

**Pervasive Computing and Public Health Research in Africa: Mobile Phones  
in the Collection, Analysis and Dissemination of Health Research**

**Alastair van Heerden**

Submitted in fulfilment of the requirements for the degree of  
DOCTOR OF PHILOSOPHY

Faculty of Health Sciences, University of the Witwatersrand  
Johannesburg, April 2013

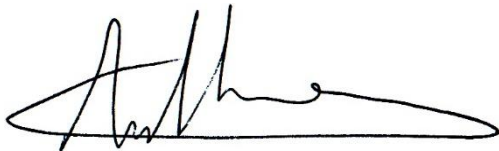
## Declaration

---

I declare that the thesis titled *Pervasive Computing and Public Health Research in Africa: Mobile Phones in the Collection, Analysis and Dissemination of Health Research* submitted for the degree of Doctor of Philosophy is the result of my own work. Any assistance received or reference made to the work of others has been fully acknowledged.

This thesis is being submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg. No part of this work has been submitted for degree or examination purposes to any other university or institution.

Alastair van Heerden



---

April 2013

## **Dedication**

---

This thesis is dedicated to the memory of Dev Griesel. Thank you for allowing me to stand on your shoulders.

## Thesis Material

---

Over the course of this PhD, the research was written up and published in a series of manuscripts. The results were also presented at various conferences, and successful grant applications attached to the study were obtained.

### **Peer-reviewed Publications**

1. Van Heerden AC, Norris SA, Richter LM. Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa. *The Journal of Adolescent Health* 2010;46(3):302-4.
2. Van Heerden AC, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. *WHO Bulletin* 2012;90(5):393-4.
3. Van Heerden AC, Norris SA, Stein AD, Richter LM, Tollman SM. On the move: Mobile phones as a platform for the delivery of questionnaires to young adults. Under review.
4. Van Heerden AC, Norris SA, Tollman SM, Richter LM. Collecting health research data: Comparing data quality, availability and costs of paper-and-pen versus mobile-phone interviews. Accepted, *Journal of Field Methods* 2012.
5. Van Heerden AC, Norris SA, Tollman SM, Richter LM, Rotheram-Borus MJ. Collecting maternal health information from HIV-positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa. Accepted, *Journal of Medical Internet Research*.

### **Student's contribution to the publications:**

1. **Van Heerden AC, Norris SA, Richter LM. Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa. *The Journal of Adolescent Health* 2010;46(3):302-4.**

This publication was a collaborative effort between all authors. I developed the survey items, analysed the data and prepared the manuscript. Dr Norris coordinated data collection and management, and assisted with the conceptualisation and preparation of the manuscript. Dr Richter assisted with the preparation of the manuscript.

**2. Van Heerden AC, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. WHO Bulletin 2012;90(5):393-4.**

I reviewed the literature and was responsible for conceptualising the recommendations made. Drs Tomlinson and Swartz edited and reviewed the manuscript.

**3. Van Heerden AC, Norris SA, Stein AD, Richter LM, Tollman SM. On the move: Mobile phones as a platform for the delivery of questionnaires to young adults. Under review, International Journal of Social Research Methodology 2012.**

I was responsible for the study design, questionnaire development, study implementation, data analysis and writing of the manuscript. Drs Norris and Richter assisted with the study design and reviewed the manuscript. Dr Stein provided guidance on the analysis and along with Drs Tollman, Richter and Norris, edited the manuscript.

**4. Van Heerden AC, Norris SA, Tollman SM, Richter LM. Collecting health research data: Comparing data quality, availability and costs of paper-and-pen versus mobile-phone interviews. Accepted, Journal of Field Methods 2012.**

I was responsible for the study design, questionnaire development and piloting, study implementation and management (including supervision of data collection

and training of field workers), data management (including data cleaning and coding), data analysis and writing of the manuscript. Drs Norris and Tollman provided guidance on the conceptualisation of the manuscripts, methodology and analysis. Drs Norris, Tollman and Richter assisted in editing the manuscript.

5. **Van Heerden AC, Norris SA, Tollman SM, Richter LM, Rotheram-Borus MJ. Collecting maternal health information from HIV-positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa. Under review, Journal of Medical Internet Research.**

Drs Richter and Rotheram-Borus were responsible for the primary trial design. I was responsible for study design, questionnaire development and piloting, study implementation and management (including supervision of data collection and training of field workers), data management (including data cleaning and coding), data analysis and writing of the manuscripts. Drs Norris, Tollman and Richter assisted with providing guidance on the conceptualisation of the manuscripts, methodology and analysis. Dr Rotheram-Borus assisted in reviewing the draft manuscript.

### **Conference Presentations**

1. Van Heerden AC. Mobile phones in health services: Perceptions and preferences of HIV-positive pregnant women in South Africa. Oral presentation at 4<sup>th</sup> South African AIDS Conference. Durban, South Africa. April 2009.
2. Van Heerden AC. Mobile phones in health services: Perceptions and preferences of HIV-positive pregnant women in South Africa. Oral presentation at XVIII International AIDS Conference. Vienna, Austria. July 2010.
3. Van Heerden AC. Birth to Twenty data collection by cellphone. 15<sup>th</sup> UKZN Postgrad Psychology Conference. Durban, South Africa. October 2011.

4. Van Heerden AC. Message delivery pending: Opportunities and challenges facing the collection of mHealth data from a cohort of young adults in South Africa. Oral presentation at the 2010 mHealth Summit. Washington, USA. December 2010.

#### **Grants and Awards**

1. 2010: Faculty of Health Sciences Individual Research Grant (PI: Alastair van Heerden; Amount: R10 000)
2. 2010: mHealth Summit Meritorious Young Investigator Award

# Abstract

---

## Introduction

With aging populations and rising health care costs, many high-income countries are exploring mobile computing technologies to improve the efficiency and effectiveness of health care provision. These technologies, which underpin the field of pervasive computing, introduce a new model of human–computer interaction. Instead of the scenario where a single user interacts with a desk-bound “personal” computer, pervasive computing envisions a world embedded with small, inexpensive, portable networked devices able to communicate seamlessly with each other. In common with resource-rich countries, the field of pervasive computing has the potential to promote and support healthy population development in middle and low-income countries, and this, therefore, has relevance for South Africa. Current estimates suggest that there are between 28 and 32 million mobile phones in South Africa. This means that around 60% of all South Africans own, or have access to, mobile telecommunication. Over 900 000 km<sup>2</sup> of the country is covered by the GSM (Global System for Mobile Communication) network of Vodacom, the largest telecommunications company in the country. Over 90% of South Africa is provided with access to mobile connectivity through shared agreements between the country’s major telecommunications networks.

## Aims

The ubiquity of mobile phones has resulted in their receiving increasing attention from public health researchers. Yet a better understanding of how mobile phones could support health research in South Africa is still an emerging field with many unanswered

questions. This thesis attempts to fill some of these gaps in our current knowledge. In particular, the primary aim of this work is to implement and evaluate the use of mobile phones as instruments with which to collect and analyse information for monitoring, evaluation and research in low-resource rural African settings.

## **Methods**

To investigate this aim, data were gathered from the development, implementation and evaluation of four health surveys in South Africa. Two surveys were conducted with Birth to Twenty, a birth cohort of South African young adults living in Greater Johannesburg. These data were used to better understand the feasibility and data-quality implications of using mobile phones as a tool for the administration of 'self-administered' surveys. Two additional surveys, completed in KwaZulu-Natal province, evaluated the same themes of feasibility, acceptability and impact of data quality in mobile-phone-assisted personal (face-to-face) interviews (MPAPI). The first, conducted with 500 HIV-positive pregnant women in eight primary health clinics and 12 interviewers trained to use the mobile-phone survey software, was used to assess the feasibility and acceptability of MPAPI. The final survey compared the difference in data quality achieved by 100 interviewers using either pen and paper, or mobile phones to conduct a short health survey. De Leeuw's conceptual model was used to frame how mode characteristics influence data quality.

## **Results**

Mobile-phone-assisted interviewing was found to have an impact on the data quality, feasibility and acceptability of health surveys. MPAPI was found to be similar in terms of accuracy and cost to small-scale paper-and-pen interviewing (PAPI) surveys. Time lines

and accessibility were improved by the use of MPAPI. Mobile-phone-assisted self-interviewing (MPASI) surveys were found to have a lower survey response but a higher item-completion rate. Acceptability was found to be moderated by technological familiarity and the use patterns of mobile-phone features. Finally, conducting health research using mobile-phone interviews in South Africa was found to be feasible; to reduce the loss of questionnaires, and photocopying and data-entry costs; and to improve the speed at which data becomes available for analysis. Factors that mediated feasibility included the technical expertise of the project management and field staff, the technological know-how of participants, the comprehensiveness of the interviewer training, the mobile communication channel used (e.g., handset-agnostic SMS) and the presence or absence of an interviewer.

## **Conclusion**

Under the right conditions, mobile-phone-assisted interviewing appears to be a feasible and practical tool for the rapid collection of health information, with data accuracy being the same or better than pen-and-paper interviews. It is argued that these benefits increase as the scale of the survey increases. Improved data can positively influence population health by providing decision makers with more rapid access to accurate data with which to monitor large-scale health systems. Small projects that do not require the rapid availability of data or where staff do not have the appropriate technical proficiencies would be better suited at present to more traditional survey data-collection techniques.

**Keywords:** mobile phones; pervasive computing; mHealth; data collection; survey error

## Acknowledgements

---

While writing a dissertation is an individual pursuit, I was never alone. I had the constant support and encouragement of family, friends, colleagues and mentors. I cannot even begin to express my gratitude for their willingness to put up with me during these years. I wish to extend my heartfelt gratitude to my colleagues, family and friends who have supported me over the past five years. In particular, I would like to thank the following people whose support, guidance and encouragement have made this work possible.

Professor Shane Norris, for the dedication, importance and time you make available to your students. Thank you for walking by my side for the entire journey and never giving up on me or showing your frustration at slow progress. I feel incredibly lucky to have made your acquaintance as a student, colleague and friend.

Professor Steve Tollman, for the clarity of thought and insight you contributed. Although our time together was limited, I learned a great deal watching the way you questioned my study design, methodology, results and conclusions.

Professor Linda Richter, for entrusting the idea for this research to me. From my undergraduate studies through to the completion of this work, you have been a guide, mentor and friend. I appreciate all the opportunities you have afforded me and owe my success this far to you.

Dr Heidi van Rooyen, for your commitment to my development. You have modelled for me a way of leadership and management that has had a profound influence on me. I know no better example of a servant leader than you.

Professor Mary-Jane Rotheram-Borus at the Centre for Community Health, University of California, Los Angeles. It was a great privilege to have had the opportunity to work with you and I learned so much from the experience.

Professor Alan Stein, for your support and contribution to this work.

I would like to thank the Wellcome Trust for their support of my work with the Birth to Twenty project, and also the University of the Witwatersrand and the Human Sciences Research Council who helped fund this doctoral study with both time and money.

To the staff at Birth to Twenty, in particular Mathew Mainwaring and Martin Manyike, thank you for assisting me with data collection in such a friendly manner.

To the 'Masi Ladies' who gave five years of their lives collecting data for Project Masihambisane, thank you!

I wish also to acknowledge and thank all those people who agreed to participate in this study. None of the insights I gained would have been possible without you.

Finally, to my family: Cath, this achievement would mean nothing if I did not have you to share it with. Mom, Dad, Kirsten, and my new family of Mom, Dad, Brian and Anthony, thank you for showing an interest in my work and for your encouragement when things were hard.

# Table of Contents

---

|                                                                                |              |
|--------------------------------------------------------------------------------|--------------|
| <b>Declaration .....</b>                                                       | <b>ii</b>    |
| <b>Dedication.....</b>                                                         | <b>iii</b>   |
| <b>Thesis Material.....</b>                                                    | <b>iv</b>    |
| <b>Abstract .....</b>                                                          | <b>viii</b>  |
| <b>Acknowledgements .....</b>                                                  | <b>xi</b>    |
| <b>Table of Contents .....</b>                                                 | <b>xiv</b>   |
| <b>List of Tables .....</b>                                                    | <b>xviii</b> |
| <b>List of Figures .....</b>                                                   | <b>xix</b>   |
| <b>Definition of Terms and Abbreviations.....</b>                              | <b>xxi</b>   |
| <b>Preface.....</b>                                                            | <b>xxiv</b>  |
| <b>Part 1: Background and Literature .....</b>                                 | <b>1</b>     |
| <b>Chapter 1: Introduction .....</b>                                           | <b>3</b>     |
| 1.1 Technology through the Ages.....                                           | 3            |
| 1.1.1 Theory of technological evolution .....                                  | 4            |
| 1.1.2 The three revolutions of computing .....                                 | 9            |
| 1.1.3 Pervasive computing .....                                                | 11           |
| 1.1.4 Summary .....                                                            | 17           |
| <b>Chapter 2: Literature Review .....</b>                                      | <b>20</b>    |
| 2.1 Theoretical Framework for Quantitative Survey Data Collection .....        | 20           |
| 2.1.1 The Total Survey Error framework.....                                    | 21           |
| 2.1.2 Multidimensional frameworks for assessing data quality .....             | 25           |
| 2.1.3 Theoretical impact of survey mode on survey error and data quality ..... | 27           |
| 2.1.4 Practical implications of the theoretical frameworks .....               | 35           |
| 2.2 Mobile Phones in Sub-Saharan Africa .....                                  | 40           |
| 2.2.1 Mobile phones for development .....                                      | 45           |
| 2.2.2 The rise of mobile phones in South Africa .....                          | 48           |
| 2.3 Application of Mobile Phones to Data Collection in Health Research .....   | 52           |
| 2.3.1 Technology behind mobile-phone health interventions .....                | 53           |
| 2.3.2 Application of mobile phones in health .....                             | 57           |
| 2.3.3 Infectious-disease surveillance .....                                    | 59           |

|                                                                                                                                                                            |           |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| 2.3.4 Treatment monitoring.....                                                                                                                                            | 61        |
| 2.3.5 Comparison with other methods .....                                                                                                                                  | 62        |
| 2.4 Summary of the Literature Review.....                                                                                                                                  | 64        |
| <b>Part 2: Empirical Papers.....</b>                                                                                                                                       | <b>67</b> |
| <b>Chapter 3: Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa. ....</b> | <b>69</b> |
| 3.1 Method .....                                                                                                                                                           | 70        |
| 3.2 Results.....                                                                                                                                                           | 70        |
| 3.3 Discussion .....                                                                                                                                                       | 72        |
| <b>Chapter 4: On the move: Mobile phones as a platform for the delivery of questionnaires to young adults. ....</b>                                                        | <b>75</b> |
| 4.1 Introduction.....                                                                                                                                                      | 75        |
| 4.1.1 Retaining contact with an increasingly mobile population .....                                                                                                       | 75        |
| 4.1.2 Collecting reliable information pertaining to sensitive behaviours.....                                                                                              | 76        |
| 4.1.3 Accurately measuring infrequently or irregularly occurring behaviours.....                                                                                           | 76        |
| 4.1.4 Mobile-phone surveys.....                                                                                                                                            | 77        |
| 4.2 Methods.....                                                                                                                                                           | 77        |
| 4.2.1 Study population and sample .....                                                                                                                                    | 77        |
| 4.2.2 Participant training.....                                                                                                                                            | 78        |
| 4.2.3 Questionnaire.....                                                                                                                                                   | 79        |
| 4.2.4 Mobenzi Researcher software .....                                                                                                                                    | 79        |
| 4.2.5 Procedures .....                                                                                                                                                     | 80        |
| 4.2.6 Data analysis and outcome indicators .....                                                                                                                           | 80        |
| 4.3 Results.....                                                                                                                                                           | 81        |
| 4.3.1 Number of participants possessing handsets capable of submitting surveys ....                                                                                        | 81        |
| 4.3.2 Texting the 435 handsets included in the study .....                                                                                                                 | 82        |
| 4.3.3 Survey response rates .....                                                                                                                                          | 83        |
| 4.3.4 Item response rates .....                                                                                                                                            | 84        |
| 4.4 Discussion .....                                                                                                                                                       | 84        |
| <b>Chapter 5: Collecting maternal health information from HIV-positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa .....</b>     | <b>90</b> |
| 5.1 Introduction.....                                                                                                                                                      | 90        |
| 5.2 Method .....                                                                                                                                                           | 92        |
| 5.2.1 Study design .....                                                                                                                                                   | 92        |
| 5.2.2 Participants and procedures .....                                                                                                                                    | 93        |

|                                                                                                                           |            |
|---------------------------------------------------------------------------------------------------------------------------|------------|
| 5.3 Statistical Analysis .....                                                                                            | 95         |
| 5.4 Fieldwork Supervision and Ethical Approval .....                                                                      | 96         |
| 5.5 Results.....                                                                                                          | 96         |
| 5.5.1 Acceptability.....                                                                                                  | 96         |
| 5.5.2 Feasibility.....                                                                                                    | 103        |
| 5.6 Discussion .....                                                                                                      | 105        |
| 5.6.1 Acceptability.....                                                                                                  | 106        |
| 5.6.2 Feasibility.....                                                                                                    | 107        |
| 5.7 Relevance to PMTCT Cascade.....                                                                                       | 109        |
| 5.8 Limitations and Future Work.....                                                                                      | 109        |
| <b>Chapter 6: Collecting health research data: Comparing paper-and-pen to mobile-phone-assisted data collection. ....</b> | <b>112</b> |
| 6.1 Introduction.....                                                                                                     | 112        |
| 6.2 Method .....                                                                                                          | 114        |
| 6.2.1 Study design and sample selection .....                                                                             | 114        |
| 6.2.2 Procedure .....                                                                                                     | 115        |
| 6.3 Results.....                                                                                                          | 118        |
| 6.3.1 Cost-effectiveness .....                                                                                            | 121        |
| 6.4 Discussion .....                                                                                                      | 123        |
| 6.5 Conclusion .....                                                                                                      | 128        |
| <b>Chapter 7: Point of care in your pocket: A research agenda for the field of mHealth...130</b>                          |            |
| 7.1 mHealth Needs to Develop an Evidence Base.....                                                                        | 130        |
| 7.2 mHealth Systems Should Be Interoperable with Existing eHealth Initiatives .....                                       | 131        |
| 7.3 mHealth Should Adopt and Implement the Same Standards Already Present in eHealth .....                                | 132        |
| 7.4 mHealth Should Take a Participatory Approach .....                                                                    | 132        |
| 7.5 mHealth Should Promote Equity in Health .....                                                                         | 133        |
| 7.6 mHealth Programmes Need to Move to a Situation of Sustainability .....                                                | 134        |
| 7.7 mHealth Needs to Focus on Health, Not on the Technology .....                                                         | 134        |
| <b>Part 3: Discussion .....</b>                                                                                           | <b>136</b> |
| <b>Chapter 8: Discussion .....</b>                                                                                        | <b>138</b> |
| 8.1 Consolidated Findings.....                                                                                            | 138        |

|                                                                                                                        |            |
|------------------------------------------------------------------------------------------------------------------------|------------|
| 8.1.1 Media-related factors .....                                                                                      | 143        |
| 8.1.2 Information transmission .....                                                                                   | 146        |
| 8.1.3 Interviewer impact .....                                                                                         | 148        |
| 8.1.4 Data quality .....                                                                                               | 149        |
| 8.2 Emerging Research Themes .....                                                                                     | 154        |
| 8.2.1 The influence of scale .....                                                                                     | 154        |
| 8.2.2 The influence of context .....                                                                                   | 156        |
| 8.3 Theoretical Relevance .....                                                                                        | 157        |
| 8.3.1 Tourangeau et al.'s conceptual model of how psychological variables, mediated by mode, affect data quality ..... | 158        |
| 8.3.2 A generalised pervasive computing framework for health research suggested from the thesis .....                  | 159        |
| 8.4 Contextual Relevance .....                                                                                         | 163        |
| 8.4.1 Treatment compliance .....                                                                                       | 164        |
| 8.4.2 Data collection .....                                                                                            | 164        |
| 8.4.3 Health information systems and point-of-care support tools for health workers .....                              | 165        |
| 8.4.4 Disease prevention and health promotion .....                                                                    | 166        |
| 8.5 Limitation and strength .....                                                                                      | 167        |
| 8.5.1 Further research .....                                                                                           | 167        |
| 8.6 Conclusion .....                                                                                                   | 169        |
| <b>References .....</b>                                                                                                | <b>171</b> |
| <b>Appendices .....</b>                                                                                                | <b>189</b> |
| <b>Appendix 1 – Approval of thesis title .....</b>                                                                     | <b>190</b> |
| <b>Appendix 2 – Ethical approval .....</b>                                                                             | <b>191</b> |
| <b>Appendix 3 – Original published papers .....</b>                                                                    | <b>192</b> |

