8] Majid M. Altuwaijri, Abdulrahman M. Sughayr, Mohamed A. Hassan, and Fahad M. Alazwari. The effect of integrating short messaging services reminders with electronic medical records on non-attendance rates. 33(2):193–196.
- [9] Morgan G. Ames. Charismatic Technology. In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives, AA '15*, pages 109–120. Aarhus University Press.
- [10] Ruth E. Anderson, Waylon Brunette, Erica Johnson, Caitlin Lustig, Anthony Poon, Cynthia Putnam, Odina Salihbaeva, Beth E. Kolko, and Gaetano Borriello. Experiences with a transportation information system that uses only GPS and SMS. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development, ICTD '10*, pages 4:1–4:10. ACM.
- [11] Caroline Asiimwe, David Gelvin, Evan Lee, Yanis Ben Amor, Ebony Quinto, Charles Katureebe, Lakshmi Sundaram, David Bell, and Matt Berg. Use of an Innovative,

- Affordable, and Open-Source Short Message Service–Based Tool to Monitor Malaria in Remote Areas of Uganda. 85(1):26–33.
- [12] Ken Banks and Erik Hersman. FrontlineSMS and Ushahidi-a demo. In *Information and Communication Technologies and Development (ICTD), 2009 International Conference On*, pages 484–484. IEEE.
- [13] Ken Banks, Sean Martin McDonald, and Florence Scialom. Mobile technology and the last mile:“Reluctant innovation” and FrontlineSMS. 6(1):7–12.
- [14] Jim Barrington, Olympia Wereko-Brobby, Peter Ward, Winfred Mwafongo, and Seif Kungulwe. SMS for Life: A pilot project to improve anti-malarial drug supply management in rural Tanzania using standard technology. 9(1):1–9.
- [15] Agathe Battestini, Vidya Setlur, and Timothy Sohn. A Large Scale Study of Text-messaging Use. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services*, MobileHCI '10, pages 229–238. ACM.
- [16] Marshall H. Becker. *The Health Belief Model and Personal Health Behavior*, volume 2. Slack.
- [17] Matt Berg, James Wariero, and Vijay Modi. *Every Child Counts: The Use of SMS in Kenya to Support the Community Based Management of Acute Malnutrition and Malaria in Children under Five*. Columbia university. Earth institute. ChildCount: with UNICEF Innovation group.
- [18] Jean Joel R. Bigna, Jean Jacques N. Noubiap, Charles Kouanfack, Claudia S. Plottel, and Sinata Koulla-Shiro. Effect of mobile phone reminders on follow-up medical care of children exposed to or infected with HIV in Cameroon (MORE CARE): A multicentre, single-blind, factorial, randomised controlled trial. 14(7):600–608.
- [19] Eric Brewer, Michael Demmer, Bowei Du, Melissa Ho, Matthew Kam, Sergiu Nedeveschi, Joyojeet Pal, Rabin Patra, Sonesh Surana, and Kevin Fall. The Case for Technology in Developing Regions. 38(6):25–38.
- [20] Jenna Burrell and Elisa Oreglia. The myth of market price information: Mobile phones and the application of economic knowledge in ICTD. 44(2):271–292.
- [21] Jessica A. Bushar, Jodie Fishman, Danielle Garfinkel, and Amy Pirretti. Enrolling Underserved Women in mHealth Programs: Results From Text4baby Outreach Campaigns. 20(2):292–299.

- [22] Josip Car, Cindy Ng, Rifat Atun, and Alan Card. SMS text message healthcare appointment reminders in England. 31(3):216–219.
- [23] Cornelia Caragea, Nathan McNeese, Anuj Jaiswal, Greg Traylor, Hyun-Woo Kim, Prasenjit Mitra, Dinghao Wu, Andrea H. Tapia, Lee Giles, Bernard J. Jansen, et al. Classifying text messages for the haiti earthquake. In *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management (ISCRAM2011)*. Citeseer.
- [24] Victoria Carrington. Txting: The end of civilization (again)? 35(2):161–175.
- [25] Dipanjan Chakraborty, Indrani Medhi, Edward Cutrell, and William Thies. Man Versus Machine: Evaluating IVR Versus a Live Operator for Phone Surveys in India. In *Proceedings of the 3rd ACM Symposium on Computing for Development*, ACM DEV '13, pages 7:1–7:9. ACM.
- [26] Alison Chatfield, Gillian Javetski, and Neal Lesh. Commcare evidence base.
- [27] Jay Chen, Brendan Linn, and Lakshminarayanan Subramanian. SMS-based Contextual Web Search. In *Proceedings of the 1st ACM Workshop on Networking, Systems, and Applications for Mobile Handhelds*, MobiHeld '09, pages 19–24. ACM.
- [28] Jay Chen, Lakshmi Subramanian, and Eric Brewer. SMS-based Web Search for Low-end Mobile Devices. In *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, MobiCom '10, pages 125–136. ACM.
- [29] Zhou-wen Chen, Li-zheng Fang, Li-ying Chen, and Hong-lei Dai. Comparison of an SMS text messaging and phone reminder to improve attendance at a health promotion center: A randomized controlled trial. 9(1):34–38.
- [30] Arul Chib. The promise and peril of mHealth in developing countries. 1(1):69–75.
- [31] Arul Chib, Michelle Helena van Velthoven, and Josip Car. mHealth Adoption in Low-Resource Environments: A Review of the Use of Mobile Healthcare in Developing Countries. 20(1):4–34.
- [32] Arul Chib, Holley Wilkin, and Bas Hoefman. Vulnerabilities in mHealth implementation: A Ugandan HIV/AIDS SMS campaign. 20:26–32.
- [33] Arul Chib, Holley Wilkin, Leow Xue Ling, Bas Hoefman, and Hajo Van Biejma. You Have an Important Message! Evaluating the Effectiveness of a Text Message HIV/AIDS Campaign in Northwest Uganda. 17:146–157.

- [34] Clara K. Chow, Julie Redfern, Graham S. Hillis, Jay Thakkar, Karla Santo, Maree L. Hackett, Stephen Jan, Nicholas Graves, Laura de Keizer, Tony Barry, Severine Bompont, Sandrine Stepien, Robyn Whittaker, Anthony Rodgers, and Aravinda Thiagalingam. Effect of Lifestyle-Focused Text Messaging on Risk Factor Modification in Patients With Coronary Heart Disease: A Randomized Clinical Trial. 314(12):1255–1263.
- [35] Carl Collins, Amy Grude, Matthew Scholl, and Robert Thompson. Txt Bus: Wait Time Information on Demand. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '07, pages 2049–2054. ACM.
- [36] Communication Commission of Kenya. Quarterly Sector Statistics Report (Oct-Dec 2013).
- [37] Jamaica Corker. Ligne Verte” toll-free hotline: Using cell phones to increase access to family planning information in the Democratic Republic of Congo. 4:23–37.
- [38] Angela Crandall. Kenyan farmers’ use of cell phones: Calling preferred over SMS. pages 119–129.
- [39] Angela Crandall, Albert Otieno, Leonida Mutuku, Jessica Colaço, Jasper Grosskurth, and Peter Otieno. *Mobile Phone Usage at the Kenyan Base of the Pyramid*.
- [40] Cristian Pop-Eleches, Harsha Thirumurthy, James P. Habyarimana, Joshua G. Zivin, Markus P. Goldstein, Damien de Walque, Leslie MacKeen, Jessica Haberer, Sylvester Kimaiyo, John Sidle, Duncan Ngare, and David R. Bangsberg. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: A randomized controlled trial of text message reminders. 25(6):825–834.
- [41] Andrew Cross, Nakull Gupta, Brandon Liu, Vineet Nair, Abhishek Kumar, Reena Kuttan, Priyanka Ivatury, Amy Chen, Kshama Lakshman, Rashmi Rodrigues, George D’Souza, Deepti Chittamuru, Raghuram Rao, Kiran Rade, Bhavin Vadera, Daksha Shah, Vinod Choudhary, Vineet Chadha, Amar Shah, Sameer Kumta, Puneet Dewan, Bruce Thomas, and William Thies. 99dots: A low-cost approach to monitoring and improving medication adherence. In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development*, ICTD '19, pages 15:1–15:12, New York, NY, USA, 2019. ACM.
- [42] Trisha M Crutchfield and Christine E Kistler. Getting patients in the door: Medical appointment reminder preferences. 11:141–150.
- [43] Sebastien Cuendet, Indrani Medhi, Kalika Bali, and Edward Cutrell. VideoKheti: Making video content accessible to low-literate and novice users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2833–2842. ACM.

- [44] Thiago Martini Da Costa, Paulo Lísias Salomão, Amilton Souza Martha, Ivan Torres Pisa, and Daniel Sigulem. The impact of short message service text messages sent as appointment reminders to patients' cell phones at outpatient clinics in Sao Paulo, Brazil. 79(1):65–70.
- [45] Catalina M. Danis, Jason B. Ellis, Wendy A. Kellogg, Hajo van Beijma, Bas Hoefman, Steven D. Daniels, and Jan-Willem Loggers. Mobile phones for health education in the developing world: SMS as a user interface. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10, pages 13:1–13:9. ACM.
- [46] Fred D. Davis, Richard P. Bagozzi, and Paul R. Warshaw. User acceptance of computer technology: A comparison of two theoretical models. 35(8):982–1003.
- [47] Dedoose. Dedoose. [Online; accessed 28-Feb-2017].
- [48] Nicola Dell and Neha Kumar. The ins and outs of HCI for development. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 2220–2232. ACM.
- [49] Nicola Dell, Vidya Vaidyanathan, Indrani Medhi, Edward Cutrell, and William Thies. "Yours is better!": Participant response bias in HCI. In *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems*, CHI '12, pages 1321–1330. ACM.
- [50] Melissa Densmore. Claim Mobile: When to Fail a Technology. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 1833–1842. ACM.
- [51] Melissa Densmore. Experiences with Bulk SMS for Health Financing in Uganda. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '12, pages 383–398. ACM.
- [52] Melissa Densmore, Ben Bellows, John Chuang, and Eric Brewer. The Evolving Braid: How an Organization in Uganda Achieved Reliable Communications. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers - Volume 1*, ICTD '13, pages 257–266. ACM.
- [53] Brian DeRenzi, Leah Findlater, Jonathan Payne, Benjamin Birnbaum, Joachim Mangilima, Tapan Parikh, Gaetano Borriello, and Neal Lesh. Improving Community Health Worker Performance Through Automated SMS. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, ICTD '12, pages 25–34. ACM.

- [54] Brian DeRenzi, Neal Lesh, Tapan Parikh, Clayton Sims, Werner Maokla, Mwajuma Chemba, Yuna Hamisi, David S hellenberg, Marc Mitchell, and Gaetano Borriello. E-imci: Improving Pediatric Health Care in Low-income Countries. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, pages 753–762. ACM.
- [55] M. Beatrice Dias, Ermine A. Teves, M. Freddie Dias, Daniel Nuffer, Hatem Alismail, Rotimi Abimbola, M. Bernardine Dias, Anthony Velázquez, Sarah Belousov, and Bradley Hall. Using Mobile Phones and Open Source Tools to Empower Social Workers in Tanzania. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*, ICTD '10, pages 8:1–8:9. ACM.
- [56] Pankaj Doke and Anirudha Joshi. Mobile Phone Usage by Low Literate Users. In *Proceedings of the 7th International Conference on HCI, IndiaHCI 2015*, IndiaHCI'15, pages 10–18. ACM.
- [57] Jonathan Donner. Research approaches to mobile use in the developing world: A review of the literature. 24(3):140–159.
- [58] Jonathan Donner. The Rules of Beeping: Exchanging Messages Via Intentional “Missed Calls” on Mobile Phones. 13(1):1–22.
- [59] Alison L. Drake, Jennifer A. Unger, Keshet Ronen, Daniel Matemo, Trevor Perrier, Brian DeRenzi, Barbra A. Richardson, John Kinuthia, and Grace John-Stewart. Evaluation of mHealth strategies to optimize adherence and efficacy of Option B+ prevention of mother-to-child HIV transmission: Rationale, design and methods of a 3-armed randomized controlled trial. 57:44–50.
- [60] Nathan Eagle. Txteagle: Mobile crowdsourcing. In *International Conference on Internationalization, Design and Global Development*, pages 447–456. Springer.
- [61] Leah Findlater, Ravin Balakrishnan, and Kentaro Toyama. Comparing Semiliterate and Illiterate Users’ Ability to Transition from Audio+Text to Text-only Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 1751–1760. ACM.
- [62] Brianna S. Fjeldsoe, Alison L. Marshall, and Yvette D. Miller. Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. 36(2):165–173.
- [63] B. J. Fogg. *Persuasive Technology: Using Computers to Change What We Think and Do*. 2002.