## List of Tables

---

|                                                                                                                                                                          |     |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Table 2.1. Common dimensions that impact survey data quality .....                                                                                                       | 26  |
| Table 2.2. Key global telecom indicators for the world telecommunication service sector in 2010 as provided by the International Telecommunication Union .....           | 42  |
| Table 2.3. South African mobile-phone ownership trends (%) from the South African Advertising Research Foundation's All Media Products Study (2011) <sup>108</sup> ..... | 49  |
| Table 2.4. An overview of mobile-phone technologies that enable the collection of survey data on the handset .....                                                       | 55  |
| Table 3.1. Importance of mobile-phone feature (including voice calls, SMS, MMS, camera and Internet access) by gender in the Birth to Twenty Cohort (2009) .....         | 71  |
| Table 3.2. Preference for mobile-phone questionnaire or face-to-face interview and technological affinity in the Birth to Twenty Cohort (2009) .....                     | 72  |
| Table 4.1. Example of text message responses received from Birth to Twenty study participants in response to survey notification SMS messages .....                      | 83  |
| Table 5.1. The twelve items of the Technology Acceptance Model Scale .....                                                                                               | 94  |
| Table 5.2. Demographic characteristics of Mothers Living with HIV (n=512) surveyed in 2010 .....                                                                         | 100 |
| Table 5.3. Questionnaire experience and preference of Mothers Living with HIV (N=512) surveyed in 2010 .....                                                             | 101 |
| Table 6.1. Comparative demographic profiles of interviewers enrolled to collect survey data using either paper and pen or MPAPI .....                                    | 115 |
| Table 6.2. Comparison of data errors for PAP and MPAPI surveys by gender, age, interview order and interviewee .....                                                     | 119 |
| Table 6.3. Comparison of the mean number of data-entry errors made by interviewers assigned to either PAP or MPAPI .....                                                 | 120 |
| Table 6.4. Effect of survey mode and the number of interviews already completed (experience) on survey duration .....                                                    | 121 |
| Table 6.5. Costs and errors associated with PAP and MPAPI survey modes .....                                                                                             | 122 |
| Table 8.1. Summary of thesis findings .....                                                                                                                              | 140 |
| Table 8.2. Table of key findings .....                                                                                                                                   | 153 |

## List of Figures

---

|                                                                                                                                                                                                                                 |     |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| Figure 1.1. The S-Curve theory of technological advancement <sup>14</sup> .....                                                                                                                                                 | 6   |
| Figure 1.2. The continuous evolution of Australian aboriginal weapons as arranged by Pitt-Rivers (1906). Radiating out from the simple stick are war clubs, boomerangs, lances, throwing sticks and shields <sup>16</sup> ..... | 8   |
| Figure 1.3. The three revolutions of computing. Adapted from Weiser <sup>19</sup> .....                                                                                                                                         | 9   |
| Figure 1.4. A working model of how context can be ascertained by mobile devices <sup>34</sup> ...                                                                                                                               | 14  |
| Figure 2.1. Groves's (1989) conceptual framework for calculating total survey error (accuracy) <sup>51</sup> .....                                                                                                              | 24  |
| Figure 2.2. The evolution of survey technology from the use of paper to computers <sup>51</sup> ..                                                                                                                              | 28  |
| Figure 2.3. De Leeuw's model of mode characteristics influencing data quality <sup>72 (p19)</sup> ...                                                                                                                           | 29  |
| Figure 2.4. Impact of mode and interviewer presence on mean number of reported sexual partners <sup>76(p277)</sup> .....                                                                                                        | 31  |
| Figure 2.5. Tourangeau et al.'s model of how psychological variables, mediated by mode, affect data quality <sup>76</sup> .....                                                                                                 | 32  |
| Figure 2.6. Worldwide growth of mobile phones and mobile broadband over the past five years <sup>97</sup> .....                                                                                                                 | 40  |
| Figure 2.7. Changing focus of information and communications technology for development (ICT4D) priorities over time .....                                                                                                      | 47  |
| Figure 2.8. Manufacturer preferences among 2 000 South African respondents, with weightings based on age, gender and all LSM levels in 2011 <sup>109</sup> .....                                                                | 50  |
| Figure 2.9. Vodacom cellular network coverage map of South Africa (2011) .....                                                                                                                                                  | 51  |
| Figure 2.10. A schematic representation of how data is collected, uploaded/processed and viewed in a mobile-phone-assisted interview .....                                                                                      | 53  |
| Figure 2.11. A sample of mobile-phone applications available in 2012 that are designed to provide feedback on various health measures to the user .....                                                                         | 58  |
| Figure 4.1. Phones not available for use from the 1 000 donated handsets that were distributed .....                                                                                                                            | 81  |
| Figure 4.2. Schematic overview of two data-collection waves .....                                                                                                                                                               | 82  |
| Figure 5.1. Estimated marginal means for the main effect of time .....                                                                                                                                                          | 102 |

**Figure 5.2. Webpage dashboard snippet summarising research assistant interview performance for a baseline interview conducted in September 2009 ..... 105**

**Figure 6.1. Extrapolation of survey cost and the number of surveys conducted..... 123**

**Figure 8.1. Thesis findings: The impact of mobile-phone-assisted data collection on data quality within De Leeuw's conceptual model of how data-collection mode affects data quality..... 142**

**Figure 8.2. A conceptual framework, adapted from Tourangeau et al.,<sup>76</sup> hypothesising the impact of electronic data-collection survey modes on psychological appraisal and response quality..... 160**

## Definition of Terms and Abbreviations

---

3G: Third-generation technology

A-CASI: Audio-based computer-assisted self-interviewing

AMPS: All Media and Products Survey

ART: Antiretroviral therapy

BT20: Birth to Twenty (cohort)

CAPI: Computer-assisted personal interviewing

CASI: Computer-assisted self-interviewing

CASIC: Computer-assisted survey information collection

CATI: Computer-assisted telephonic interviews

CAW: Computer-assisted web interviewing

CHW: Community health worker

DHIS: District Health Information System

EDGE: Enhanced data rates for GSM evolution (2.75G)

EMA: Ecological momentary assessment

FTF: Face-to-face interviewing

GPRS: General packet radio service (2.5G)

GPS: Global positioning system

GSM: Global system for mobile communication (2G)

GSMA: Global System for Mobile Communication Association

HSDPA: High-speed downlink packet access (also called 3.5G)

HSRC: Human Sciences Research Council

ICT: Information and communications technology

ICT4D: Information and communications technology for development

ITU: International Telecommunication Union

IVR: Interactive Voice Response

J2ME: Java 2 Micro Edition

LMICs: Low and middle-income countries

LSM: Living standards measure

MDR-TB: Multi-drug resistant tuberculosis

mHealth: Mobile health

MHL: Mothers living with HIV and AIDS

Mobile phone penetration rate: The number of active mobile phone numbers (usually as a percentage) within a specific population.

Mode effect: *The unique measurement properties [of modes] may mean that differences observed in data can be attributed to how the data have been collected rather than to real differences in the survey population*<sup>1</sup>

MPAI: Mobile-phone-assisted interviewing

MPAPI: Mobile-phone-assisted personal interviewing

MPASI: Mobile-phone-assisted self-interviewing

OS: Operating system

PAPI: Paper-and-pencil interviewing

PDA: Personal digital assistant

RCT: Randomised control trial

SAARF: South African Advertising Research Foundation

SAQ: Self-administered questionnaires

T-CASI: Text-based computer-assisted self-interviewing

TDE: Touchtone data entry

TSE: Total survey error: an accumulation of all errors resulting from the design, collection, processing and analysis of survey data.

USSD: Unstructured supplementary service data

V-CASI: Video-based computer-assisted self-interviewing

XDR-TB: Extensively drug-resistant tuberculosis

*Three objects were considered essential across all participants, cultures and genders: keys, money and mobile phone.*

*Nokia trend-spotter Jan Chipchase, 2007<sup>2</sup>*

## Preface

---

Towards the end of 2007, I was approached by an old friend and mentor and asked to consider applying for a PhD internship position in her research programme. The programme had just received a grant to complete a five-year clinic-based effectiveness trial in KwaZulu-Natal. In order to prevent contamination with other studies already underway in the area, eight geographically isolated primary health care clinics were chosen for this new trial. It soon became evident that the solution to how best to collect data from these widely dispersed, low-resourced rural clinics was not a trivial problem to solve. Having audited the Provincial Department of Health's informatics system in 2005, I was well aware of the challenges faced in collecting and aggregating accurate, timely and useful data from provincial clinics and hospitals. Broken computers, missing registers, double counting, undercounting, incorrect versions of the data-collection forms and staff shortages all erode the quality of data produced by the District Health Information System (DHIS).

Faced with the challenge of implementing my own mini DHIS, I set up appointments with a number of colleagues, hoping that, three years on, a better solution had been found. Sadly, I found that the exact same system, with all its inefficiencies and challenges, was still in use. Having come up empty-handed, the team and I went on to consult with researchers and academics working in the field. Through this process a number of candidate solutions to our problem began to emerge. The research I present in this thesis documents my experience with the novel approach chosen to collect research data in such resource-constrained environments.

I have always been interested in technology and was well aware of the rapid advances being made in mobile-phone hardware. But even I have been astounded by how fast miniaturisation, increased mobile network speeds, decreased hardware costs and electronic device convergence have affected the mobile telecommunications industry. In the five years it has taken me to formulate, propose, implement and write up this thesis, mobile phones have changed dramatically from the basic communication platforms that were capable of making calls, sending text messages and, if you had the technical know-how, connecting to basic websites with a preferred computing device. Taking photos, listening to music, getting turn-by-turn navigational direction, watching movies and emailing are now all common activities to perform on a mobile phone, with many of these features available on entry-level handsets. These features have been enabled through the inclusion of more and more hardware into the mobile-phone handset. Cameras, GPS units, accelerometers, proximity sensors and WiFi hardware are now all available in a single handset. This trend is only set to gain momentum in the coming years.

It is not, however, only device hardware that has undergone rapid change. In July 2008 Apple launched their app market, which allowed owners of the iPhone to download and install applications on their phones. The introduction of the app market fundamentally altered the way that people use and think about their mobile phones. Phones are no longer simply mobile telephones, but have rather become portable computers that enable users to book a table at their favourite restaurant, manage their personal finances, translate text or go virtual bass fishing. While sightseeing on a recent trip to Vienna, I came across a very large and majestic building. Unable to locate it in my guide book, I

pulled out my mobile phone and took a photo with the “Google Goggles” application. Within seconds I was informed by the application that I was standing in front of the National Fine Arts Museum and that it was closed on Mondays. On entering the gallery I was particularly taken by one of the works hanging in the gallery. Again, within a few seconds of firing up Goggles I was presented with a wealth of information about the Van Gogh self-portrait I was standing in front of.

None of this was yet possible in late 2007, but the potential for mobile phones to double as research tools was already beginning to be noted both within academia and the private sector. A research partnership between the University of Cape Town and the Cape Peninsula University of Technology was formed in 2001. The collaboration aimed to develop a system that would allow nurses to be linked to and communicate with home-based carers responsible for monitoring the treatment of HIV-positive patients. A system was built that relied on mobile phones and the JavaRosa data-collection platform. This system facilitated the flow of information between nursing staff and home-based carers. The research project was expanded into a not-for-profit company called Cell-Life in 2005. In 2006, Clyral CC introduced the commercial Mobile Researcher platform. The software enabled mobile phones capable of running Java 2 Micro Edition software to be turned into survey data-collecting tools. Surveys are first prepared online using the provided set of survey-design tools. These surveys are then downloaded onto mobile phones running the Mobile Researcher software. Although any data can be collected using the Mobile Researcher platform, the company has tailored its product to meet the needs of social and medical science researchers.

The collection of health information using mobile phones is only one aspect of a field now known as mHealth (mobile health). Although there is still debate about the exact definition of mHealth, in the context of this thesis mHealth refers to leveraging the unique attributes of mobile communications and network technologies in order to improve detection, diagnosis, health monitoring and care. The unique attributes that lend the use of mobile devices to health care are that they are personal, they are permanently carried, they are always on and they provide a two-way communication channel to a large percentage of the population — even in low-resource settings. mHealth seeks to harness these unique factors in such a way that barriers of time and location are removed from health care provision. It aims to improve health care and make it available at any time, to anyone, living anywhere.

These early initiatives by Cell-Life and Mobile Researcher showed great potential and ultimately influenced our decision to use mobile phones to collect our research data. This thesis implements and evaluates the use of mobile phones as research instruments with which to collect and analyse information for research in low-resource South African settings. This primary aim was achieved through the completion of three studies, each designed to address one aspect of the use of mobile phones for this purpose.

1. How does the use of mobile phones for data collection impact on costs, quality and time-to-availability of data?
2. What perceptions do vulnerable communities and fieldworkers in rural South Africa hold about mobile phones as a data-collection tool?
3. Can mobile phones be used by adolescents to complete self-report surveys unaided?

#### 4. What are the policy and research agenda implications of this research?

These general objectives are designed to examine different facets of the primary aim.

The thesis by publication was undertaken in the School of Public Health at the University of the Witwatersrand. Part 1 is composed of two chapters. The first, Chapter 1, provides a contextual background to the evolution of technology through time. Chapter 2 builds on this introduction by presenting a summary of published literature in the field of mobile data-collection methods. It relies primarily on peer-reviewed articles but, where appropriate, draws on work from other publication sources. Part 2 (chapters 3 to 7) forms the body of this thesis. Each chapter presents one of the five submitted journal papers that were generated from this research. Chapter 3 is a research letter titled “Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa.” and was published in the *Journal of Adolescent Health*. It presents data collected on mobile-phone ownership and use in the Birth to Twenty cohort. Chapter 4 is under review for publication in the *International Journal of Social Research Methodology*. The research presented is a follow-on study of the pilot work presented in Chapter 3. The publication is titled “On the move: Mobile phones as a platform for the delivery of questionnaires to young adults”. Chapter 5, titled “Collecting maternal health information from HIV- positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa” is under review for publication in the *Journal of Medical Internet Research*. The chapter presents a cross-sectional study on the views of HIV-positive pregnant women about being interviewed using a mobile-phone survey. Chapter 6 has been accepted for publication in the journal

of *Field Methods*. Titled “Collecting health research data: Comparing paper-and-pen to mobile-phone-assisted data collection”, the chapter describes the differences in cost and data accuracy achieved using pen-and-paper and mobile-phone surveys. Chapter 7 was published in a special eHealth-themed issue of the *WHO Bulletin*. The paper presents seven key considerations to be taken into account when formulating a research agenda for mHealth. The paper was titled “Point of care in your pocket: A research agenda for the field of mHealth”. Finally, in Part 3, Chapter 8 concludes with a synthesis of the findings from the papers presented in Part 2 and returns to address the primary aim of this research. It provides a reference point for looking back on what has been learnt, while at the same time suggesting some of the steps that now need to be taken.

---

# **Part 1: Background and Literature**

---

# **Chapter 1**

---

## **Introduction**

## Chapter 1: Introduction

---

This chapter presents a framework in which to situate the studies undertaken in this thesis. The core of this dissertation is organised around three primary themes: survey research, pervasive computing and mobile phones applied to health research. To traverse these seemingly disparate content areas successfully, the organisation of this chapter reflects these themes and presents an overview of the relevant literature, divided into five sections. The first two sections provide an introduction to the conceptual frameworks underpinning this thesis; the third explores the use and impact of mobile phones in Sub-Saharan Africa; the fourth ties together survey research and pervasive computing by considering the convergence of mobile phones and survey data collection; and the final section presents a summary of key points from the literature.