- [64] BJ Fogg. A Behavior Model for Persuasive Design. In *Proceedings of the 4th International Conference on Persuasive Technology*, Persuasive '09, pages 40:1–40:7. ACM.
- [65] BJ Fogg and Enrique Allen. 10 Uses of Texting to Improve Health. In *Proceedings of the 4th International Conference on Persuasive Technology*, Persuasive '09, pages 38:1–38:6. ACM.
- [66] V. L. Franklin, A. Waller, C. Pagliari, and S. A. Greene. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. *23(12):1332–1338*.
- [67] Vanessa Frias-Martinez and Jesus Virseda. On the Relationship Between Socio-economic Factors and Cell Phone Usage. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, ICTD '12, pages 76–84. ACM.
- [68] Elsa Friscira, Hendrik Knoche, and Jeffrey Huang. Getting in touch with text: Designing a mobile phone application for illiterate users to harness SMS. In *Proceedings of the 2nd ACM Symposium on Computing for Development*, ACM DEV '12, pages 5:1–5:10. ACM.
- [69] FrontlineSMS. Frontline Cloud. [Online; accessed 28-Feb-2017].
- [70] Gareth V. Furber, Ann E. Crago, Kevin Meehan, Tom D. Sheppard, Ken Hooper, Dorothy T. Abbot, Stephen Allison, and Clive Skene. How Adolescents Use SMS (Short Message Service) to Micro-Coordinate Contact With Youth Mental Health Outreach Services. *48(1):113–115*.
- [71] Genevieve Gebhart. Zero-rating in Emerging Mobile Markets: Free Basics and Wikipedia Zero in Ghana. In *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development*, ICTD '16, pages 16:1–16:9. ACM.
- [72] Nathan Georgette, Mark J. Siedner, Carter R. Petty, Brian C. Zanoni, Stephen Carpenter, and Jessica E. Haberer. Impact of a clinical program using weekly Short Message Service (SMS) on antiretroviral therapy adherence support in South Africa: A retrospective cohort study. *17:18*.
- [73] S. Germann, A. Jabry, J. Njogu, and R. Osumba. The Illness of “Pilotitis” In mHealth—Early Lessons from the KimMNCHip Partnerships in Kenya. In *Geneva Health Forum, Geneva, Switzerland*.

- [74] Karen Glanz, Barbara K. Rimer, and Kasisomayajula Viswanath. *Health Behavior and Health Education: Theory, Research, and Practice*. John Wiley & Sons.
- [75] Edward C. Green and Elaine Murphy. Health belief model.
- [76] Robert Gross, Justin Ritz, Michael D. Hughes, Robert Salata, Peter Mugenyi, Evelyn Hogg, Linda Wieclaw, Catherine Godfrey, Carole L. Wallis, John W. Mellors, Victor O. Mudhune, Sharlaa Badal-Faesen, Beatriz Grinsztejn, and Ann C. Collier. Two-way mobile phone intervention compared with standard-of-care adherence support after second-line antiretroviral therapy failure: A multinational, randomised controlled trial. 1(1):e26–e34.
- [77] World Bank Group. *World Development Report 2016: Digital Dividends*. World Bank Publications.
- [78] Aditi Sharma Grover, Karen Calteaux, Etienne Barnard, and Gerhard van Huyssteen. A Voice Service for User Feedback on School Meals. In *Proceedings of the 2Nd ACM Symposium on Computing for Development*, ACM DEV '12, pages 13:1–13:10. ACM.
- [79] Tilly A. Gurman, Sara E. Rubin, and Amira A. Roess. Effectiveness of mHealth Behavior Change Communication Interventions in Developing Countries: A Systematic Review of the Literature. 17:82–104.
- [80] Jessica E. Haberer, Angella Musiimenta, Esther C. Atukunda, Nicholas Musinguzi, Monique A. Wyatt, Norma C. Ware, and David R. Bangsberg. Short message service (SMS) reminders and real-time adherence monitoring improve antiretroviral therapy adherence in rural Uganda. 30(8):1295–1299.
- [81] Margaret Hagan, Nan Zhang, and Joseph 'Jofish' Kaye. Safe Mathare: A Mobile System for Women's Safe Commutes in the Slums. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services Companion*, MobileHCI '12, pages 47–52. ACM.
- [82] David A. Hanauer, Katherine Wentzell, Nikki Laffel, and Lori M. Laffel. Computerized Automated Reminder Diabetes System (CARDS): E-Mail and SMS Cell Phone Text Messaging Reminders to Support Diabetes Management. 11(2):99–106.
- [83] Elizabeth K. Harrington, Alison L. Drake, Daniel Matemo, Keshet Ronen, Alfred O. Osofi, Grace John-Stewart, John Kinuthia, and Jennifer A. Unger. An mHealth SMS intervention on Postpartum Contraceptive Use Among Women and Couples in Kenya: A Randomized Controlled Trial. 109(6):934–941.

- [84] Carl Hartung, Adam Lerer, Yaw Anokwa, Clint Tseng, Waylon Brunette, and Gaetano Borriello. Open Data Kit: Tools to Build Information Services for Developing Regions. In *Proceedings of the 4th ACM/IEEE International Conference on Information and Communication Technologies and Development*, ICTD '10, pages 18:1–18:12. ACM.
- [85] Katharine J. Head, Seth M. Noar, Nicholas T. Iannarino, and Nancy Grant Harrington. Efficacy of text messaging-based interventions for health promotion: A meta-analysis. 97:41–48.
- [86] Kurtis Heimerl, R. J. Honicky, Eric Brewer, and Tapan Parikh. Message phone: A user study and analysis of asynchronous messaging in rural uganda. In *SOSP Workshop on Networked Systems for Developing Regions (NSDR)*, pages 15–18.
- [87] Johan Hellström and Anna Karefelt. Participation Through Mobile Phones: A Study of SMS Use During the Ugandan General Elections 2011. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, ICTD '12, pages 249–258. ACM.
- [88] Johan Hellström and Per-Einar Tröften. *The Innovative Use of Mobile Applications in East Africa*. Swedish international development cooperation agency (Sida).
- [89] Hoefman BAS1 Apunyu BONNY. Using SMS for HIV/AIDS education and to expand the use of HIV testing and counselling services at the AIDS Information Centre (AIC) Uganda. page 40.
- [90] Muhammad Imran, Patrick Meier, Carlos Castillo, Andre Lesa, and Manuel Garcia Heranz. Enabling Digital Health by Automatic Classification of Short Messages. In *Proceedings of the 6th International Conference on Digital Health Conference*, DH '16, pages 61–65. ACM.
- [91] Ann Irvine, Jonathan Weese, and Chris Callison-Burch. Processing Informal, Romanized Pakistani Text Messages. In *Proceedings of the Second Workshop on Language in Social Media*, LSM '12, pages 75–78. Association for Computational Linguistics.
- [92] William Jack and Tavneet Suri. Mobile Money: The Economics of M-PESA.
- [93] Larissa Jennings and Laina Gagliardi. Influence of mhealth interventions on gender relations in developing countries: A systematic literature review. 12:85.
- [94] Robert Jensen. The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector *. 122(3):879–924.

- [95] Douglas Johnson, Randall Juras, Pamela Riley, Minki Chatterji, Phoebe Sloane, Soon Kyu Choi, and Ben Johns. A randomized controlled trial of the impact of a family planning mHealth service on knowledge and use of contraception. 95(1):90–97.
- [96] Caroline OH Jones, Beatrice Wasunna, Raymond Sudoi, Sophie Githinji, Robert W. Snow, and Dejan Zurovac. “Even if you know everything you can forget”: Health worker perceptions of mobile phone text-messaging to improve malaria case-management in Kenya. 7(6):e38636.
- [97] Anirudha Joshi. Mobile phones and economic sustainability: Perspectives from India. In *Proceedings of the First International Conference on Expressive Interactions for Sustainability and Empowerment*, pages 2–2. BCS Learning & Development Ltd.
- [98] Anirudha Joshi, Mandar Rane, Debjani Roy, Nagraj Emmadi, Padma Srinivasan, N. Kumarasamy, Sanjay Pujari, Davidson Solomon, Rashmi Rodrigues, D.G. Saple, Kamalika Sen, Els Veldeman, and Romain Rutten. Supporting Treatment of People Living with HIV / AIDS in Resource Limited Settings with IVRs. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems, CHI '14*, pages 1595–1604. ACM.
- [99] R. J. Kaitho, J. W. Stuth, and A. A. Jama. Application of Information and Communication Technology in Livestock Marketing in the Pastoral Areas of Eastern Africa.
- [100] Juhee Kang and Moutusy Maity. Texting Among the Bottom of the Pyramid: Facilitators and Barriers to SMSs Use Among the Low-Income Mobile Users in Asia.
- [101] Simon Kemp. *Digital Around The World In 2019: Essential Insights Into How People Around The World Use The Internet Mobile Devices, Social Media, and E-Commerce*. Hootsuite.
- [102] S. Kinkade, K. Verclas, and K. Toyama. *Reflections on MobileActive 2008 and the M4D Landscape Wireless Technology for Social Change*.
- [103] Risa Kitagawa. Texting and Sexual Health: Experimental Evidence from an Information Intervention in Kenya. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development, ICTD '15*, pages 18:1–18:10. ACM.
- [104] Govind Kothari, Sumit Negi, Tanveer A. Faruquie, Venkatesan T. Chakaravarthy, and L. Venkata Subramaniam. SMS Based Interface for FAQ Retrieval. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL '09, pages 852–860. Association for Computational Linguistics.

- [105] Lucy I. Kululanga, Johanne Sundby, Lucy I. Kululanga, Address Malata, and Ellen Chirwa. Male involvement in maternity health care in Malawi : Original research article. 16(1):145–157.
- [106] Setor Kunutsor, John Walley, Elly Katabira, Simon Muchuro, Hudson Balidawa, Elizabeth Namagala, and Eric Ikoona. Using Mobile Phones to Improve Clinic Attendance Amongst an Antiretroviral Treatment Cohort in Rural Uganda: A Cross-sectional and Prospective Study. 14(6):1347–1352.
- [107] Hmwe Hmwe Kyu, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, et al. Global, regional, and national disability-adjusted life-years (dalys) for 359 diseases and injuries and healthy life expectancy (hale) for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1859–1922, 2018.
- [108] Amnesty E. LeFevre, Pierre Dane, Charles J. Copley, Cara Pienaar, Annie Neo Parsons, Matt Engelhard, David Woods, Marcha Bekker, Peter Benjamin, Yogan Pillay, Peter Barron, Christopher John Seebregts, and Diwakar Mohan. Unpacking the performance of a mobile health information messaging program for mothers (MomConnect) in South Africa: Evidence on program reach and messaging exposure. 3:e000583.
- [109] Kelly L. L’Engle, Heather L. Vahdat, Elizabeth Ndakidemi, Christine Lasway, and Trinity Zan. Evaluating feasibility, reach and potential impact of a text message family planning information service in Tanzania. 87(2):251–256.
- [110] Natalie Leon, Rebecca Surender, Kirsty Bobrow, Jocelyn Muller, and Andrew Farmer. Improving treatment adherence for blood pressure lowering via mobile phone SMS-messages in South Africa: A qualitative evaluation of the SMS-text Adherence Support (StAR) trial. 16:80.
- [111] Adam Lerer, Molly Ward, and Saman Amarasinghe. Evaluation of IVR data collection UIs for untrained rural users. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV ’10, pages 2:1–2:8. ACM.
- [112] Richard T. Lester. Ask, Don’t Tell — Mobile Phones to Improve HIV Care. 369(19):1867–1868.
- [113] R.T Lester, P Ritvo, E.J Mills, A Kariri, S Karanja, M.H Chung, W Jack, J Habyarimana, M Sadatsafavi, M Najafzadeh, C.A Marra, B Estambale, E Ngugi, T.B Ball, L Thabane, L.J Gelmon, J Kimani, M Ackers, and F.A Plummer. Effects of a mobile

- phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): A randomised trial. 376(9755):1838–1845.
- [114] Sebastian Linnemayr, Haijing Huang, Jill Luoto, Andrew Kambugu, Harsha Thirumurthy, Jessica E. Haberer, Glenn Wagner, and Barbara Mukasa. Text messaging for improving antiretroviral therapy adherence: No effects after 1 year in a randomized controlled trial among adolescents and young adults. 107(12):1944–1950.
- [115] R. Manguni, M.L. Navarro, K. Rosario, and C.A. Festin. chitSMS: Community Health Information Tracking System Using Short Message Service. In *2010 3rd International Conference on Human-Centric Computing (HumanCom)*, pages 1–6.
- [116] Lawrence Mbuagbaw, Lehana Thabane, and Pierre Ongolo-Zogo. Opening communication channels with people living with HIV using mobile phone text messaging: Insights from the CAMPS trial. 6:131.
- [117] Lawrence Mbuagbaw, Lehana Thabane, Pierre Ongolo-Zogo, Richard T. Lester, Edward J. Mills, Marek Smieja, Lisa Dolovich, and Charles Kouanfack. The Cameroon Mobile Phone SMS (CAMPS) Trial: A Randomized Trial of Text Messaging versus Usual Care for Adherence to Antiretroviral Therapy. 7(12):e46909.
- [118] Lawrence Mbuagbaw, Mia L. van der Kop, Richard T. Lester, Harsha Thirumurthy, Cristian Pop-Eleches, Chenglin Ye, Marek Smieja, Lisa Dolovich, Edward J. Mills, and Lehana Thabane. Mobile phone text messages for improving adherence to antiretroviral therapy (ART): An individual patient data meta-analysis of randomised trials. 3(12):e003950.
- [119] Susan McRoy, Vishnuvardhan Vaidhynathan, Amy May, and Hayeon Song. An Open Architecture for Messaging-based Consumer-health Question-answering. In *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium, IHI '12*, pages 761–766. ACM.
- [120] Indrani Medhi, S.N. Nagasena Gautama, and Kentaro Toyama. A Comparison of Mobile Money-transfer UIs for Non-literate and Semi-literate Users. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09*, pages 1741–1750. ACM.
- [121] Indrani Medhi, Somani Patnaik, Emma Brunskill, S.N. Nagasena Gautama, William Thies, and Kentaro Toyama. Designing Mobile Interfaces for Novice and Low-literacy Users. 18(1):2:1–2:28.

- [122] Indrani Medhi Thies. User interface design for low-literate and novice users: Past, present and future. 8(1):1–72.
- [123] Jamison Merrill, Rebecca Hershow, Katherine Gannett, and Chris Barkley. Pretesting an mHealth Intervention for At-risk Adolescent Girls in Soweto, South Africa: Studying the Additive Effects of SMSs on Improving Sexual Reproductive Health & Rights Outcomes. In *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development: Notes - Volume 2*, ICTD '13, pages 96–99. ACM.
- [124] Matt Morio, Hallie Goertz, Brian Taliesin, and Kate Wilson. mHealth Mobile Messaging Toolkit: Considerations When Selecting a Mobile Messaging Platform Vendor. [Online; accessed 28-Feb-2017].
- [125] Nathan Morrow, Nancy Mock, Adam Papendieck, and Nicholas Kocmich. Independent evaluation of the Ushahidi Haiti project. 8:2011.
- [126] Preeti Mudliar, Jonathan Donner, and William Thies. Emergent Practices Around CGNet Swara, Voice Forum for Citizen Journalism in Rural India. In *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development*, ICTD '12, pages 159–168. ACM.
- [127] Preeti Mudliar and Joyojeet Pal. ICTD in the Popular Press: Media Discourse Around Aakash, the 'World's Cheapest Tablet'. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers - Volume 1*, ICTD '13, pages 43–54. ACM.
- [128] National Council For Population and Development. Sessional Paper No. 3 of 2012 on Population Policy for National Development.
- [129] Nexmo. Nexmo. [Online; accessed 28-Feb-2017].
- [130] Fidele Ngabo, Judith Nguimfack, Friday Nwaigwe, Catherine Mugeni, Denis Muhoza, David R Wilson, John Kalach, Richard Gakuba, Corrine Karema, and Agnes Binagwaho. Designing and Implementing an Innovative SMS-based alert system (RapidSMS-MCH) to monitor pregnancy and reduce maternal and child deaths in Rwanda. 13.
- [131] José António Nhavoto, Åke Grönlund, and Walter Ponce Chaquilla. SMSaúde: Design, Development, and Implementation of a Remote/Mobile Patient Management System to Improve Retention in Care for HIV/AIDS and Tuberculosis Patients. 3(1).

- [132] Daniel Ninsiima. "Buuza Omulimisa" (Ask the Extension Officer): Text Messaging for Low Literate Farming Communities in Rural Uganda. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*, ICTD '15, pages 54:1–54:4. ACM.
- [133] Daniel Ninsiima. Factors affecting adoption of an information communications technology system for agriculture in Uganda.
- [134] Simonette Alzona Nisperos and Yoshiyori Urano. SMS text messaging: A ubiquitous tool for cheaper medicine in the Philippines. In *13th International Conference on Advanced Communication Technology (ICACT2011)*, pages 1415–1419. IEEE.
- [135] Thomas A. Odeny, Robert C. Bailey, Elizabeth A. Bukusi, Jane M. Simoni, Kenneth A. Tapia, Krista Yuhas, King K. Holmes, and R. Scott McClelland. Effect of Text Messaging to Deter Early Resumption of Sexual Activity After Male Circumcision for HIV Prevention: A Randomized Controlled Trial. 65(2):e50–e57.
- [136] Thomas A. Odeny, Elizabeth A. Bukusi, Craig R. Cohen, Krista Yuhas, Carol S. Camlin, and R. Scott McClelland. Texting improves testing: A randomized trial of two-way SMS to increase postpartum prevention of mother-to-child transmission retention and infant HIV testing. 28(15):2307–2312.
- [137] Billy Odero, Brian Omwenga, Mokeira Masita-Mwangi, Pauline Githinji, and Jonathan Ledlie. Tangaza: Frugal group messaging through speech and text. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10, pages 1:1–1:10. ACM.
- [138] Erick Oduor, Carman Neustaedter, Tejinder K. Judge, Kate Hennessy, Carolyn Pang, and Serena Hillman. How Technology Supports Family Communication in Rural, Suburban, and Urban Kenya. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, CHI '14, pages 2705–2714. ACM.
- [139] Communications Authority of Kenya. Annual Report for the Financial Year 2015-2016.
- [140] Kenya National Bureau of Statistics, Ministry of Health/Kenya, National AIDS Control Council/Kenya, Kenya Medical Research Institute, and National Council for Population and Development/Kenya. *Kenya Demographic and Health Survey 2014*.
- [141] Ory Okolloh. Ushahidi, or 'testimony': Web 2.0 tools for crowdsourcing crisis information. 59(1):65–70.