### 1.1 Technology through the Ages

*The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.*

Mark Weiser

To locate the current study of mobile-phone data collection within both the health research field and the broader trends of computing, it is necessary first to provide an overview of the field of pervasive computing which aims to *deliver information to computer users wherever and whenever they might need it*<sup>4 (p13)</sup>. To achieve this goal, the section begins with providing some background to the concept of technological evolution. This framework is then used to examine how advances in technology have specifically

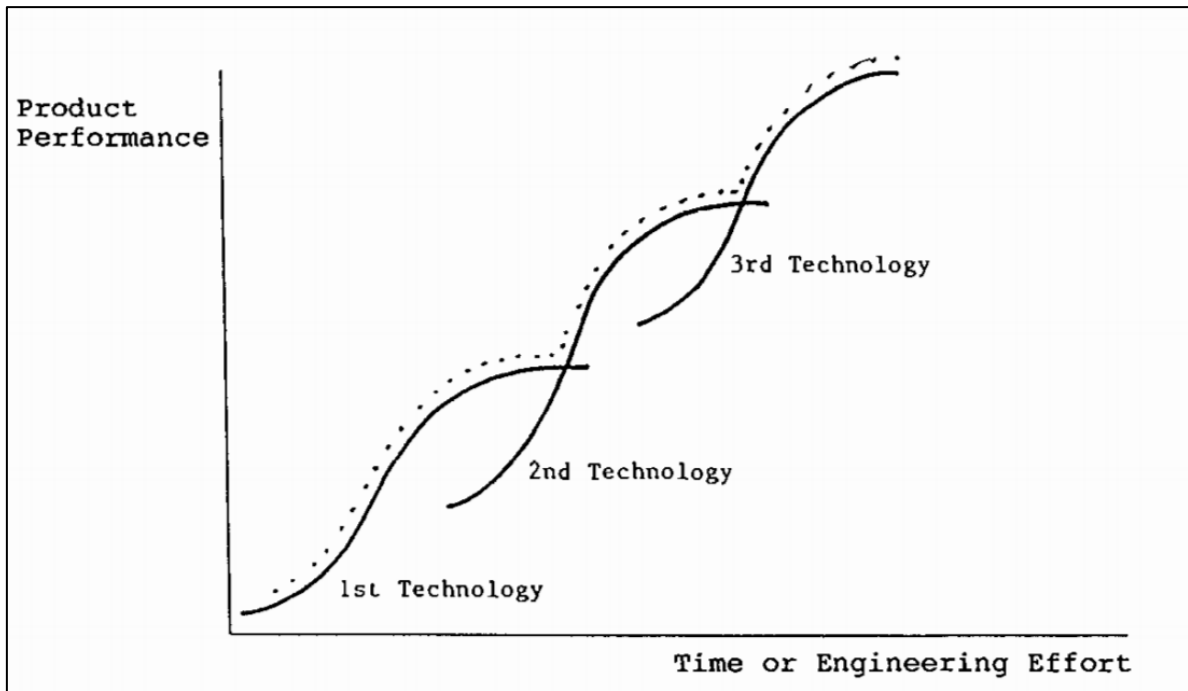
impacted on the development of computing devices. A framework for understanding these advances is then offered with the aim of describing the field of pervasive computing. Finally, mobile phones are discussed within the context of health research.

### **1.1.1 Theory of technological evolution**

In 1859 Charles Darwin published *On the Origin of Species*, in which he outlined his theory of evolutionary biology<sup>5</sup>. Four years later in a newspaper article, Samuel Butler commented that although machines are inanimate, they themselves may be governed by similar evolutionary processes<sup>6</sup>. Nearly 100 years passed before this concept of technological evolution was described in the scientific literature by the Czech philosopher, Radovan Richta. He proposed that technology, defined as *a material entity created by the application of mental and physical effort to nature in order to achieve some value*,<sup>7(p13)</sup> evolved through three stages. The first era of human technological advance was the “tool”. These assistive devices gave early humans a mechanical advantage with which to complete physical tasks. Examples given by Richta include the container, spear, arrow and plough. In the second technological era, tools were joined by machines. What distinguishes a tool from a machine is that tools are powered by human or animal effort, while machines are self-powered, relying on an operator only to control their function. Examples include motorbikes, aeroplanes, computers and stoves. The third era of technological evolution proposed by Richta is automation. Having removed the need for direct human physical effort when moving from tool to machine, automation removes the need for human control of machines. Examples of this type of technology include computer programs, robotic vacuum cleaners and pacemakers. Richta makes the point that all three eras or stages simply highlight the introduction of different classes of technology, each overlapping and evolving from what came before and all still in common

use. A nail and a computer keyboard, although created many years apart, are both examples of tools in use today. As noted astutely by McLennan<sup>8</sup>, technology has evolved so rapidly that we exist today in an era beyond the three-stage model proposed by Richta. Many of today's technological breakthroughs are made by automated machines that we have designed.

The model proposed by Richta is therefore useful only up to a point. He provides no means by which to predict how new machines arise or evolve from current models, what sparks innovation or what factors led to humans progressing through the three stages proposed by his theory. Two competing theories of change arose to address these issues. The first postulates that technology changes in small, continuous, incremental shifts; the other that long periods of equilibrium and stability are punctuated by rapid revolutionary change<sup>9,10</sup>. Slow, incremental change is grounded in the Darwinian concept of gradualism. Change is continuous, incremental and cumulative. Punctuated equilibrium<sup>11,12</sup> suggests that change can sometimes be slow and incremental while at other times rapid and abrupt. This paradigm was modelled to an S-Curve (logistic curve) function<sup>13,14</sup>. Initially, new technology struggles to improve upon the performance of existing technologies. Slow improvement is followed by rapid and disruptive growth in the maturity of this new technology, which sees it supersede the capabilities of what old technologies could achieve. This growth is then followed after some variable time by a period of stagnation (Figure 1.1).

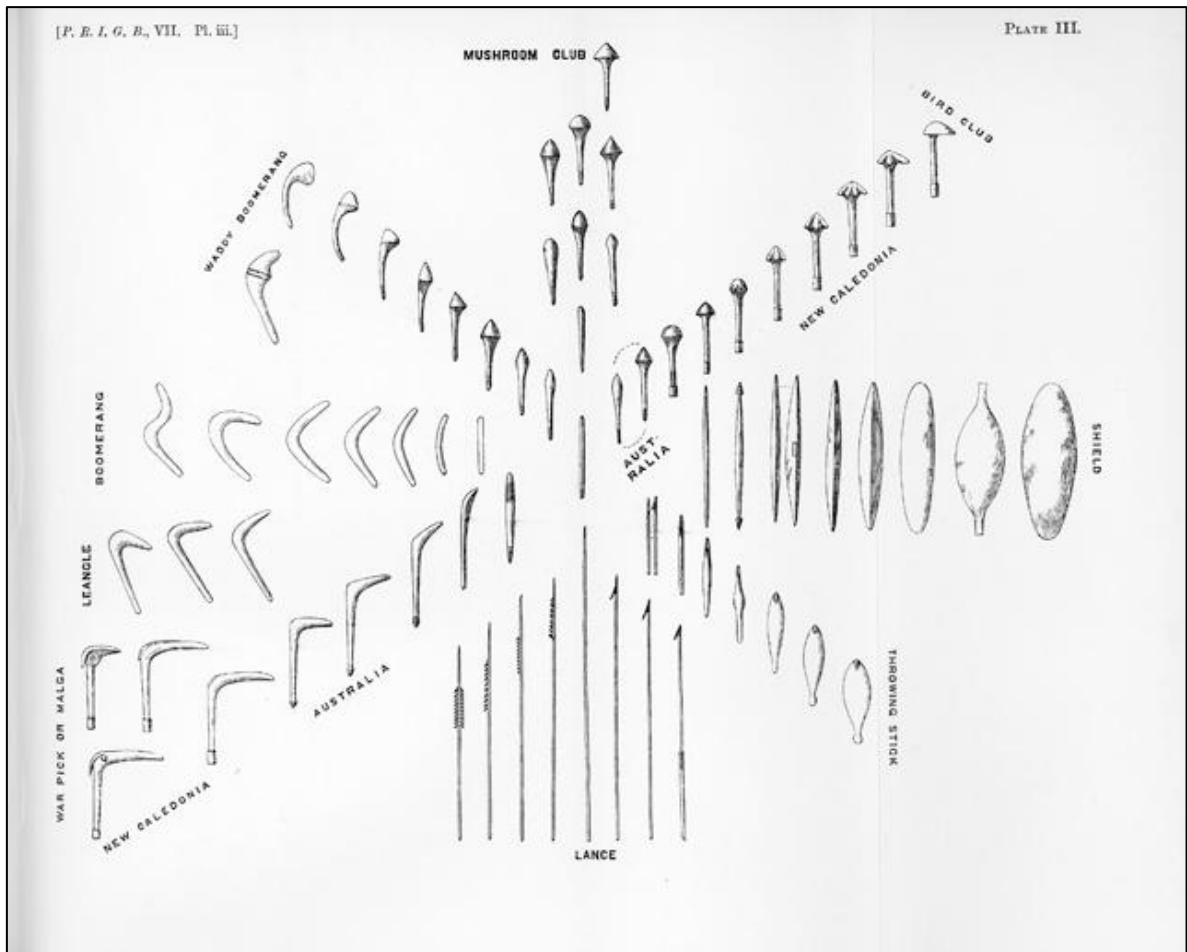


**Figure 1.1. The S-Curve theory of technological advancement**<sup>14</sup>

The logistic model is not without its critics. A historical review of performance trends in computer memory, desktop printers and display monitors led Sood and Tellis<sup>15</sup> to conclude that rather than an S-Shape, punctuated equilibrium should be modelled by a step function. Long periods of stagnation in performance are interspersed with large “jumps” in performance and innovation before returning to a state of low innovation. They also conclude that new technologies may enter above or below the performance of existing technologies and do not necessarily have only one crossing point. An old technology may enter a period of renewed innovation, placing it above the new technology in terms of performance or functionality.

The discontinuity of each new technology as demonstrated in Figure 1.1, is discounted by Basalla<sup>16</sup>. Rather than *the heroic view of invention*<sup>16 (p21)</sup> where new discoveries emerge fully formed from the minds of brilliant individuals, he suggests that technological

evolution happens in a continuous manner with each new technology emerging from its antecedents. Gilfillan provides an early example of this by tracing the evolution of early dugout canoes to modern-day ships<sup>17</sup>. Arthur supports this continuous view of development with a theory he describes as combinatorial evolution<sup>18</sup>. His basic premise is that newer technologies evolve from previous technologies through novel recombination. Arthur argues that technologies are like building blocks. They can be combined and recombined to produce improved technologies. Every now and again a novel combination of building blocks allows us to capture new phenomena. Encapsulating these building blocks into a new block allow for the possibility that a whole new range of combinations will occur. A simple analogy can be drawn with organic chemistry. Simpler molecules (hard drives, fans, microprocessors) can be put together to form new molecular combinations (the computer), which can again be combined to produce further new molecules (the Internet)<sup>18</sup>. Strong arguments are made to support both the view that change happens in disruptive bursts or that it occurs through small incremental changes. A more balanced view would be not to consider these processes as mutually exclusive but rather complementary. The implications of this idea for the work of this thesis are that any new methods that are developed for the collection of data are most likely to emerge from the novel combination of already existing technologies and methods (see Figure 1.2).

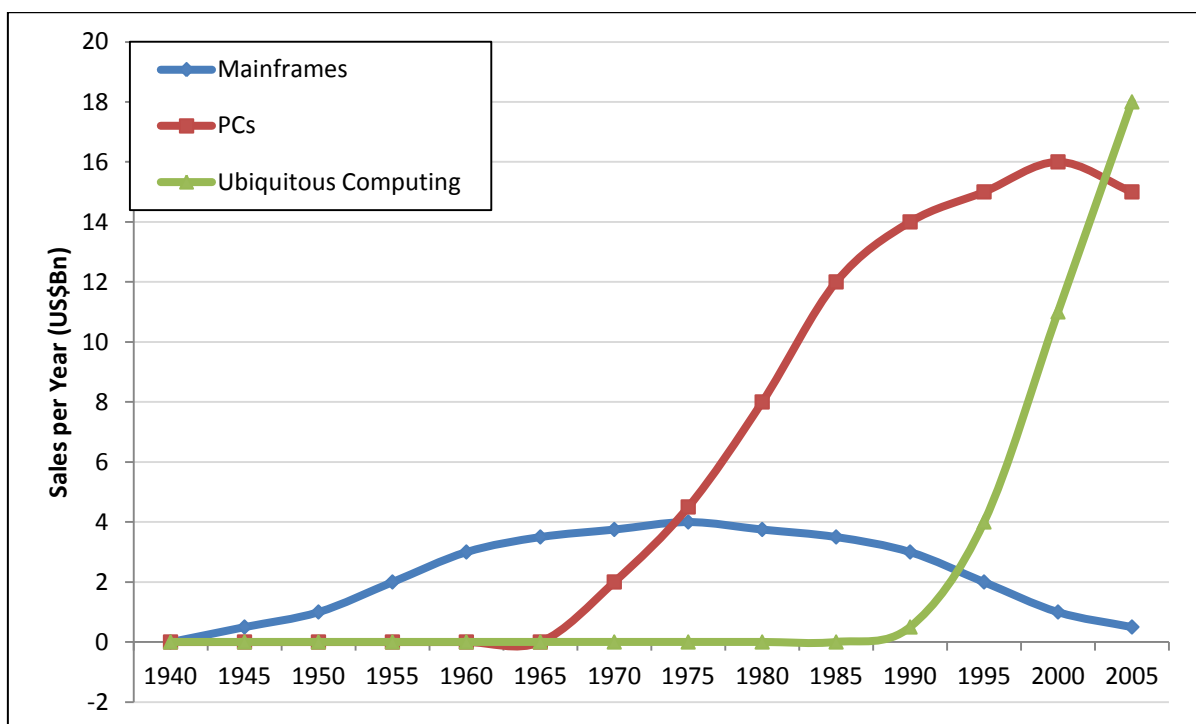


**Figure 1.2. The continuous evolution of Australian aboriginal weapons as arranged by Pitt-Rivers (1906). Radiating out from the simple stick are war clubs, boomerangs, lances, throwing sticks and shields<sup>16</sup>**

In summary, over a relatively short period of historical time technology has progressed from simple tools to complicated automatic machines. This process has occurred through four main mechanisms: 1) standard engineering, i.e. small, incremental fixes and improvements; 2) the invention of fundamentally new technologies; 3) structural deepening, i.e. basic technology adding new features over time (for example, an automobile costs about the same now (in real terms) as when it was invented but what you get for your money is much greater); 4) an emerging body of technologies building up over time and transforming the industries that give rise to them<sup>18</sup>.

### 1.1.2 The three revolutions of computing

Having outlined the thinking on evolutionary theory of general technological progress, specific attention is now paid to how this theory has played out in the development of computing devices. The history of computing, culminating in computers being an everyday tool in many people's personal and professional lives, can be broadly described in terms of three distinct waves (Figure 1.3).



**Figure 1.3. The three revolutions of computing. Adapted from Weiser<sup>19</sup>**

The first wave of computing, from the 1940s to the beginning of the 1980s, comprised of large room-size mainframe computers designed to be used by a number of people. This one-machine, many-users paradigm began to shift in the 1980s with the arrival of the personal desktop computer. Decreasing hardware size and availability at consumer-level prices meant that each individual could now own and interact with his or her own computer (one-machine, one-user). Not everyone was convinced that this new paradigm

had merit. As Ken Olson, one time president, chairman and founder of Digital Equipment Corporation now infamously quipped in 1977, *There is no reason anyone would want a computer in their home*<sup>20(p8)</sup>. As it turned out, he and many others were wrong, and were overtaken by the second wave of computing.

Like the shift from the first to the second wave, there was much debate as to what would succeed desktop computers. In the 1980s two views prevailed. The first view was that computers, through improved software, would become *intelligent agents* with whom we interacted directly<sup>21,22</sup>. Using the metaphor of a smart robotic butler, this vision for the future was one where computers retained their physical presence and allowed us to interact intelligently with them in a one-on-one manner. This view was supported by those who described machines of the future with phrases like *intimate computer*<sup>23</sup>, or *rather like a human assistant*<sup>24</sup>. The second view, favoured by Mark Weiser, the father of pervasive computing, was that computers would physically disappear from our awareness. He hypothesised that miniature computers would proliferate, becoming embedded in everyday objects such as clothes, beds, pens and shoes<sup>25</sup>. These devices would work non-intrusively and invisibly throughout our physical environment, collecting data and providing services. Computers would no longer be disconnected tools confined to a desktop but would rather become networked and pervasive, and would penetrate the environment around us. Thus in Weiser's view, the third wave of computing would come to be represented by many computers being used by one individual. Weiser thus argued that the defining words of the third wave would not be *intelligent* or *agent*, but rather *invisible* and *connection*<sup>21</sup>.

Although used widely and interchangeably, there is a subtle difference in the concepts put forward by ubiquitous and pervasive computing. Although they share the same goal of having computers available *everywhere at every time*,<sup>26</sup> the primary difference is that pervasive computing more often refers to devices carried on one's person, like the mobile phones of today. Ubiquitous computing in its strictest sense refers to a world in which we interact with devices embedded in the world around us.

### **1.1.3 Pervasive computing**

Although ongoing research supports both the *intelligent agent* and *ubiquitous computing* views, the conceptual paradigm of pervasive computing is of greatest relevance to this dissertation. Consideration of anthropological literature on people's work life led Weiser to the conclusion that people work mainly in a collaborative environment on shared problems with computer interaction at the forefront of attention. These computer interactions were often found to get in the way of work<sup>25</sup> with people spending more time on trying to get the technology to work than on the actual work itself. Weiser suggested that as devices became smaller, more mobile and less expensive, people would stop being aware that they were working on a computer. Instead, people would interact with computers without ever becoming aware of the computers' presence. He maintained that the goal of pervasive computing would be achieved when users were no longer aware of the fact that they were interacting with a computer. At the All Things Digital conference in 2010, Steve Jobs presented the following analogy of the current shift beyond the desktop computer. *When we were an agrarian nation, all cars were trucks. But as people moved more towards urban centres, people started to get into cars. I think PCs are going to be like trucks. Less people will need them*<sup>27</sup>. The maturation and

availability of two technologies, distributed computer networking and mobile computing, are at the heart of the shift away from the desktop model.

A combination of the two has enabled the move towards a pervasive computing paradigm. Distributed computer networking allows networked computers to share resources, access remote information and communicate with other connected computers, whether by wire or wireless networks. Through a simple mechanism designed to locate resources available on the network, the web has integrated information distributed between many computers into a single addressable structure. Distributed computing itself relies on the technologies of remote communication – having to device communicate with each other from a distance, fault tolerance – techniques for detecting the inaccurate transmission of information and resending a corrected copy, high availability – increasing the amount of time a machine is responsive and functional, remote information access and distributed security for communication. The predominance of the World Wide Web has seen it become the primary network through which devices connect with each other<sup>28,29</sup>.

The emergence of portable devices such as laptops and early mobile phones linked to wireless networks produced a new set of challenges not faced by desk-bound computers networked together by cable. In particular, four key challenges are introduced by device mobility<sup>30,31</sup>.

- **Mobile systems are resource-limited**

Mobile devices are both smaller in dimension and lighter in weight than desktop machines. These physical limitations reduce the computational resources of the device. Although processor speed, memory resources and disk size will continue

to improve on mobile devices, they will always be relatively limited when compared with desktop computers.

- **Mobile systems have a finite energy source**

Large displays, background application tasks and sensors built into mobile devices all require energy. Relying on the finite source of energy provided by battery power has implications for both hardware and software design.

- **Mobile systems have variable connectivity**

Network strength and availability are variable for non-wired devices. In some cases, users may choose to turn off network connectivity. Managing this variability is an important issue in mobile computing.

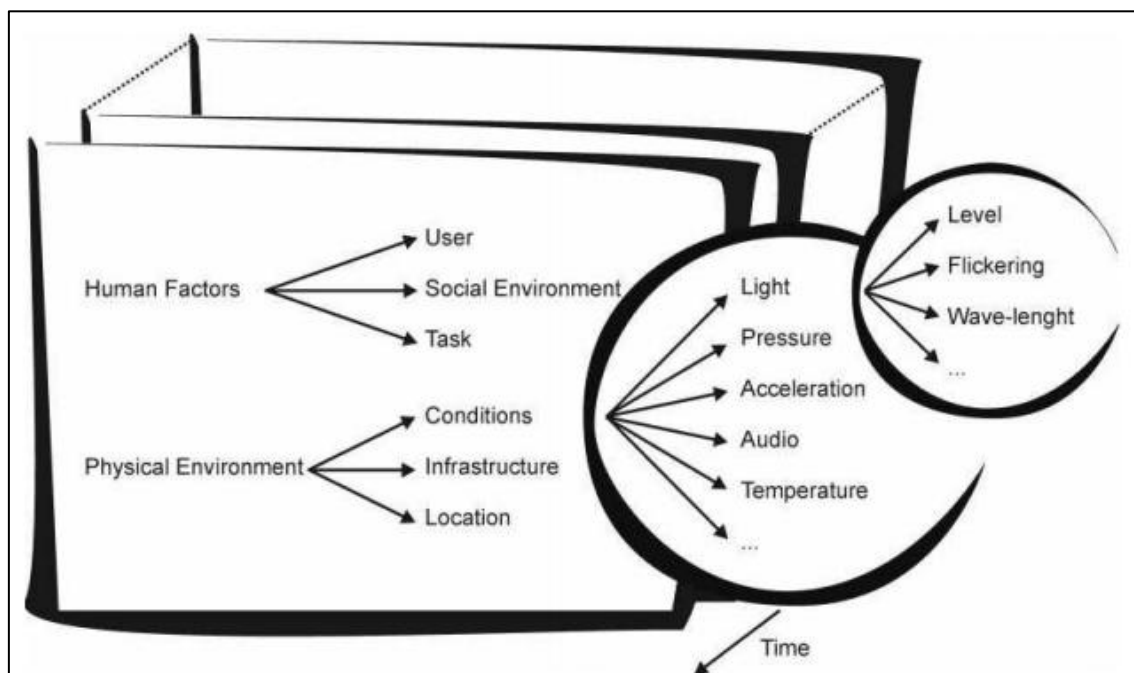
- **Mobile systems are less secure**

Desktop computers are inherently more secure than mobile devices as physical access has to be gained to the building in which the machine is housed. With mobile devices, the risk is much higher for unintended access to information through loss or theft, due to the open environment in which they are used.

In addition to the technologies associated with distributed networking, mobile computing relies on five further innovations to address the issues listed above. Broadly these can be described as mobile networking, mobile information access, support for adaptive applications, system-level energy saving and location sensitivity<sup>29</sup>. Together the advances in distributed networking and mobile computing have made portable devices such as mobile phones and tablets possible, while at the same time hiding much of the complexity involved. The possibilities offered and challenges faced by devices that are

mobile, simple to use and widely available in our environment are the concern of pervasive computing researchers.

One area in which pervasive computing has expanded on previous work is the notion of a context-aware device<sup>32</sup>. Due to the increased mobility of devices, the context in which they are used changes constantly. To best support the user in these changing contexts, a device automatically needs to be able to discover and track changes to people, objects and resources in close proximity to it<sup>33</sup>. Put simply, devices must “gain awareness” as to where you are, who you are with and what resources you are near. Schmidt<sup>34</sup> emphasises that context is more than simply knowing the device’s physical location. Figure 1.4 depicts his working model for context. It includes both physical environment (conditions, surrounding infrastructure and location) and human factors (user, social environment and task). In Schmidt’s model, a fully described local context can be achieved through the monitoring of both human and physical factors.



**Figure 1.4. A working model of how context can be ascertained by mobile devices<sup>34</sup>**

- **Human contextual factors**

Human factors can be classified into three categories: user information, user's social environment and the user's tasks. User information describes the user's habits, emotional state and physiological condition. Knowledge of a user's social environment would allow the device to monitor who else is around, ongoing social interaction between users, and group dynamics. The final category allows the device to determine what tasks the user is engaging in.

- **Physical environment**

The physical environment is also important in relation to context. The three environmental factors are location, awareness of surrounding resources and awareness of physical conditions. Location is critical to context, and it includes the user's absolute geographic position and his or her position relative to other people and resources. This positioning information facilitates an awareness of surrounding resources and infrastructure. For example, if a wireless network is available, the device should become aware of it. Finally, the device should be aware of physical conditions such as sound, light, movement and temperature.

The development of inexpensive, portable and powerful devices open many new opportunities within the field health research. The term *health research*, otherwise known as *medical or clinical research*, refers to research that is done to learn more about human health. Health research aims to better understand how to prevent and treat disease. Health research is an important way to help improve the care and treatment of

people worldwide. The relevance of pervasive computing and contextually aware devices such as mobile phones for health research can first be seen in the work of Pentland and one of his PhD students, Nathan Eagle. Coining the term *reality mining*, Pentland realised that the data generated by the rise of pervasive computing may reveal interesting patterns and insights about people and communities. Eagle defines reality mining as

*the collection of machine-sensed environmental data pertaining to human social behavior. This new paradigm of data mining makes possible the modeling of conversation context, proximity sensing, and temporospatial location throughout large communities of individuals.*<sup>35</sup> (paragraph 1)

Mobile phones are currently the most valuable and widely available source for acquiring both user initiated (cell, SMS) and non-initiated (on-board sensors) data about a user's physical and logical context<sup>36</sup>. There are three reasons for this. Firstly, sensors capable of recording both physical and logical context are commonly available in mobile phones through the inclusion of cameras, microphones, light sensors, accelerometers and Global Positioning System (GPS) chips<sup>37-39</sup>. Data detected by these sensors can be queried to gain insight into human behaviour. The second factor that lends mobile phones to the task of acquiring user data is their mobility<sup>40</sup>. Finally mobile phones are widely available in modern society.

The current challenge for researchers in this field is how to take the flood of discrete data events emanating from the multitude of mobile sensors and transform them into a meaningful real-world event. The example given by Ye<sup>41</sup> is the process of integrating

individual pieces of sensor data about heart rate, body position, galvanic skin response and blood pressure as they relate to the situation *heart attack*.

#### **1.1.4 Summary**

This section introduced how technology advanced through small incremental combinations of existing ideas and techniques. The theory of technological evolution was described in some detail. This theory was used to underpin the three evolutionary stages computers have undergone in the past 70 years. Early mainframes where one machine served many people gave way in the 1980s to personal desktop computers used in a one-on-one manner. Now, as we enter the third stage, devices will become even more ubiquitous and pervasive in our environment. Individuals will each be served by a large number of computing devices. The section concluded with a discussion on the importance of pervasive devices being “context-aware” and how this awareness will play out with respect to the most common pervasive device available at present — the mobile phone.

Running in parallel to the opportunities presented by each new wave of technology are the traditional challenges faced by researchers interested in understanding population-level health outcomes when collecting large-scale survey data. Managing geographically dispersed field teams, keeping track of surveys from interview through to data capture and implementing effective quality-assurance checks are some of the familiar challenges faced by these researchers. As new technology has become available, researchers have moved quickly to innovate in the hope of expediting the process and improving data quality. With advancements in mobile technology, this study implemented and evaluated

the use of mobile phones as research instruments with which to collect and analyse information for research in low-resource Sub-Saharan African settings.

## **Chapter 2**

---

### **Literature Review**

### 2.1 Theoretical Framework for Quantitative Survey Data Collection

*The movement to computer-based survey methods is not an option. It seems as inexorable as the transition to computers in most other organized human activities in modern society. The issue is not whether this change will occur but what its effects will be. – Mick Couper<sup>42</sup>*

Social, economic, public opinion and health research all share the survey questionnaire as a primary data-collection tool. The field of survey methodology is broad in its scope. The primary methodological concern is identifying and minimising sources of error in order to improve the accuracy and validity of responses. Questionnaire design, sources of survey error, sampling frames, data-collection methods and post-collection processing of data are all of interest to researchers in the field. This section introduces the concepts relevant to understanding the impact that the choice of data-collection method has on the quality of the survey response. The conceptual frameworks introduced are drawn upon in the discussion to evaluate how the use of mobile phones in health survey research impacts on data quality. The section ends with an examination of how mobile phones, as extensions of computer-aided personal interviewing, compare with other better known methods of data collection.

### 2.1.1 The Total Survey Error framework

Up until the late 19th century very little attention was given to the wording and structuring of health research surveys. Armed with a set of objectives such as age, occupation and education, researchers would create their questions during face-to-face interviews with participants. Standardised questioning came about as the result of work performed by psychometricians interested in quantifying human intelligence. Still, until the advent of the Likert scale<sup>43</sup>, surveys were long, cumbersome and time-consuming to administer<sup>44</sup>.

Despite question design having remained more or less unchanged over the past 80 years, many new data-collection methods have emerged during this time. In the mid-1940s, Deming<sup>45</sup> was one of the first to note the influence that these new methods of data collection were having on survey responses. In a list of 13 factors influencing the usefulness of a survey, Deming suggests that *the problem is not whether differences exist [between data collection methods] but how great the differences are, and why do they exist* (p362). Deming's paper was one of the first to recognise and acknowledge that surveys included both sampling and non-sampling error. In 1979, Anderson, Kasper, and Frankel<sup>46</sup> introduced the Total Survey Error (TSE) framework. The framework attempts to include the impact of a variety of error sources on survey estimates and is currently the dominant paradigm within the field of survey methodology<sup>47,48</sup>. The TSE paradigm aims to facilitate the optimal allocation of resources in such a way as to minimise total survey error. This is achieved in practice by firstly trying to identify major sources of error that may have been introduced through research and survey design. Study resources can then be optimally allocated to activities that reduce these errors, while remaining within

specified cost limitations and data-collection timelines<sup>44,47</sup>. The sources of error that make up TSE have been classified in various ways by different researchers (in the context of TSE, errors need to be understood as differences between the obtained value and the true value, rather than *mistakes*<sup>49</sup>). In the first edition of *Survey Errors and Survey Costs* (2004), Groves<sup>44</sup>, building on the work of Kish<sup>50</sup> and Anderson et al.<sup>46</sup>, identifies four types of error: coverage error, non-response error, sampling error and measurement error. In a second edition, Groves adds processing error and validity to his typology.

- **Coverage error**

Coverage errors, non-response errors and sampling errors are three types of error that can be introduced through non-observation. Coverage errors occur when not everyone in the population of interest is given a chance to be selected into the sample. Such omissions have an impact on the ability to generalise the results obtained<sup>49,51</sup> as the sample is no longer representative. Obtaining national estimates from a web survey, where the online population is known not to approximate the general population, would introduce such an error.

- **Non-response error**

Two types of non-response error have been identified. The first, unit-level non-response, occurs when participants in a randomly selected sample choose not to participate in a non-random way. When actual respondents vary systematically from those of the initial sample, the respondents may not be representative and may introduce a survey bias. Fricker et al.<sup>52</sup> noted such a finding in a 2005 mode comparison study. The responding portion of the sample contained a higher proportion of highly educated, middle-aged female participants than the non-responding portion. In contrast to unit-level non-response, item-level non-

response occurs through the selective skipping of certain survey items by a homogenous group within the sample.

- **Sampling**

The most widely known source of survey errors is sampling error. Sampling errors arise when resources are not sufficient to survey the entire population of interest. Controlling for cost, a sample of the population is selected. If the sample is not completely representative, sampling error arises. This error can be calculated for randomly selected probability samples, but is not measurable in non-probability samples (as are most online surveys)<sup>49</sup>.

- **Measurement error**

Measurement error relates to any inaccuracy between the true value and the recorded value that arises during the measurement process. These errors can be introduced by the survey instrument being poorly designed, worded or laid out; by the effect the interviewer has on the participant; by respondents being unwilling or incapable of responding clearly and truthfully; and finally, by the mode of data collection used. The mode of data collection is an important factor as it influences both the presentation and the role of the interviewer<sup>49</sup>. Theory that underpins how the mode of data collection exerts its influence will be examined in more detail shortly.

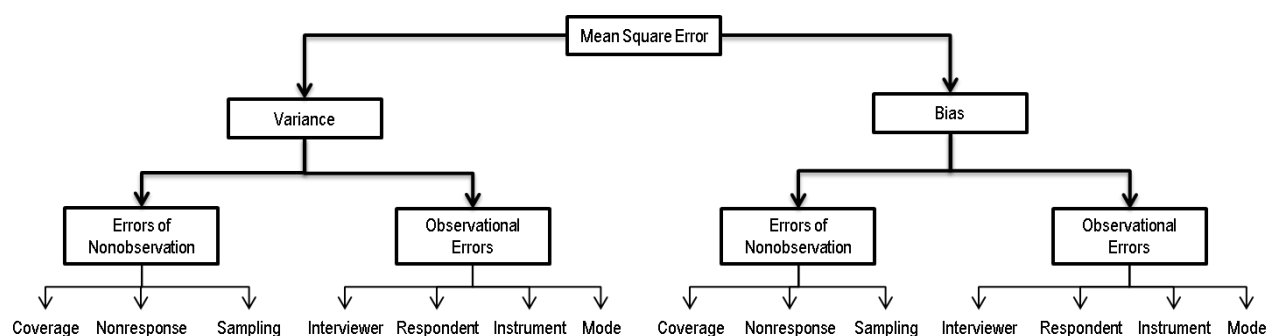
- **Post-survey error**

Not included in early TSE typologies, post-survey errors are important and their presence and impact should not be underestimated. Errors of this type can be introduced at data-processing and survey-reporting stages<sup>49</sup>. During data

processing, editing and cleaning should be done carefully so as not to “over clean” the data and thereby introduce unwanted errors.

- **Validity**

In his initial conceptualisation (see Figure 2.1), Groves divides all errors into those of variance and bias. Each is then subdivided into errors of observation and errors of non-observation. The final level of disaggregation is covered by the classification detailed above. In his later work, Groves reconstructs the concept of variance as *validity*, drawing on the language of psychometrics<sup>48</sup>. In this context, validity is used to refer to the gap introduced between construct and measurement. For example, if a researcher is interested in the abstract concept of artistic ability, then a great deal of effort would need to go towards creating a survey instrument capable of fully capturing this ability. If the final measurement does not adequately measure artistic ability, a bias is introduced. The test is said to lack construct validity<sup>53</sup>.



**Figure 2.1. Groves’s (1989) conceptual framework for calculating total survey error (accuracy)<sup>51</sup>**

### 2.1.2 Multidimensional frameworks for assessing data quality

Despite the widespread use of the TSE framework, it has been criticised for not taking a broad enough perspective on what contributes to survey quality. In addition to error, the relevance of the information (the degree to which data meets the needs of its end user) and the credibility of the data-collection process are equally important<sup>48</sup>. Juran and Gryna<sup>54</sup> were among the first to consider data quality in this wider context. They advanced the idea that data quality should take into account its *fitness for use*. This implies that, depending on the intended use of data, different perspectives may be held on what quality means. Biemer<sup>47</sup> advances this thesis by suggesting that data users and data producers hold different notions of what constitutes *data quality*. For data producers, data accuracy is often a key concern. A large proportion of resources is therefore often allocated to improving sample size, coverage, response rates and other drivers of data accuracy (through the reduction of errors). For users, the usefulness of the data is often paramount. So from this perspective, timeliness, supporting documentation and accessibility are considered important elements of high-quality data.

The TSE only considers data quality from a producer perspective. Biemer and Lyberg<sup>55</sup> reason that the TSE framework needs to be replaced by a more multidimensional quality model that focuses on measures for assessing quality with respect to the fitness of data to its intended use by data users. By concentrating on the fitness of data for its intended use, study design could strive to meet the requirements of users, both with respect to the accuracy of data and in relation to other dimensions of quality (e.g. relevance, interpretability, data security and timeliness). Through this lens, the relative merits of competing data-quality dimensions would need to be assessed in the context of the

study. Ensuring quality on one dimension (completeness) may be at odds with ensuring quality on a second dimension (timeliness). There are a number of multidimensional quality frameworks that have emerged to address these concerns with TSE<sup>56–60</sup>. Common dimensions along which data quality is measured are listed in Table 2.1.

**Table 2.1. Common dimensions that impact survey data quality**<sup>56–60</sup>

| Dimension        | Description                                                                                                                                                                                          |
|------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Believability    | Data collected by a credible institution are considered to be trustworthy.                                                                                                                           |
| Accuracy         | The total survey error (TSE). A lower TSE is synonymous with improved accuracy.                                                                                                                      |
| Relevance        | Data are relevant and fit for its intended use.                                                                                                                                                      |
| Timeliness       | Data are available in a timeframe that is sensitive to the urgency with which data is needed.                                                                                                        |
| Completeness     | Surveys are broad enough in range to allow key objectives to be addressed while at the same time not overburdening participants with irrelevant items not likely ever to be included in an analysis. |
| Interpretability | Good documentation and clear metadata are provided.                                                                                                                                                  |
| Accessibility    | Suitable data-curation procedures are applied in such a way as to make data access user-friendly.                                                                                                    |
| Data Security    | Data are kept in a secure environment with participant identifiers stored separately.                                                                                                                |

Maximising data quality is always counterbalanced by the need to limit survey costs<sup>44</sup>.