- [142] Peter O. Ouma, Anna M. van Eijk, Mary J. Hamel, Evallyne S. Sikuku, Frank O. Odhiambo, Kaendi M. Munguti, John G. Ayisi, Sara B. Crawford, Piet A. Kager, and Laurence Slutsker. Antenatal and delivery care in rural western Kenya: The effect of training health care workers to provide" focused antenatal care". 7(1):1.
- [143] Niranjana Pai, Pradnya Supe, Shailesh Kore, Y. S. Nandanwar, Aparna Hegde, Edward Cutrell, and William Thies. Using Automated Voice Calls to Improve Adherence to Iron Supplements During Pregnancy: A Pilot Study. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers - Volume 1*, ICTD '13, pages 153–163. ACM.
- [144] Carolyn E. Pang, Carman Neustaedter, Bernhard E. Riecke, Erick Oduor, and Serena Hillman. Technology preferences and routines for sharing health information during the treatment of a chronic illness. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1759–1768. ACM.
- [145] T. S. Parikh. Using mobile phones for secure, distributed document processing in the developing world. 4(2):74–81.
- [146] Ruth M. Parker, Elena Dmitrieva, Sergei Frolov, and Julie A. Gazmararian. Text4baby in the United States and Russia: An Opportunity for Understanding How mHealth Affects Maternal and Child Health. 17:30–36.
- [147] Neil Patel, Sheetal Agarwal, Nitendra Rajput, Amit Nanavati, Paresh Dave, and Tapan S. Parikh. A Comparative Study of Speech and Dialed Input Voice Interfaces in Rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 51–54. ACM.
- [148] Neil Patel, Deepti Chittamuru, Anupam Jain, Paresh Dave, and Tapan S. Parikh. Avaaj Otalo: A Field Study of an Interactive Voice Forum for Small Farmers in Rural India. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 733–742. ACM.
- [149] Somani Patnaik, Emma Brunskill, and William Thies. Evaluating the Accuracy of Data Collection on Mobile Phones: A Study of Forms, Sms, and Voice. In *Proceedings of the 3rd International Conference on Information and Communication Technologies and Development*, ICTD'09, pages 74–84. IEEE Press.
- [150] Trevor Perrier, Nicola Dell, Brian DeRenzi, Richard Anderson, John Kinuthia, Jennifer Unger, and Grace John-Stewart. Engaging Pregnant Women in Kenya with a Hybrid Computer-Human SMS Communication System. In *Proceedings of the 33rd Annual*

- ACM Conference on Human Factors in Computing Systems, CHI '15*, pages 1429–1438. ACM.
- [151] Trevor Perrier, Brian DeRenzi, and Richard Anderson. Usd: The third universal app. In *Proceedings of the 2015 Annual Symposium on Computing for Development*, pages 13–21. ACM.
- [152] Trevor Perrier, Elizabeth K. Harrington, Keshet Ronen, Daniel Matemo, John Kinuthia, Grace John-Stewart, Richard Anderson, and Jennifer A. Unger. Male Partner Engagement in Family Planning SMS Conversations at Kenyan Health Clinics. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, COMPASS '18*, pages 3:1–3:11. ACM.
- [153] Fahad Pervaiz, Trevor Perrier, Sompasong Phongphila, and Richard Anderson. User errors in SMS based reporting systems. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*, page 55. ACM.
- [154] Angela Pfammatter, Bonnie Spring, Nalini Saligram, Raj Davé, Arun Gowda, Linelle Blais, Monika Arora, Harish Ranjani, Om Ganda, Donald Hedeker, Sethu Reddy, and Sandhya Ramalingam. mHealth Intervention to Improve Diabetes Risk Behaviors in India: A Prospective, Parallel Group Cohort Study. 18(8).
- [155] Jillian Pintye, John Kinuthia, D. Allen Roberts, Anjuli D. Wagner, Kenneth Mugwanya, Felix Abuna, Harison Lagat, George Owiti, Carol E. Levin, and Ruanne V. Barnabas. Brief Report: Integration of PrEP Services Into Routine Antenatal and Postnatal Care Experiences From an Implementation Program in Western Kenya. 79(5):590–595.
- [156] Praekelt.org. Praekelt.org Pilots Whatsapp for Social Impact on Behalf of MomConnect.
- [157] Coimbatore K. Prahalad. *The Fortune at the Bottom of the Pyramid, Revised and Updated 5th Anniversary Edition: Eradicating Poverty through Profits*. FT Press.
- [158] Purity Mwendwa. Assessing the fit of RapidSMS for maternal and new-born health: Perspectives of community health workers in rural Rwanda. 26(1):38–51.
- [159] Nancy Puttkammer, Steven Zeliadt, Jean Gabriel Balan, Janet Baseman, Rodney Destine, Jean Wysler Domercant, Garilus France, Nathaelf Hyppolite, Valerie Pelletier, and Nernst Atwood Raphael. Development of an electronic medical record based alert for risk of HIV treatment failure in a low-resource setting. 9(11):e112261.