One of the key decisions affecting the cost of a survey is the method used to collect the data. In this respect, researchers have a wide and ever-increasing range of methods from which to choose. Traditionally, the most commonly used methods or ‘modes’ of data collection have included face-to-face, fixed-line telephone and mail. The evolution of mainframe to desktop computing (described previously) allowed a new range of complementary computerised techniques to emerge that are collectively known as computer-assisted survey information collection (CASIC). Among the most widely researched and adopted of these techniques are computer-assisted personal interviewing

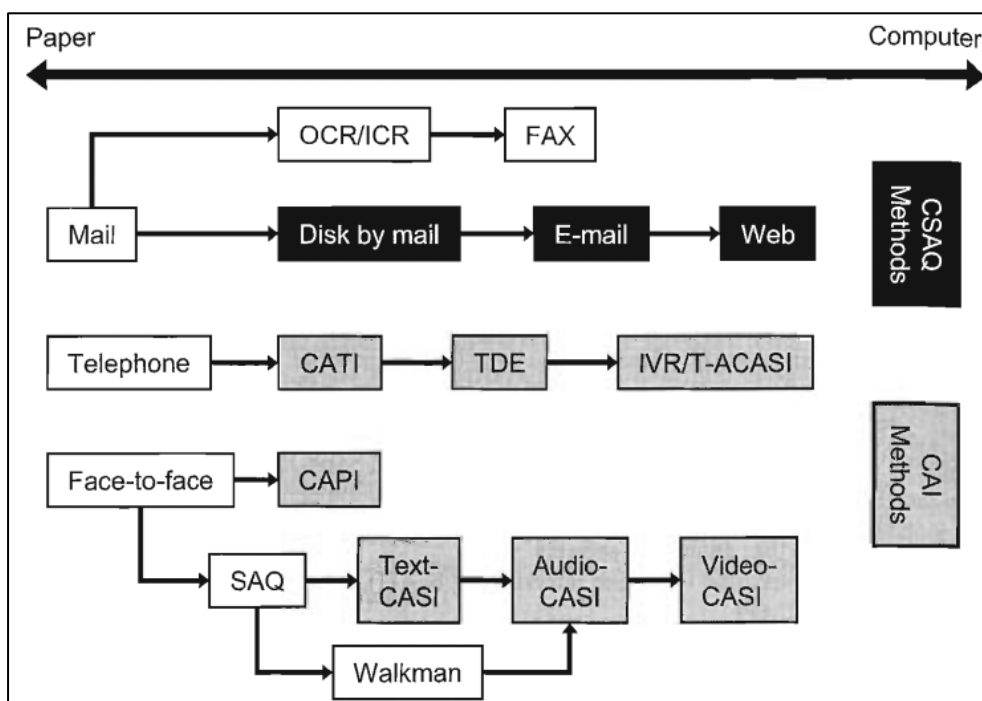
(CAPI), computer-assisted self-interviewing (CASI) and computer-assisted telephonic interviewing (CATI)<sup>61</sup>.

The rise of the Internet spawned its own set of techniques. Jointly known as online surveys, they include web, email and social network-based surveys<sup>62,63</sup>. The third wave of development in computer technology has ushered in a third set of techniques for collecting survey data. The key distinguishing feature of these methods is that they embrace the mobility of pervasive devices<sup>62</sup>. It is into this category that the mode of interest in this thesis falls, namely mobile-phone interviewing. By leveraging one of the many communication channels available on mobile phones (e.g. text message, voice, Internet survey and software application), it is possible to collect survey data. It is well established that the data-collection method can affect the quality of the responses received. To understand the implications for mobile-phone surveys, the existing theory on the impact of the mode of data collection on data quality is reviewed below.

### **2.1.3 Theoretical impact of survey mode on survey error and data quality**

Survey data-collection modes vary in the method used to contact participants, in the presentation of the questionnaire, and in the way in which questions are administered. These variations give each data-collection mode a unique set of advantages, disadvantages and costs. They also introduce the theoretical question of whether or not response is affected by mode. The answer to this question has been studied extensively<sup>64–73</sup>. A brief summary of these findings is presented in the next section. A drawback to only performing comparison studies between two different modes of data collection is that the continuous proliferation of new techniques means constantly assessing the impact of these new methods on data accuracy, quality and costs. An alternative, complementary

approach is the use of a theoretical model to inform expectations of how a new method might affect the data collected. Groves et al.<sup>51</sup> provide a useful taxonomy of the modes available in 2004 (Figure 2.2). The diagram promotes a better understanding of how new methods have arisen from the adaptation of older methods. Groves identifies mail, telephone and face-to-face interviewing as the cornerstone methods. This would imply that what we know about the impact of telephone surveys on data accuracy (e.g. sample coverage) could equally be applied to interactive voice response (IVR) data-collection systems.



**Figure 2.2.<sup>1</sup> The evolution of survey technology from the use of paper to computers<sup>51</sup>**

<sup>1</sup> OCR/ICR — Optical character recognition / Intelligent character recognition

CSAQ — Computerized self-administered questionnaires

CASI — Computer-assisted self-interviewing

CATI — Computer-assisted telephone interviewing

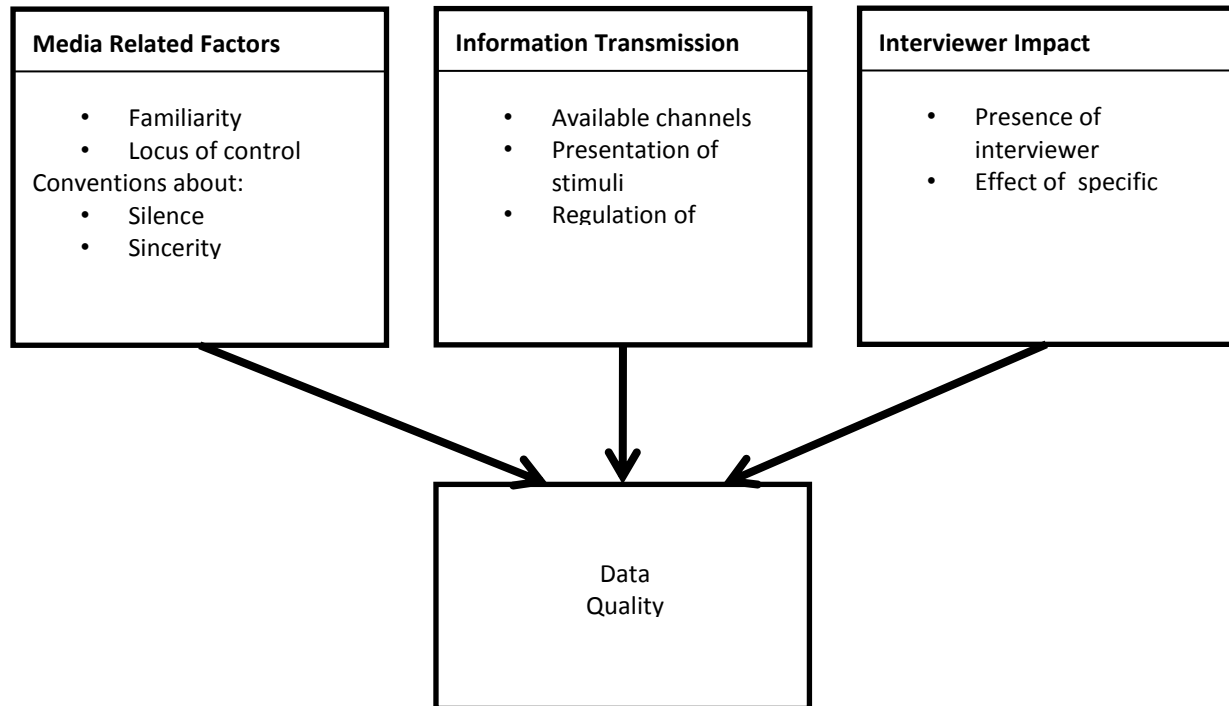
TDE — Touchtone data entry

IVR/T-ACASI — Interactive voice response / Telephone-audio computer-assisted self-interviewing

CAPI — Computer-assisted personal interviewing

SAQ — Self-administered questionnaires

Although useful, the approach is limited in its application as it does not make clear where new methods should be situated in this taxonomy. De Leeuw<sup>72</sup> suggests a more generalisable model by defining three characteristics that differ by mode (Figure 2.3).



**Figure 2.3. De Leeuw's model of mode characteristics influencing data quality** <sup>72</sup> (p19)

- **Media-related factors**

Communication channels are each governed by different sets of social conventions. Thus, receiving an email may be met with the expectation of a prompt reply, while a snail-mail letter may not. These conventions can shift through time and are related to the familiarity people have with the medium. De Leeuw argues that the channel of communication also mediates the acceptability of long silences and the sincerity with which people receive a request for participation. Finally, locus-of-control is affected by the communication channel

chosen. In a mail survey, locus-of-control rests completely with the participant, while a shared locus-of-control exists in a face-to-face interview.

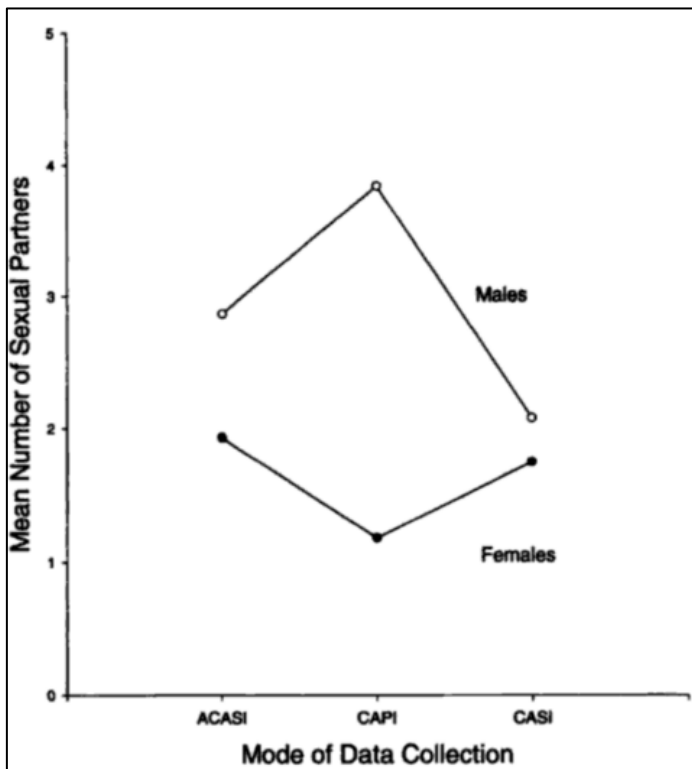
- **Information transmission**

In addition to social conventions associated with the communication channel used, the second component of this model focuses on technical features of the mode in question. All communication happens through one of three channels: verbal, non-verbal and paralinguistic (emotional tone, timing of speech, emphasis, etc.). Depending on the choice of mode, certain information from one of these three channels may be lost. Further, in some modes the order in which questions are presented can be controlled by the participant, while in other situations the order is carefully controlled by the interviewer. In a similar vein, regulation of communication flow can be managed in some modes but not in others. Collectively, these features all affect the quantity and quality of the information that can be transmitted.

- **Interviewer impact**

The importance of the presence or absence of an interviewer is further highlighted in this model by the inclusion of a factor relating exclusively to the impact of the interviewer. The distinction between interviews that include an interviewer and those that don't is known as "interviewer-led" versus "self-complete" surveys. Some studies have found the presence of an interviewer to result in underreporting of sensitive information<sup>74,75</sup>. Figure 2.4 illustrates this concept using data from Tourangeau et al.<sup>76</sup>. Using three different modes (ACASI, CAPI and CASI), two of which were "self-complete" (ACASI and CASI), a significant difference

between the responses of men and women was noted in the “interviewer-led” CAPI survey. The authors offer a number of possible reasons for this finding.

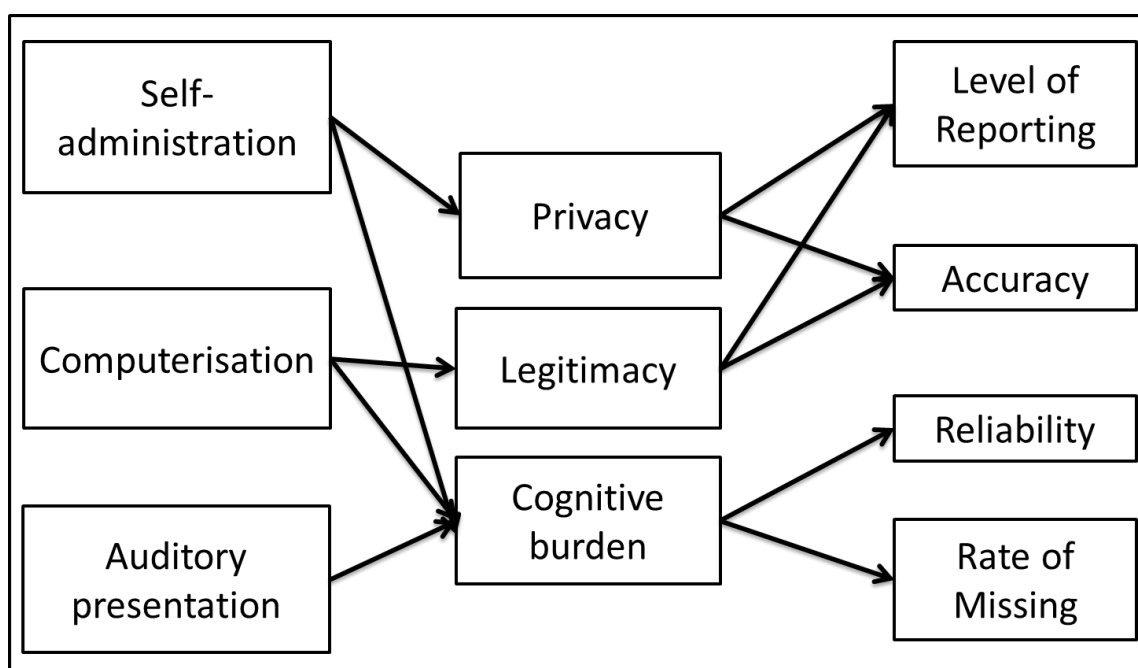


**Figure 2.4. Impact of mode and interviewer presence on mean number of reported sexual partners**<sup>76(p277)</sup>

Tourangeau<sup>76</sup> suggests the evaluation of new methods along five key dimensions: how respondents were contacted, the presentational medium, the method of administration, the sensory input channel, and the response mode. By varying one or more of these factors in combination with classic face-to-face, mail and telephone surveys, it is possible to describe and categorise the many data-collection modes available at present. Groves et al.<sup>51</sup> also presents a five-feature model according to which mode comparisons can be made: degree of interviewer involvement, level of integration with respondent, degree of privacy, degree of technology and, like De Leeuw, channels of communication. Although both models offer similar insights into the key features differentiating between data-

collection modes, neither provides a clear framework for unpacking the influence of mode on response quality.

To answer this question and better understand how mode affects the process of hearing a question, formulating a response and answering, an information-processing model can be used. Summarising seven such models, Jobe<sup>77</sup> distils four discrete processing stages: comprehension of the question; retrieval of the answer from autobiographical memory; use of heuristic and decision processes to estimate the answer; and response formulation<sup>77</sup>. Using this generic information-processing model, Tourangeau et al.<sup>76</sup> suggests the framework presented in Figure 2.5 as a way of tying mode (left-hand column), information processing (middle column) and data quality (right-hand column) into a unified conceptual framework.



**Figure 2.5. Tourangeau et al.'s model of how psychological variables, mediated by mode, affect data quality<sup>76</sup>**

The model highlights the role of three key mode features that influence a person's psychological appraisal, i.e. a person's evaluation or appraisal of an event or situation will cause specific reactions. The psychological appraisal of a situation will cause an emotional, or affective, response based on that appraisal. In this model these situations are:

- Whether or not the questionnaire is interviewer- or self-administered
- Whether or not a computer-assisted survey mode is used, and
- Whether or not the questions are presented aurally or visually.

Together these mode features influence the psychological appraisal of the question being asked. This psychological processing includes:

- *The degree to which the method promotes a sense of impersonality or privacy.* A survey that is self-administered increases a respondent's willingness to report sensitive information as they experience the interaction as more private than disclosing to an individual. This response pattern may be driven by the fear of being embarrassed<sup>78</sup>.
- *The perceived sense of the legitimacy of the survey that the choice of mode creates for respondents.* The method of data collection can affect this sense of legitimacy in two ways. If the initial contact is not done in person, the legitimacy of the survey is harder to establish as respondents will not have access to the visual cues available in a face-to-face meeting (e.g. uniform, letterheads and ethics committee stamps). Computerisation may lend further legitimacy to the study because people tend to assess things that they do not fully understand favourably.

- *The cognitive burden placed on respondents in order to formulate a response.*

Different modes require varying degrees of effort. Reading, listening, keying, or navigating a web page all place a cognitive load on the respondent. Cognitive burden can also be increased by poorly designed surveys that make question comprehension or answer retrieval difficult.

Finally, the model enumerates four indicators of data quality and postulates how the cognitive tasks associated with responding can influence levels of reporting, accuracy, reliability and the rate of missing values. Impersonality and legitimacy are hypothesised to influence survey response rates and data accuracy, while cognitive burden decreased survey quality through its interplay with the reliability of responses and item non-response rates<sup>76</sup>. Robling et al.<sup>79</sup> recently made some notable additions to this model. The first amendment they suggest is to separate the role of psychological appraisal (impersonality, legitimacy and cognitive burden) from response. The authors focus their attention on three general theories of response formation.

- *Taking the easy way out*

Optimizing and *satisficing* define the two ends on the “thoughtfulness of response” continuum. Depending on the complexity of the question or survey, respondents may be less diligent about the effort they put into information retrieval. This leads to incomplete responses and is described formally as satisficing. Satisficing is a decision-making strategy that aims to achieve only satisfactory results because the satisfactory position is familiar, hassle-free, and secure, whereas aiming for the best-achievable result would call for costs, effort, and incurring of risks. Three factors mediate satisficing: respondent ability,

respondent motivation and task difficulty<sup>80</sup>. The mode of data collection interacts with this continuum through the cognitive burden it places on people to respond. Dillman and Smith<sup>81</sup> report that an online web survey study increased the percentage of correctly formatted date responses from 45% to 96% by changing the way the question was presented. Though question layout, the preferred format reduced cognitive load and led to improved data quality.

- *Social Desirability*

Mode effect, in particular the presence of an interviewer, is well known to affect the social desirability of responses. This tendency for individuals to present themselves in a socially desirable way is particularly pronounced when asking about sensitive information.

- *Acquiescence*

Dichotomous questions (e.g. yes/no, agree/disagree) are popular among survey designers. They are, however, controversial as this type of question format is known to increase acquiescence, defined as the tendency to agree with a statement. Acquiescence is effected by personality, culture and educational level<sup>82</sup>.

#### **2.1.4 Practical implications of the theoretical frameworks**

To explore the practical implications of Tourangeau et al.'s conceptual model, a brief overview of relevant comparative studies is provided next. Specific attention is given to survey administration, response rates and survey mode characteristics, in particular item non-response.

##### *Survey administration*

Face-to-face interviewing, where an interviewer uses paper and pen to record the responses of an interviewee, predates all other survey modes. Through time, face-to-face interviewing has forked into two main branches. Computer-assisted personal interviewing (CAPI) is a form of face-to-face interviewing where an interviewer uses a computer as a replacement for a paper questionnaire. The interviewer reads questions off the screen and captures responses onto the computer or similar device. Computer-assisted self-interviewing (CASI, sometimes called self-administered questionnaires or SAQ) describes the second branch of face-to-face interviews. This method requires the interviewee to sit in front of a computer screen and self-complete a questionnaire. The questionnaire may be presented on the screen using text (T-CASI), audio prompts (A-CASI) or video (V-CASI)<sup>51</sup>.

Mail surveys have given rise to the second major branch of new survey modes. From computer disk by mail in the 1980s to email in the 1990s and now web surveys at the start of the 21<sup>st</sup> century, the principle of sending a survey to participants and waiting for them to return their answers endures<sup>51</sup>. Email and web surveys, which rely on the Internet, are finding increasing use, with particular importance in assessing public opinion<sup>83</sup>.

As with the other survey modes, telephonic surveys have embraced technology to produce a number of hybrid modes. Computer-assisted telephonic interviews (CATI) rely on software to display and record item responses. Touchtone data entry (TDE) and interactive voice response (IVR) systems differ only in the way data is entered into the system. With TDE, participants key in their response on the handset. Slight variations in

tone represent the key that was pressed. IVR systems use automated text-to-speech and voice recognition software to deliver and record participant responses to survey questions<sup>51</sup>.

### *Response rates*

In a 45-study meta-analysis comparing the response received using mail, telephone or face-to-face survey, Hox and De Leeuw<sup>84</sup> report an average face-to-face response rate of 70.3%. This was slightly lower than a previous estimate of 75% (sd = 18)<sup>85</sup>. Since these meta-analyses, many other studies have reported similar face-to-face response rates (for example, 70.7%<sup>86</sup>, 80%<sup>87</sup> and 74.5%<sup>88</sup>). Although often the most expensive, face-to-face interviews are considered to have the best response rate when compared to mail or telephone surveys<sup>89</sup>.

Response rates for postal mail and web-based surveys have been found to be considerably lower than face-to-face (FTF) response rates. In a recent meta-analysis of 39 web-versus-mail studies from the past 10 years, Shih and Fan<sup>90</sup> report average response rates of 45% for mail surveys and 34% for web surveys. In a second meta-analysis by Shih and Fan<sup>91</sup> in which 35 email versus paper mail comparison studies were analysed, a slightly higher response rate of 53% was found for mail survey and 33% for email surveys. These results are similar to the findings of mail-in response rates of 58% by Goyder<sup>87</sup>, email response rates of 39.6% (SD = 19.6%) found by Cook et al.<sup>92</sup> and web survey response rates of 52.5%<sup>93</sup>.

Using data collected from the 2000/2001 Funen County Health Telephonic Survey, Ekholm et al.<sup>88</sup> calculated a telephonic-survey response rate of 69.2%. This figure is in broad agreement with others (cf. 70%<sup>66</sup>, 60.2%<sup>87</sup> and 69% (sd = %19)<sup>85</sup>).

### *Item non-response*

Kaplan et al.<sup>86</sup> report that in their sample of 601 young women asked about their smoking behaviour, item non-response was 3.7% for the face-to-face (FTF) group and 17.9% for the self-administering group (SAQ). Slightly lower rates were found by Tourangeau et al.<sup>94</sup> with item non-response in PAPI (2.2%), SAQ (5.4%), CAPI (0.8%) and CASI (2.1%) when asked about abortion, sexual behaviour and drug use. Some studies report lower item non-response rates for non-sensitive FTF items (for example, 0.4% for medical questions<sup>68</sup> and 0.01% for items focused on attitudes toward immigrants<sup>93</sup>).

Item non-response, as with other self-completion techniques, has been found to be higher than FTF item non-response where an interviewer can probe for answers. O'Toole et al.<sup>68</sup> report an average item non-response rate of 5.5% over 84 questions in a mail survey. This is similar to the values reported by Harris et al.<sup>95</sup> of 7.1% (+/- 12.3%).

Heerwegh and Loosveldt<sup>93</sup> do not report on item non-response rates, but do report a breakoff rate of 16.6% for web surveys. This means that 83.4% of the 1 888 people who logged on to the survey completed the entire survey.

Item non-response for telephonic interviews has been found to be around 1%. Harris et al.<sup>95</sup> report item non-response rates of 0.7% (+/- 2.39%) in a telephonic survey with 122

participants on their experience at an inner-city hospital, while O'Toole et al.<sup>68</sup> found similarly low item non-response to a telephonic interview (0.4%).

One of the more recent iterations in survey mode has been the use of mobile phones as an instrument for the collection of survey data. Derived from the convergence of web and telephonic survey methodologies, mobile-phone surveying offers a number of alternative ways to collect data. With mobile Internet now common on all but the most basic of handsets, survey applications are the most recent way to collect data from users on the move. The rise of mobile phones in Sub-Saharan Africa will be discussed in section 2.2.

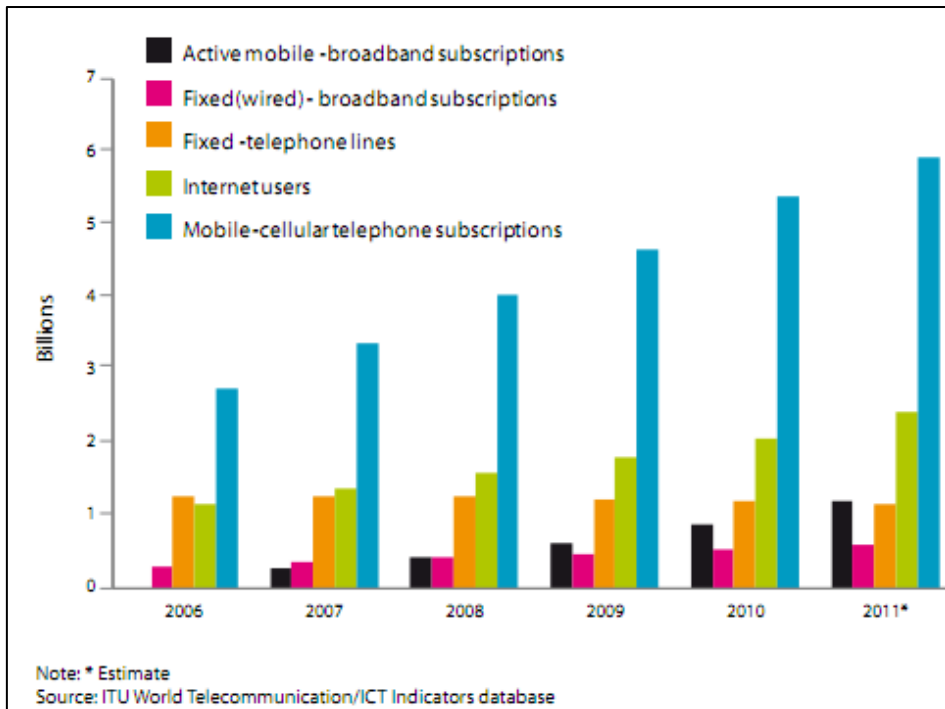
In this section, a theoretical basis has been introduced for understanding the impact that the data-collection mode has on survey error and data quality. The TSE framework was referred to as a way to enumerate the various sources of error that appear during the survey process. This framework was developed by suggesting a shift in focus from survey error to the multidimensional construct of survey data quality. The section ended with the theoretical and then the practical influence of data-collection mode on survey data quality. It is at the interface of survey error, data quality and survey mode that this thesis positions itself theoretically.

## 2.2 Mobile Phones in Sub-Saharan Africa

*Poverty is not deprivation. It is isolation.*

Malcolm Gladwell<sup>96</sup>

Mobile phones have found universal acceptance and rapid global diffusion<sup>97</sup>. As of 2011, there were only slightly fewer mobile phones than people, with six billion mobile-phone subscriptions. This gives a global mobile-phone penetration rate of 87%. Access in developing countries is not far behind this global trend, with a penetration rate of 79%. In 2011, broadband mobile Internet access surpassed fixed-line telephones. With over 90% of the world now living in a place with at least 2G mobile Internet access, the vision that Weiser proposed over 20 years ago that computing devices would become pervasive and embedded in our environments, now seems a distinct possibility.



**Figure 2.6. Worldwide growth of mobile phones and mobile broadband over the past five years<sup>97</sup>**

Figure 2.6 provides global summary statistics for information and communications technology (ICT). The key trend evident is the rise in mobile technologies and the stagnation of geographically static technologies. This is consistent with the model of pervasive computing discussed above. The diffusion of mobile communication technologies has occurred more rapidly than any other technology to date and is now an essential part of many people's day-to-day lives<sup>98</sup>. Table 2.2 disaggregates the data in Figure 2.6 by region. On average, 12 out of every 100 people living in low and middle-income countries had access to a fixed-line telephone in 2011. The average for Africa is one person in every 100 having a fixed line. In contrast, two thirds of people on the African continent have a mobile-phone subscription. The average across all developing nations is 78%. Still far behind in terms of uptake, mobile Internet access shows a similar pattern of mobile broadband overtaking wired broadband connections in both developed and developing countries.

**Table 2.2. Key global telecom indicators for the world telecommunication service sector in 2010 as provided by the International Telecommunication Union**

|                                      | (millions)  |             |             |             |             |              | Per 100 inhabitants |             |             |             |             |              |
|--------------------------------------|-------------|-------------|-------------|-------------|-------------|--------------|---------------------|-------------|-------------|-------------|-------------|--------------|
| <b>Fixed telephone lines</b>         | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> | <b>2006</b>         | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> |
| World                                | 1'261       | 1'254       | 1'250       | 1'214       | 1'189       | 1'159        | 19.2                | 18.8        | 18.6        | 17.8        | 17.2        | 16.6         |
| Developed <sup>a</sup>               | 565         | 546         | 544         | 524         | 513         | 494          | 46.5                | 44.8        | 44.5        | 42.7        | 41.6        | 39.8         |
| Developing <sup>a</sup>              | 696         | 708         | 705         | 690         | 676         | 665          | 13                  | 13          | 12.8        | 12.4        | 11.9        | 11.6         |
| Africa                               | 11          | 11          | 11          | 12          | 12          | 12           | 1.5                 | 1.5         | 1.5         | 1.6         | 1.5         | 1.4          |
| Arab States                          | 31          | 33          | 35          | 34          | 35          | 35           | 9.6                 | 10.1        | 10.3        | 9.8         | 9.8         | 9.7          |
| Asia & Pacific                       | 577         | 579         | 567         | 547         | 527         | 511          | 15.5                | 15.4        | 14.9        | 14.2        | 13.6        | 13           |
| CIS                                  | 69          | 72          | 73          | 73          | 73          | 74           | 24.7                | 25.8        | 26.1        | 26.3        | 26.2        | 26.3         |
| Europe                               | 273         | 265         | 261         | 255         | 250         | 242          | 45.3                | 43.7        | 42.9        | 41.6        | 40.7        | 39.1         |
| The Americas                         | 284         | 278         | 287         | 275         | 274         | 268          | 31.9                | 30.9        | 31.5        | 30          | 29.5        | 28.5         |
| <b>Mobile cellular subscriptions</b> | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> | <b>2006</b>         | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> |
| World                                | 2'747       | 3'370       | 4'035       | 4'650       | 5'373       | 5'981        | 41.8                | 50.6        | 59.9        | 68.3        | 78          | 86.7         |
| Developed                            | 1'127       | 1'243       | 1'328       | 1'391       | 1'408       | 1'461        | 92.9                | 102         | 108.5       | 113.2       | 114.2       | 117.8        |
| Developing                           | 1'620       | 2'127       | 2'708       | 3'259       | 3'965       | 4'520        | 30.2                | 39.1        | 49.1        | 58.4        | 70.1        | 78.8         |
| Africa                               | 129         | 174         | 246         | 296         | 360         | 433          | 17.9                | 23.5        | 32.4        | 38          | 45.2        | 53           |
| Arab States                          | 126         | 175         | 214         | 264         | 310         | 349          | 39.3                | 53          | 63.4        | 76.5        | 87.9        | 96.7         |
| Asia & Pacific                       | 1'074       | 1'398       | 1'773       | 2'161       | 2'690       | 2'897        | 28.8                | 37.1        | 46.6        | 56.2        | 69.2        | 73.9         |
| CIS                                  | 227         | 267         | 313         | 357         | 376         | 399          | 81.8                | 96.1        | 112.5       | 128         | 134.8       | 143          |
| Europe                               | 610         | 677         | 717         | 725         | 724         | 741          | 101.2               | 111.7       | 117.7       | 118.4       | 117.7       | 119.5        |

|              |     |     |     |     |     |     |    |      |      |      |      |       |
|--------------|-----|-----|-----|-----|-----|-----|----|------|------|------|------|-------|
| The Americas | 553 | 649 | 741 | 814 | 878 | 969 | 62 | 72.1 | 81.5 | 88.5 | 94.5 | 103.3 |
|--------------|-----|-----|-----|-----|-----|-----|----|------|------|------|------|-------|

| <b>Active mobile broadband subscriptions</b> |      |      |      |      |      |       |      |      |      |      |      |       |
|----------------------------------------------|------|------|------|------|------|-------|------|------|------|------|------|-------|
|                                              | 2006 | 2007 | 2008 | 2009 | 2010 | 2011* | 2006 | 2007 | 2008 | 2009 | 2010 | 2011* |
| World                                        | N/A  | 268  | 422  | 614  | 870  | 1'186 | N/A  | 4    | 6.3  | 9    | 12.6 | 17    |
| Developed                                    | N/A  | 225  | 336  | 450  | 569  | 701   | N/A  | 18.5 | 27.4 | 36.6 | 46.2 | 56.5  |
| Developing                                   | N/A  | 43   | 86   | 165  | 301  | 484   | N/A  | 0.8  | 1.6  | 2.9  | 5.3  | 8.5   |
| Africa                                       | N/A  | 2    | 7    | 11   | 20   | 31    | N/A  | 0.2  | 0.9  | 1.4  | 2.5  | 3.8   |
| Arab States                                  | N/A  | 3    | 8    | 17   | 36   | 48    | N/A  | 0.8  | 2.4  | 5    | 10.2 | 13.3  |
| Asia & Pacific                               | N/A  | 116  | 164  | 205  | 289  | 421   | N/A  | 3.1  | 4.3  | 5.3  | 7.4  | 10.7  |
| CIS                                          | N/A  | 1    | 2    | 20   | 31   | 42    | N/A  | 0.2  | 0.8  | 7.2  | 11.2 | 14.9  |
| Europe                                       | N/A  | 89   | 148  | 200  | 254  | 336   | N/A  | 14.7 | 24.2 | 32.6 | 41.3 | 54.1  |
| The Americas                                 | N/A  | 58   | 93   | 149  | 224  | 286   | N/A  | 6.4  | 10.3 | 16.2 | 24.1 | 30.5  |

| <b>Fixed (wired) broadband subscriptions</b> |      |      |      |      |      |       |      |      |      |      |      |       |
|----------------------------------------------|------|------|------|------|------|-------|------|------|------|------|------|-------|
|                                              | 2006 | 2007 | 2008 | 2009 | 2010 | 2011* | 2006 | 2007 | 2008 | 2009 | 2010 | 2011* |
| World                                        | 284  | 351  | 414  | 472  | 527  | 591   | 4.3  | 5.3  | 6.1  | 6.9  | 7.6  | 8.5   |
| Developed                                    | 188  | 223  | 252  | 274  | 291  | 319   | 15.5 | 18.3 | 20.6 | 22.3 | 23.6 | 25.7  |
| Developing                                   | 96   | 128  | 162  | 197  | 236  | 272   | 1.8  | 2.3  | 2.9  | 3.5  | 4.2  | 4.8   |
| Africa                                       | N/A  | 1    | 1    | 1    | 1    | 1     | 0.1  | 0.1  | 0.1  | 0.1  | 0.2  | 0.2   |
| Arab States                                  | 2    | 3    | 4    | 6    | 7    | 8     | 0.5  | 0.9  | 1.3  | 1.7  | 1.9  | 2.2   |
| Asia & Pacific                               | 103  | 126  | 151  | 181  | 212  | 243   | 2.8  | 3.3  | 4    | 4.7  | 5.5  | 6.2   |
| CIS                                          | 4    | 6    | 12   | 17   | 23   | 27    | 1.3  | 2.3  | 4.5  | 6.1  | 8.3  | 9.6   |
| Europe                                       | 89   | 112  | 126  | 137  | 146  | 160   | 14.8 | 18.4 | 20.7 | 22.4 | 23.8 | 25.8  |
| The Americas                                 | 81   | 99   | 113  | 124  | 131  | 145   | 9.1  | 11   | 12.4 | 13.5 | 14.1 | 15.5  |

| <b>Households with<br/>a computer</b> | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> |
|---------------------------------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|
| World                                 | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 27.8        | 30.2        | 32.7        | 34.7        | 36.1        | 38.1         |
| Developed                             | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 57          | 62.3        | 66.1        | 69.1        | 71.1        | 74           |
| Developing                            | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 16.2        | 17.6        | 19.8        | 21.5        | 22.7        | 24.5         |
| Africa                                | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 3.3         | 3.8         | 5.6         | 6.3         | 7.1         | 7.9          |
| Arab States                           | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 17.4        | 19.7        | 22.7        | 25.8        | 28          | 30.7         |
| Asia & Pacific                        | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 21.3        | 22.6        | 24.4        | 26.2        | 26.9        | 28.5         |
| CIS                                   | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 12.2        | 24.6        | 31.8        | 36.8        | 39.2        | 43.8         |
| Europe                                | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 56.3        | 61.2        | 65.1        | 68.5        | 71.9        | 75.5         |
| The Americas                          | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 42.8        | 45.5        | 48.1        | 49.3        | 51.5        | 53.6         |

| <b>Households with<br/>Internet access<br/>at home</b> | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> | <b>2006</b> | <b>2007</b> | <b>2008</b> | <b>2009</b> | <b>2010</b> | <b>2011*</b> |
|--------------------------------------------------------|-------------|-------------|-------------|-------------|-------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|
| World                                                  | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 20.2        | 23          | 25.2        | 27.6        | 30.2        | 34.4         |
| Developed                                              | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 46.8        | 53.4        | 57.7        | 62.6        | 66.5        | 71.4         |
| Developing                                             | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 9.6         | 11.2        | 12.6        | 14.1        | 16.4        | 20.5         |
| Africa                                                 | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 1.3         | 1.7         | 2.6         | 3.4         | 4.3         | 5.7          |
| Arab States                                            | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 10.9        | 12.9        | 14.7        | 18.2        | 21.7        | 26.1         |
| Asia & Pacific                                         | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 13.5        | 15.8        | 17          | 18.5        | 20.5        | 24.9         |
| CIS                                                    | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 6.7         | 16.8        | 20.7        | 27          | 32.3        | 38.5         |
| Europe                                                 | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 46.4        | 51.7        | 57.7        | 62.4        | 67.8        | 72.2         |
| The Americas                                           | N/A         | N/A         | N/A         | N/A         | N/A         | N/A          | 34.5        | 36.8        | 38.9        | 42          | 45          | 49.7         |

Note: \* Estimate.