- [160] Julia R. G. Raifman, Heather E. Lanthorn, Slawa Rokicki, and Günther Fink. The Impact of Text Message Reminders on Adherence to Antimalarial Treatment in Northern Ghana: A Randomized Trial. 9(10):e109032.
- [161] K. Ramamritham, A. Bahuman, S. Sahni, M. Baru, C. Bahuman, A. Chandran, and M. Joshi. The aAQUA Approach: Innovative Web 2.0 Tools for Developing Countries. 12(2):62–70.
- [162] Arun Ramanujapuram and Anup Akkihal. Improving Performance of Rural Supply Chains Using Mobile Phones: Reducing Information Asymmetry to Improve Stock Availability in Low-resource Environments. In *Proceedings of the Fifth ACM Symposium on Computing for Development*, ACM DEV-5 '14, pages 11–20. ACM.
- [163] Pranav Ramkrishnan, Aditya Vashistha, Edward Cutrell, and William Thies. DocTalk: Extending Doctors' Visits with Personalized Voice Messages. In *Proceedings of the 3rd ACM Symposium on Computing for Development*, ACM DEV '13, pages 40:1–40:2. ACM.
- [164] Aadia I. Rana, Jacob J. van den Berg, Eric Lamy, and Curt G. Beckwith. Using a Mobile Health Intervention to Support HIV Treatment Adherence and Retention Among Patients at Risk for Disengaging with Care. 30(4):178–184.
- [165] Yashodhara Rana, Jessica Haberer, Haijing Huang, Andrew Kambugu, Barbara Mukasa, Harsha Thirumurthy, Peter Wabukala, Glenn J. Wagner, and Sebastian Linnemayr. Short Message Service (SMS)-Based Intervention to Improve Treatment Adherence among HIV-Positive Youth in Uganda: Focus Group Findings. 10(4):e0125187.
- [166] RapidPro. RapidPro. [Online; accessed 28-Feb-2017].
- [167] Martin Ravallion. How long will it take to lift one billion people out of poverty? 28(2):139–158.
- [168] Kyle Rector and Joshua Hailpern. MinEMail: SMS Alert System for Managing Critical Emails. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 783–792. ACM.
- [169] A. Rodgers, T. Corbett, D. Bramley, T. Riddell, M. Wills, R.-B. Lin, and M. Jones. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. 14(4):255–261.
- [170] Keshet Ronen, Jennifer A. Unger, Alison L. Drake, Trevor Perrier, Pamela Akinyi, Lusi Osborn, Daniel Matemo, Gabrielle O'Malley, John Kinuthia, and Grace John-Stewart.

- SMS messaging to improve ART adherence: Perspectives of pregnant HIV-infected women in Kenya on HIV-related message content. 30(4):500–505.
- [171] S Lund. Mobile phones as a health communication tool to improve skilled attendance at delivery in Zanzibar: A cluster-randomised controlled trial. 119(10):1256–1264.
- [172] Ricardo Sabates, J Westbrook, K Akyeampong, and Frances Hunt. School drop out: Patterns, causes, changes and policies.
- [173] Nithya Sambasivan, Ed Cutrell, Kentaro Toyama, and Bonnie Nardi. Intermediated Technology Use in Developing Communities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2583–2592. ACM.
- [174] Nithya Sambasivan, Paul Lee, Greg Hecht, Paul M. Aoki, Maria-Ines Carrera, Jenny Chen, David Pablo Cohn, Pete Kruskall, Everett Wetchler, Michael Youssefmir, and Astrid Twenebowa Larssen. Chale, How Much It Cost to Browse?: Results from a Mobile Data Price Transparency Trial in Ghana. In *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development: Full Papers - Volume 1*, ICTD '13, pages 13–23. ACM.
- [175] Rudy Schusteritsch, Shailendra Rao, and Kerry Rodden. Mobile Search with Text Messages: Designing the User Experience for Google SMS. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '05, pages 1777–1780. ACM.
- [176] Christopher Seebregts, Peter Barron, Gaurang Tanna, Peter Benjamin, and Thomas Fogwill. MomConnect: An exemplar implementation of the Health Normative Standards Framework in South Africa. pages 125–135.
- [177] Jennifer R. Shapiro, Tina Koro, Neal Doran, Sheri Thompson, James F. Sallis, Karen Calfas, and Kevin Patrick. Text4Diet: A randomized controlled study using text messaging for weight loss behaviors. 55(5):412–417.
- [178] Ernesto Sirolli. *Ripples from the Zambezi: Passion, Entrepreneurship, and the Rebirth of Local Economies*. New Society Publishers.
- [179] Kirsten Smillie, Natasha Van Borek, Mia L. van der Kop, Abigael Lukhwaro, Neville Li, Sarah Karanja, Anik R. Patel, David Ojaka, and Richard T. Lester. Mobile health for early retention in HIV care: A qualitative study in Kenya (WelTel Retain). 13(4):331–338.

- [180] Kirsten Smillie, Natasha Van Borek, Joshua Abaki, Neora Pick, Evelyn J. Maan, Karen Friesen, Rebecca Graham, Sarah Levine, Mia L. van der Kop, Richard T. Lester, and Melanie Murray. A Qualitative Study Investigating the Use of a Mobile Phone Short Message Service Designed to Improve HIV Adherence and Retention in Care in Canada (WelTel BC1). 25(6):614–625.
- [181] R. W. Snow, D. Zurovac, G. Jagoe, J. Barrington, S. Githinji, S. Kigen, A. M. Mbiti, D. Memusi, A. N. Muturi, A. Nyandigisi, and A. Wamari. Reducing Stock-Outs of Life Saving Malaria Commodities Using Mobile Phone Text-Messaging: SMS for Life Study in Kenya.
- [182] David B. Springer, Kirsten L. Bobrow, Naomi Levitt, Andrew Farmer, and Lionel Tarassenko. The SMS-text Adherence Support (StAR) Study: Hardware and Software Infrastructure. In *Proceedings of the Sixth International Conference on Information and Communications Technologies and Development: Notes - Volume 2*, ICTD '13, pages 147–150. ACM.
- [183] Richard Ssekibuule, John A. Quinn, and Kevin Leyton-Brown. A Mobile Market for Agricultural Trade in Uganda. In *Proceedings of the 4th Annual Symposium on Computing for Development*, ACM DEV-4 '13, pages 9:1–9:10. ACM.
- [184] Hyewon Suh, John R. Porter, Robert Racadio, Yi-Chen Sung, and Julie A. Kientz. Baby Steps Text: Feasibility Study of an SMS-Based Tool for Tracking Children's Developmental Progress. 2016:1997–2006.
- [185] Minhyang (Mia) Suh and Gary Hsieh. Designing for Future Behaviors: Understanding the Effect of Temporal Distance on Planned Behaviors. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 1084–1096. ACM.
- [186] Mukda Suktarachan, Patthrawan Rattanamanee, and Asanee Kawtrakul. The Development of a Question-answering Services System for the Farmer Through SMS: Query Analysis. In *Proceedings of the 2009 Workshop on Knowledge and Reasoning for Answering Questions*, KRAQ '09, pages 3–10. Association for Computational Linguistics.
- [187] Africa's Talking. Africa's Talking. [Online; accessed 28-Feb-2017].
- [188] Mauro Teófilo, Luiz Cavalcanti, and Vicente Ferreira de Lucena. A SMS-based Application Store for Emerging Market: A Case Study. In *SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications*, SA '13, pages 66:1–66:4. ACM.