N/A: Not available.

<sup>a</sup> The developed/developing country classifications are based on the UN M49, see: <http://www.itu.int/ITU-D/ict/definitions/regions/index.html>.

### 2.2.1 Mobile phones for development

Until the middle to late 2000s, mobile phones inspired much less scientific research than the Internet. For example, Manuel Castells' influential and ambitious trilogy *Information Age*<sup>99</sup>, published in 1996, concentrates primarily on Internet technologies and pays little attention to the potential of mobile phones. The sentiment is echoed by Townsend:

*The advent of inexpensive mass-produced mobile communications in particular, has avoided scholarly attention, perhaps because it seems pedestrian compared to the nebulous depths of cyberspace. Yet the cellular telephone, merely the first wave of an imminent invasion of portable digital communications tools to come, will undoubtedly lead to fundamental transformations in individuals' perceptions of self and the world, and consequently the way they collectively construct that world.*<sup>100 (p85)</sup>

Ten years later, the rapid worldwide diffusion of mobile phones has made them a vibrant area for research and practice. Castells himself produced another influential work in 2007 titled *Mobile Communication and Society*<sup>98</sup>. The widespread adoption of mobile phones has occurred independently of cultural habits, values and norms. With it have come new possibilities for development on the African continent. Used by both rural and urban dwellers, wealthy and economically impoverished people, mobile phones provide a connection to individuals, information resources, economic markets and public services<sup>101</sup>.

Yet the ubiquity of mobile networks and access to handsets is necessarily only the beginning. If mobile phones and ICT are to have an impact on economic and social development, our focus needs to move from potential to the realisations of this potential. Heeks<sup>102</sup> provides a useful value-chain model (Figure 2.7), based on the logistic growth curve discussed earlier, for understanding the shifting focus from system inputs (readiness) to outputs (impact). Four domains in the value chain are proposed:

- **Readiness**

National-level preparedness is a prerequisite of information and communications technology for development (ICT4D) initiatives. A country needs to invest in infrastructure, create policy guidelines, develop skills and gain political support to secure investment. For mobile phones this would mean the construction of a national cellular network infrastructure, the training of technicians and enacting electronic communications legislation.

- **Availability**

After implementing strategies agreed upon in the previous stage, a new technology (e.g. mobile phone) or ITC service (e.g. Internet cafes) becomes available to the public. In our example, mobile phones would become easily and cheaply available for sale to the public.

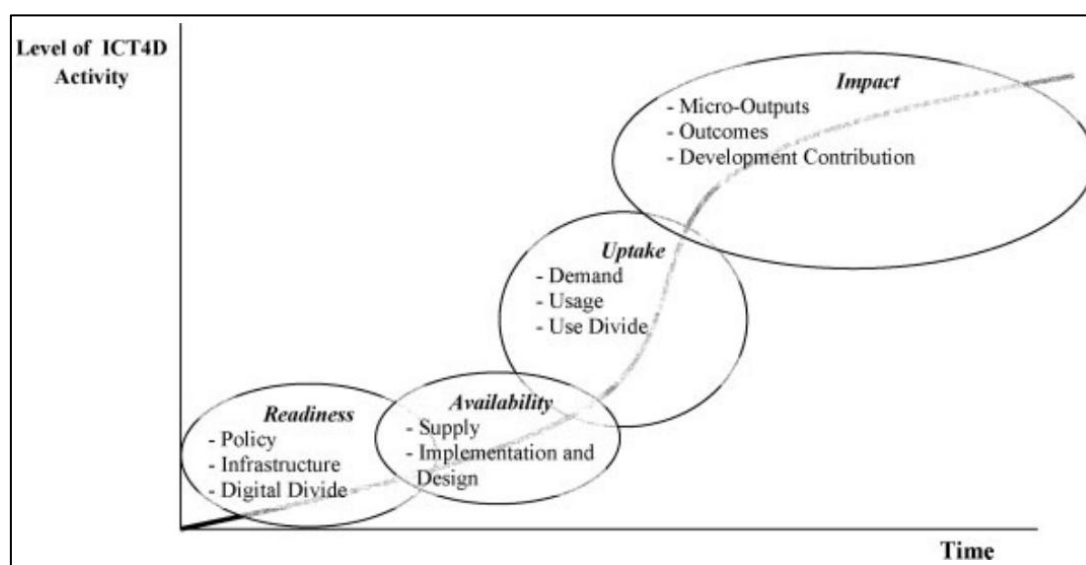
- **Uptake**

During this stage availability is converted to adoption and widespread use. The rapid uptake and wide-scale adoption are often characterised by intense interest and excitement over the newly discovered ‘silver bullet’ for development in low-income countries. Yet as Anker and Mbiti<sup>101</sup> report, the enthusiasm often wanes

when the impact becomes hard to show. A more detailed discussion of how a new technology diffuses through a country is beyond the scope of this work. A succinct literature review is provided by Geroski<sup>103</sup>. A more thorough review is provided by Rodgers<sup>104</sup>.

- **Impact**

Impact can be subdivided into the categories of outputs, outcomes, and developmental impact. New outputs, such as a change in communication patterns or the adoption of mobile banking, can lead to improved social relationships and financial outcomes. Aggregated, these improved outcomes have an impact on larger public-health development goals.



**Figure 2.7. Changing focus of information and communications technology for development (ICT4D) priorities over time**

For mobile phones to have an impact on health outcomes in low and middle-income countries, common ICT challenges faced during implementation need to be addressed fully. These challenges include: connectivity issues; the development of contextually and culturally relevant materials; enhancing human-resource skills and capacity; identifying,

involving and incorporating community feedback; and raising capital to support the ICT initiative<sup>105</sup>.

### **2.2.2 The rise of mobile phones in South Africa**

As evident in Table 2.2, fixed-line telecommunications infrastructure in Africa is scarce<sup>97,106</sup>. As a result, less than 2% of the world's fixed-line telephones are installed in Africa. In contrast, Africa has seen an explosion in mobile-phone ownership. This growth has been driven mainly by cheaper infrastructure, affordable mobile handsets, and a prepaid billing model orientated towards the poorer segments of the population. In 2008, estimates put the number of mobile-phone owners in Africa at approximately 200 million — close to a third of the 680 million people who live on the continent<sup>107</sup>. In the past five years this figure has doubled, with the International Telecommunication Union (ITU) now estimating mobile subscriptions in Africa to have risen to 433 million.

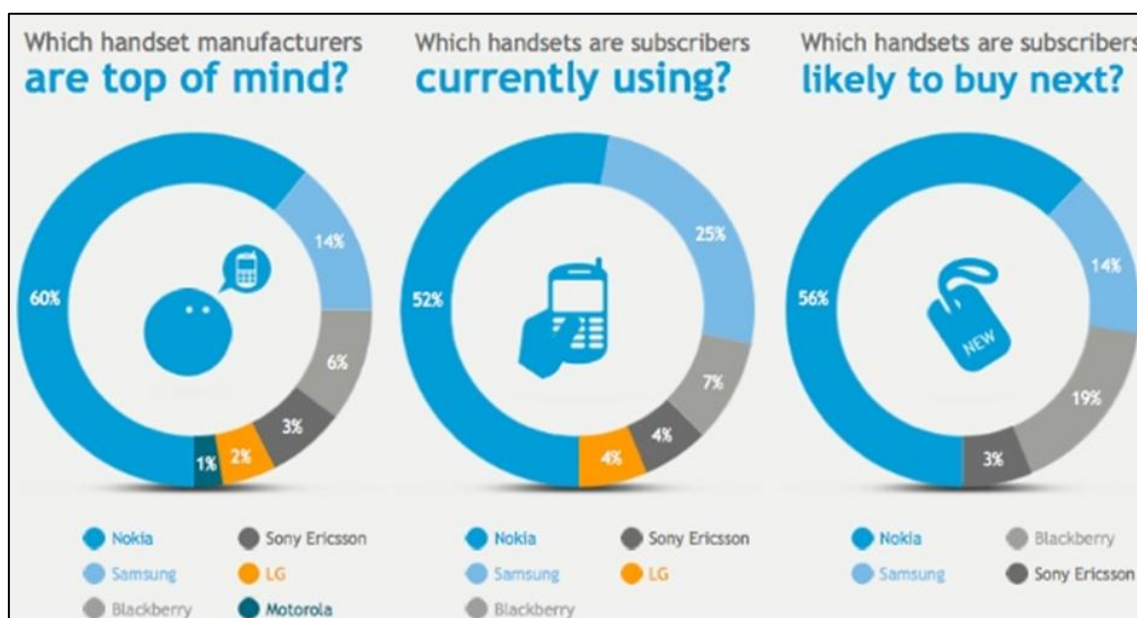
According to the All Media Products Study (AMPS) conducted by the South African Advertising Research Foundation (SAARF), 79.5% of South Africa's adult population (over the age of 15 years) own or use a mobile phone. Using the Bureau for Market Research (BMR) population estimate, based on the 2001 Census, this suggests that of South Africa's 34 million adults, 27 million own or use a mobile<sup>108</sup>. The majority of these phones are privately maintained (78.6%), pre-paid (71.1%) and connected through the two largest mobile operators, namely MTN (34.8%) and Vodacom (35.2%). Table 2.3 provides an overview.

**Table 2.3. South African mobile-phone ownership trends (%) from the South African Advertising Research Foundation's All Media Products Study (2011)**<sup>108</sup>

|                                                                                                                                                       | 2006   | 2007   | 2008   | 2009   | 2010   | 2011   |
|-------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------|--------|--------|--------|--------|
| AMPS Universe<br>(Adult Population in '000)                                                                                                           | 30,903 | 31,109 | 31,305 | 31,524 | 34,020 | 34,020 |
| Cell Phones                                                                                                                                           |        |        |        |        |        |        |
| Own/Rent/Use                                                                                                                                          | 49.4   | 56.3   | 67.9   | 71.3   | 76.3   | 79.5   |
| Company/Private Maintained:                                                                                                                           |        |        |        |        |        |        |
| Privately Maintained                                                                                                                                  | 48.3   | 55.1   | 66.7   | 70     | 75.1   | 78.6   |
| Company Maintained                                                                                                                                    | 1.1    | 1.3    | 1.1    | 1.3    | 1.2    | 1      |
| Contract/Pre-paid:                                                                                                                                    |        |        |        |        |        |        |
| Contract                                                                                                                                              | 6.9    | 7.7    | 8.6    | 8.7    | 8.4    | 8.4    |
| Pre-paid                                                                                                                                              | 42.5   | 48.6   | 59.3   | 62.6   | 67.9   | 71.1   |
| Service Provider:                                                                                                                                     |        |        |        |        |        |        |
| Cell C                                                                                                                                                | 4      | 4.5    | 5.8    | 8.3    | 8.5    | 9.2    |
| MTN                                                                                                                                                   | 20.1   | 22.9   | 29.4   | 30.5   | 33.9   | 34.8   |
| Virgin Mobile                                                                                                                                         | -      | -      | 0.3    | 0.2    | 0.2    | 0.2    |
| Vodacom                                                                                                                                               | 25.4   | 28.8   | 32.3   | 32.4   | 33.6   | 35.2   |
| South African Advertising Research Foundation. All Media Products Study.<br>Available at: <a href="http://www.saarf.co.za">http://www.saarf.co.za</a> |        |        |        |        |        |        |

In a 2011 market research study, Nielsen<sup>109</sup> reported that the majority (52%) of South Africans currently own a mobile phone manufactured by Nokia (Figure 2.8). The most popular of the Nokia handsets is the Nokia 1100 series. Samsung holds the second largest market share (25%). Nokia also holds the largest mind share (60%) and is the most likely phone to be purchased next by respondents (56%). Ninety percent of respondents report owning one handset and 85% one SIM card. Roughly half of both males (49%) and females (48%) prefer making calls to sending an SMS (41% and 44% respectively). Using the South African Advertising Research Foundation's Socioeconomic Living Standards Measure (LSM), which divides the population into 10 LSM groups depending on asset ownership, Internet access on mobile phones was found to vary. One out of three respondents (30%) with higher living standards (LSM 8-10) was found to have accessed social media over the Internet in the previous month. For LSM groups 1-5, this figure was

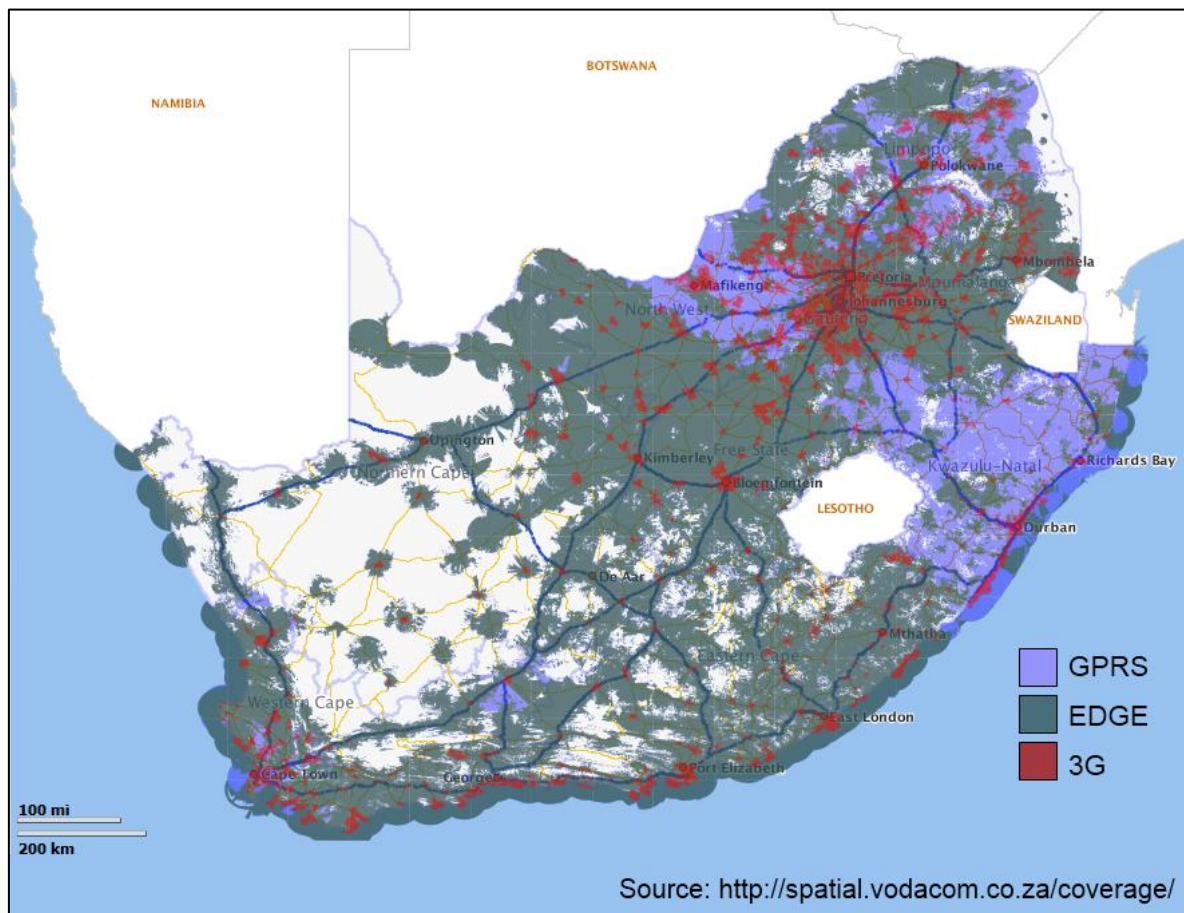
15%. One out of four respondents (25%) in LSM groups 8-10 accessed the Internet directly on their phone in the previous month, but this figure was less than one out of ten (6%) for LSM groups 1-5.