- [189] Jay Thakkar, Rahul Kurup, Tracey-Lea Laba, Karla Santo, Aravinda Thiagalingam, Anthony Rodgers, Mark Woodward, Julie Redfern, and Clara K. Chow. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. 176(3):340–349.
- [190] Kentaro Toyama. Technology As Amplifier in International Development. In *Proceedings of the 2011 iConference*, iConference '11, pages 75–82. ACM.
- [191] Katherine Tumlinson, Chinelo C. Okigbo, and Ilene S. Speizer. Provider barriers to family planning access in urban Kenya. 92(2):143–151.
- [192] Twilio. Twilio. [Online; accessed 28-Feb-2017].
- [193] J. A. Unger, Keshet Ronen, Trevor Perrier, Brian DeRenzi, Jennifer Slyker, A. L. Drake, Dyphna Mogaka, John Kinuthia, and Grace John-Stewart. Short message service communication improves exclusive breastfeeding and early postpartum contraception in a low-to middle-income country setting: A randomised trial. 125(12):1620–1629.
- [194] Ushahidi. SMSsync. [Online; accessed 28-Feb-2017].
- [195] Mia L. van der Kop, Sarah Karanja, Lehana Thabane, Carlo Marra, Michael H. Chung, Lawrence Gelmon, Joshua Kimani, and Richard T. Lester. In-Depth Analysis of Patient-Clinician Cell Phone Communication during the WelTel Kenya1 Antiretroviral Adherence Trial. 7(9):e46033.
- [196] Trevor van Mierlo, Rachel Fournier, Anathalie Jean-Charles, Jacinthe Hovington, Isabelle Ethier, and Peter Selby. I'll Txt U if I Have a Problem: How the Société Canadienne du Cancer in Quebec Applied Behavior-Change Theory, Data Mining and Agile Software Development to Help Young Adults Quit Smoking. 9(3):e91832.
- [197] Aditya Vashistha, Edward Cutrell, Gaetano Borriello, and William Thies. Sangeet Swara: A Community-Moderated Voice Forum in Rural India. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, CHI '15, pages 417–426. ACM.
- [198] Aditya Vashistha and William Thies. Ivr junction: Building scalable and distributed voice forums in the developing world.
- [199] R. Veeraraghavan, N. Yasodhar, and K. Toyama. Warana Unwired: Replacing PCs with mobile phones in a rural sugarcane cooperative. In *International Conference on Information and Communication Technologies and Development, 2007. ICTD 2007*, pages 1–10.

- [200] Viswanath Venkatesh and Fred D. Davis. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. 46(2):186–204.
- [201] David S. Wald, Shahena Butt, and Jonathan P. Bestwick. One-way versus two-way text messaging on improving medication adherence: Meta-analysis of randomized trials. 128(10):1139–e1.
- [202] Ying Wang. Mobile appointment reminders in patient-centered care: Design and evaluation.
- [203] Thomas Webb, Judith Joseph, Lucy Yardley, and Susan Michie. Using the Internet to Promote Health Behavior Change: A Systematic Review and Meta-analysis of the Impact of Theoretical Basis, Use of Behavior Change Techniques, and Mode of Delivery on Efficacy. 12(1):e4.
- [204] Lu Wei-Chih, Matt Tierney, Jay Chen, Faiz Kazi, Alfredo Hubbard, Jesus Garcia Pasquel, Lakshminarayanan Subramanian, and Bharat Rao. UjU: SMS-based applications made easy. In *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10, pages 16:1–16:11. ACM.
- [205] Susan Wyche. Exploring Women's Everyday Mobile Phone Experiences in Nairobi, Kenya. 29(3):391–402.
- [206] Susan Wyche, Nightingale Simiyu, and Martha E. Othieno. Mobile Phones As Amplifiers of Social Inequality Among Rural Kenyan Women. 23(3):14:1–14:19.
- [207] Susan Wyche, Charles Steinfield, Tian Cai, Nightingale Simiyu, and Martha E. Othieno. Reflecting on Video: Exploring the Efficacy of Video for Teaching Device Literacy in Rural Kenya. In *Proceedings of the Eighth International Conference on Information and Communication Technologies and Development*, page 8. ACM.
- [208] Susan P. Wyche, Melissa Densmore, and Brian Samuel Geyer. Real Mobiles: Kenyan and Zambian Smallholder Farmers' Current Attitudes Towards Mobile Phones. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*, ICTD '15, pages 9:1–9:10. ACM.
- [209] Susan P. Wyche and Laura L. Murphy. "Dead China-make" Phones off the Grid: Investigating and Designing for Mobile Phone Use in Rural Africa. In *Proceedings of the Designing Interactive Systems Conference*, DIS '12, pages 186–195. ACM.
- [210] Khou Xiong, Joy Kamunyor, and Jane Sebidi. The MomConnect helpdesk: How an interactive mobile messaging programme is used by mothers in South Africa. 3:e000578.

- [211] Judith Yargawa and Jo Leonardi-Bee. Male involvement and maternal health outcomes: Systematic review and meta-analysis. 69(6):604–612.
- [212] Tae-Jung Yun and Rosa I. Arriaga. A Text Message a Day Keeps the Pulmonologist Away. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1769–1778. ACM.
- [213] Tae-Jung Yun, Hee Young Jeong, Tanisha D. Hill, Burt Lesnick, Randall Brown, Gregory D. Abowd, and Rosa I. Arriaga. Using SMS to Provide Continuous Assessment and Improve Health Outcomes for Children with Asthma. In *Proceedings of the 2Nd ACM SIGHIT International Health Informatics Symposium*, IHI '12, pages 621–630. ACM.
- [214] Dejan Zurovac, Gabriel Otieno, Samuel Kigen, Agneta M. Mbithi, Alex Muturi, Robert W. Snow, and Andrew Nyandigisi. Ownership and use of mobile phones among health workers, caregivers of sick children and adult patients in Kenya: Cross-sectional national survey. 9:20.
- [215] Dejan Zurovac, Raymond K Sudoi, Willis S Akhwale, Moses Ndiritu, Davidson H Hamer, Alexander K Rowe, and Robert W Snow. The effect of mobile phone text-message reminders on Kenyan health workers' adherence to malaria treatment guidelines: A cluster randomised trial. 378(9793):795–803.