**Figure 2.8. Manufacturer preferences among 2 000 South African respondents, with weightings based on age, gender and all LSM levels in 2011**<sup>109</sup>

Over 900 000 km<sup>2</sup> of the country is covered by the GSM network of Vodacom, the largest telecommunications company in the country. Ongoing infrastructure investment in the region has seen a 2008 prediction of mobile coverage reaching 90% in 2011 become a reality<sup>110</sup>. The determinants of mobile cellular network coverage are explored by Buys et al.<sup>111</sup> Using high-resolution, spatially disaggregated datasets of Sub-Saharan Africa, they correlated cell tower positions to geographic features of the surrounding terrain. As expected, the analysis found a strong positive correlation between the presence of a cell tower and population density, per capita income and the degree of operator competition. Reduced coverage was noted in areas where installation and

maintenance were hampered by geographic factors such as elevation, steepness of slopes and distance from main road or major city.



**Figure 2.9. Vodacom cellular network coverage map of South Africa (2011)**

Figure 2.9 shows that although well covered by GPRS (2.75G) and EDGE (Enhanced data rates for GSM evolution), Vodacom's coverage of South Africa with the faster 3G network is much lower. Nevertheless, given the expanded network footprint, high saturation rate and relatively low cost of mobile technologies, widespread enthusiasm has been generated about the potential of information and communication technologies (ICTs) to affect development positively. This enthusiasm, unsupported by evidence, prompted people to promote information and communication technologies as a way to bring about

revolutionary change in the development of low and middle-income countries<sup>112</sup>.

Education, agriculture, finance and health are all domains in which mobile phones are believed to have the potential to affect development positively. In the final section of this literature review, I will consider the impact mobile phones are having on health research.

### **2.3 Application of Mobile Phones to Data Collection in Health Research**

*To date, the vast majority of research on the human condition has relied on single-shot, self-report data: a yearly census, public polls, focus groups, and the like.*

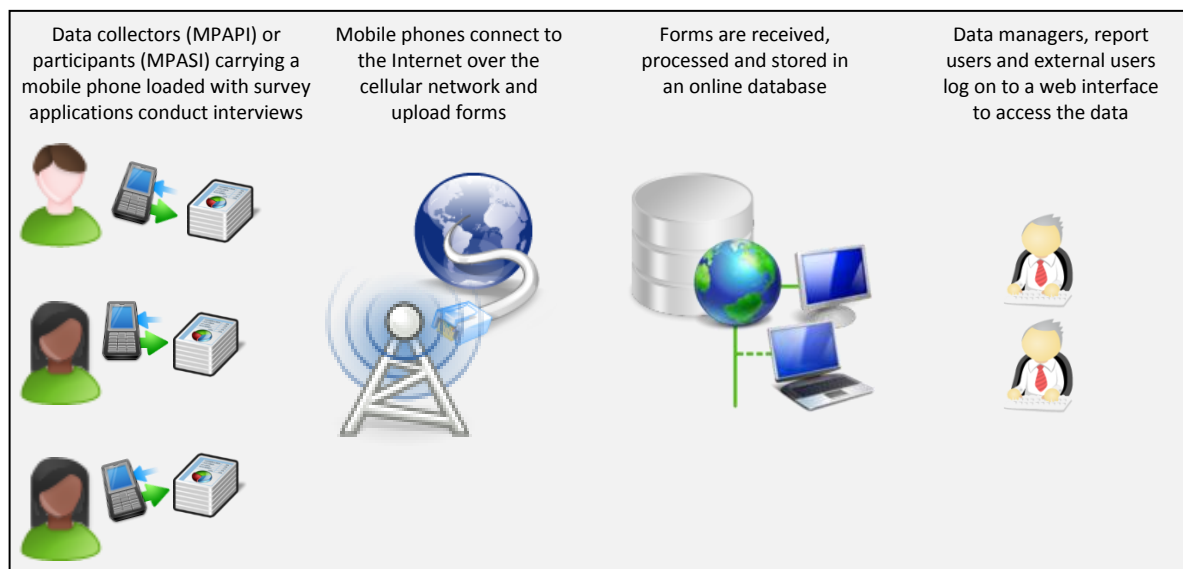
- Sandy Petland<sup>36</sup>

With aging populations, rising health care costs and a growing level of e-literacy, many high-income nations are exploring how pervasive computing technologies can improve the efficiency and effectiveness of health care provision<sup>113–115</sup>. In these countries, information and communication technologies are one of the avenues being explored as a way to transform health care delivery, for example, through outpatient monitoring of chronic disease<sup>116,117</sup>. Information about the prevention and treatment of many diseases is widely available for those with access to the World Wide Web<sup>118</sup>. Yet, with a complex transitional health picture with highly prevalent infectious disease as well as fast-emerging non-communicable diseases and a “digital divide” in available technology, people living in middle and low-income countries have not benefited to the same degree<sup>119</sup>. However, as mentioned, the grouping of existing technologies in new arrangements can lead to incremental innovation<sup>18</sup> and new solutions to old problems. The combining of pervasive mobile computing and the survey approach to data collection has given rise to mobile-phone-assisted data collection. This new approach may provide

one solution to the problem of information collection, knowledge dissemination and data quality of health services in low-resource settings<sup>120</sup>.

### 2.3.1 Technology behind mobile-phone health interventions

Using mobile phones, survey software, a cellular network and an online database, it is possible to conduct survey research. Figure 2.10 illustrates the components that make up a mobile-phone-assisted interview.



**Figure 2.10. A schematic representation of how data is collected, uploaded/processed and viewed in a mobile-phone-assisted interview**

- **Mobile-phone operator**

The mobile phone can either be operated by a trained data collector who asks a respondent the questions presented on the mobile phone (mobile-phone-assisted personal interviewing – MPAPI), or surveys can be sent directly to the mobile phone of the survey respondent (mobile-phone-assisted self-interviewing – MPASI).

- **Mobile phone**

The data-collection options available on a mobile phone are mediated by both the hardware and the operating system of the phone. As seen in Figure 2.8, the majority of South Africans own handsets manufactured by Nokia. These handsets have traditionally run the Symbian operating system and are able to install applications that conform to the Java 2 Micro-Edition (J2ME) platform. In addition to Symbian, there are three other popular mobile operating systems: iOS, Android and Windows Phone. The data-collection software to be installed on a handset needs to be designed to take into account the differences between these operating systems (OS) as well as differences between different versions of the OS<sup>121,122</sup>.

- **Survey presentation**

Surveys can be delivered to a mobile phone in a number of ways. The most popular methods to present surveys on mobile phones are through the use of the fixed-format short message service (SMS, also known as a text message), USSD (unstructured supplementary service data), dedicated survey software and a web-based form. Table 2.4 describes each in more detail and gives an example of products available in each category. Although traditional voice and interactive voice response (IVR) surveys could also be conducted with respondents using their mobile phones, they are not included in Table 2.4 since they are more popular as a fixed-line telephone method.

**Table 2.4. An overview of mobile-phone technologies that enable the collection of survey data on the handset**

|                       | Description                                                                                                                                                                                                                                                 | Supported handsets                                                                                | Example tools                                    | Key data-collection literature |
|-----------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|--------------------------------------------------|--------------------------------|
| SMS                   | The short message service (SMS) allows the sending and receiving of short alphanumeric messages to and from mobile telephones <sup>123</sup> . Data collection can be achieved through the use of simple single-item questions.                             | All                                                                                               | RapidSMS, Frontline                              | Patnaik et al. <sup>124</sup>  |
| USSD                  | Allows for the transmission of information via the cellular network. An interactive service can be set up between mobile phones and an application on the network <sup>125</sup> . The use of USSD menus allows the asking of simple single-item questions. | All                                                                                               | No user-friendly tool available                  | Medhi et al. <sup>126</sup>    |
| On-device application | Survey applications allow full-featured, multi-item surveys to be administered on a mobile phone. Was predated by personal digital assistant (PDA) based survey software.                                                                                   | Any device that is running the version of operating system that the application was designed for. | EpiSurveyor, JavaRosa, Mobile Researcher, REDCap | None to date                   |

- **Uploading surveys**

There are two common methods of uploading surveys from the handset into a data store. If a cellular network signal is available, data are uploaded using this data connection. In the absence of a signal, data are stored on the handset until such time that a signal does become available. If no cellular coverage is available in the area at all, data can be uploaded from the phone by connecting it directly to a computer with an Internet connection.

- **Data storage**

Data uploaded from the handset can be stored in any number of ways.

Commonly, this involves an online database set up to receive data in the format of the survey.

- **Web interface**

The final component of a typical online system will be the presence of a web interface. The functionality of such an interface can vary from minimal, e.g. providing an option to download the data to a local file, to a full-featured system with a well-designed user interface and access to reports, graphs and interviewer statistics.

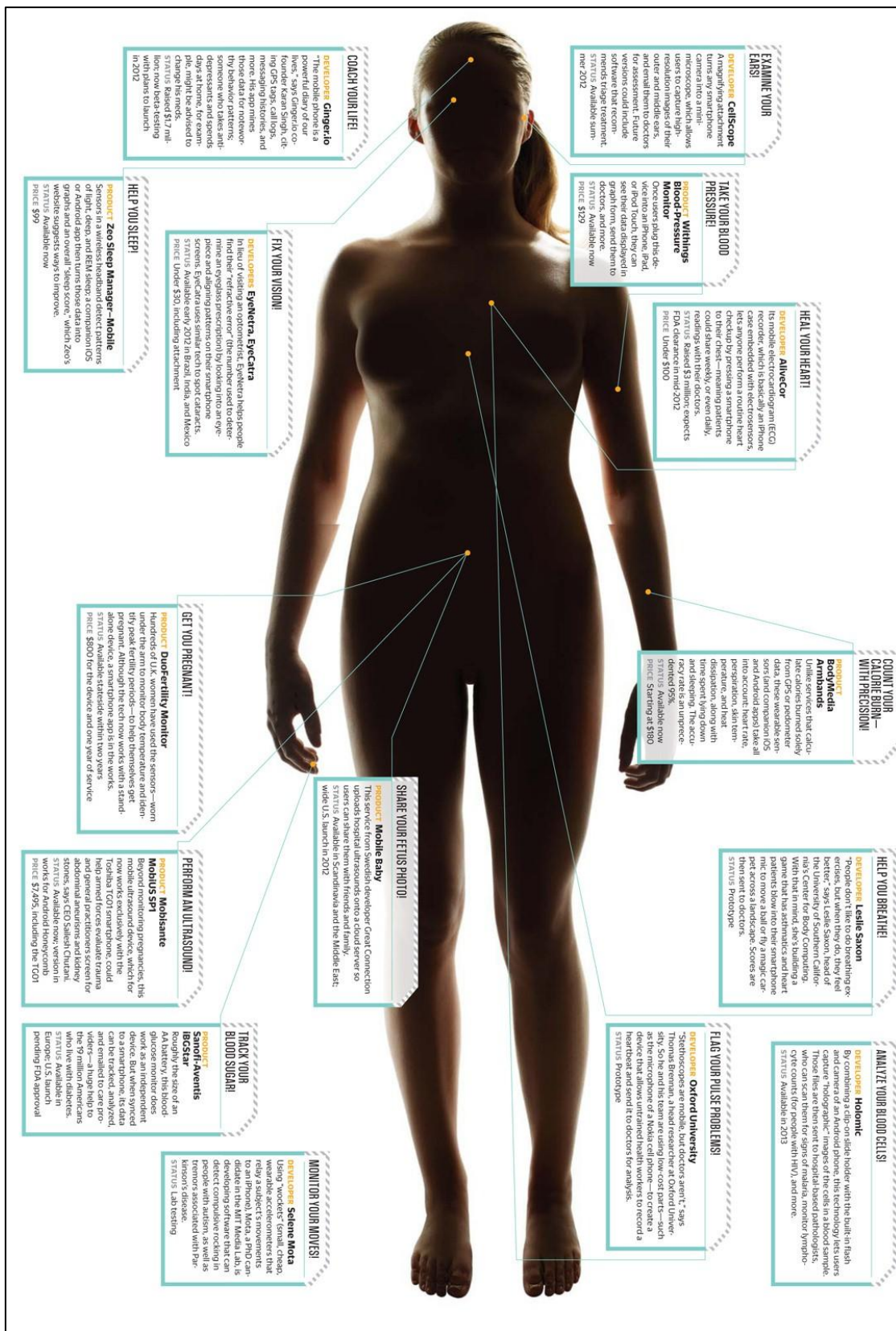
In 2006, Kaplan<sup>127</sup> documented the dearth of literature on using mobile telephones for health care in developing countries. The past five years has seen the landscape change dramatically with a proliferation of studies seeking to harness the capabilities of mobile phones to monitor health and treatment-related behaviour<sup>122</sup>. Studies generally fall into one of three domains: tracking health behaviours (sometimes with an associated intervention), infectious-disease surveillance, and treatment monitoring. The various uses

of mobile phones in health has come to be known as mobile Health<sup>128–132</sup>. Mobile health, or mHealth, has been defined by the Global Observatory for eHealth (GOe) as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices”<sup>133</sup> (p6).

### **2.3.2 Application of mobile phones in health**

Numerous health-related commercial applications and health research studies exist that leverage the mobility, ubiquity and availability of on-board sensors (e.g. cameras) in mobile phones. Health behaviours such as weight control<sup>134</sup>, smoking cessation<sup>135</sup> and blood glucose monitoring<sup>136,137</sup> have all been tackled with mobile-phone data-collection and tracking tools. The support of mental health through mHealth is also growing<sup>138</sup>.

Klasnja<sup>122</sup> categorises applications into those that provide users with a journaling application that tracks data and provides feedback on progress; those that use mobile phones to receive data from other devices (such as a pedometer) and upload the data for telemonitoring; and mobile-phone games designed to teach health-related behaviours. Figure 2.11, reproduced from an article on mHealth in the *Fast Company* magazine<sup>139</sup>, depicts the diverse range of commercially available mHealth monitoring tools currently available. The availability and use of mobile-phone tools to promote health is of particular interest to proponents of a more participatory model of health care and health reporting<sup>140</sup>.



**Figure 2.11. A sample of mobile-phone applications available in 2012 that are designed to provide feedback on various health measures to the user** <sup>139</sup>

### 2.3.3 Infectious-disease surveillance

Growing global connectedness makes systems that monitor the early outbreak of new infectious diseases critical. In 2002, Frenk and Gómez-Dantés<sup>141</sup> reported that there were three million international travellers per day, triple the number in 1980. This increase resulted in a corresponding increase in *microbial traffic* (p325). To combat the threat, the malERA Consultative Group on Monitoring, Evaluation, and Surveillance<sup>142</sup> suggested the rigorous evaluation of new forms of mobile and web-enabled surveillance technologies. The advantages of electronic capturing applied to public-health surveillance data may bolster its usefulness through early detection and registration, rapid reporting, analyses and action<sup>143,144</sup>. As an example, a sentinel surveillance system based on mobile computing technologies, such as mobile phones, was shown to increase the detection and response speeds to outbreaks of malaria in Madagascar<sup>145</sup>.

The authors of the Madagascar trial conclude that for less than US\$1 a month for data-transmission costs, their findings *should encourage other African countries to develop their own syndromic surveillance systems*<sup>145 (p1)</sup>. A South African study reported similar success in a mobile-phone surveillance system set up to allow the rapid reporting of zoonotic diseases in rural South Africa<sup>146</sup>. The case study reports no specifically measureable outcomes but does describe the successful implementation of an SMS-based system that allowed users to report unusual symptoms to a state veterinarian. These mobile-phone-based systems allow for the capturing of health data from remote rural settings where resources, a rugged environment and limited infrastructure make traditional techniques less feasible, yet at the same time provide all the benefits of a fully fledged electronic surveillance system<sup>147</sup>.

### **2.3.3.1 Community health workers**

A subset of surveillance systems are those designed specifically to support and monitor community health workers. Much of the early innovation in mHealth data collection arose from the desire to improve monitoring and speed up data collection by community-based health workers providing disease surveillance and treatment support. The Cell-PREVEN trial in Peru was one of the first to collect health data using mobile phones<sup>148</sup>. By dialling into an interactive voice response (IVR) system tailored to provide audio computer-assisted personal interviewing, the researchers were able to collect real-time adverse-events data from health workers administering metronidazole to female sex workers. Equipping community health workers (CHW) with mobile phones for data collection and supervision has been explored in a number of studies, and has shown positive results<sup>149–151</sup>. Reviewing the role of mobile phones in community health worker systems, De Rezni et al.<sup>152</sup> report on how each of the six functions of a CHW health system (data collection, training and access to materials, inter-CHW communication, decision support, staff supervision, and promotion of health behaviours) are impacted by mobile phones. In each instance, mobile phones were found to offer an additional contribution to a paper-based system. The challenges the authors observed in using mobile phones for low and middle-income CHW programmes included data interoperability, managing workflow complexities, loss and theft of devices, ensuring batteries remain charged and network coverage constraints. In addition to the phone-maintenance challenges, a peer health worker programme in Uganda reported key challenges that included privacy concerns and variable levels of patient access to mobile phones, limiting the usefulness of communication between peer health workers and patients<sup>153</sup>.

### 2.3.4 Treatment monitoring

Adherence and treatment monitoring is the third active area of mHealth research.

Evidence is still mixed as to its benefits. Lester et al.<sup>154</sup> report on a randomised control trial (RCT) in Kenya that was designed to test the impact of weekly SMS reminders on self-report antiretroviral therapy (ART) adherence and viral load. Adherence was reported to be significantly higher in the intervention (SMS) (62%) than in the control group (50%; relative risk [RR] for non-adherence 0.81, 95% CI 0.69–0.94;  $p=0.006$ ). Viral-load suppression was also significantly higher in the intervention group (57% versus 48%; RR for virological failure 0.84, 95% CI 0.71–0.99;  $p=0.04$ ). Haberer et al.<sup>155</sup>, working in Uganda, report on the uses of text messages (SMS) and an interactive voice response system (IVR) to collect adherence data from the caregivers of 19 HIV-infected children. A number of technical challenges resulted in low numbers (0-33%) of caregivers submitting adherence data. A programme using SMS to send treatment reminders to tuberculosis (TB) patients in Cape Town, South Africa, reported similar challenges<sup>156</sup>. Examining the cure and completion rates of 221 patients revealed no difference between the control and intervention groups. The authors cite major challenges in project implementation and note that patients were often not using the service as directed. These shortcomings reduce the value and validity of the findings and make interpretation problematic. In a second randomised control trial using an intention-to-treat analysis, Pop-Eleches et al.<sup>157</sup> divided 431 people into a control group and four intervention groups. Intervention groups differed in the length (short versus long) and frequency (daily versus weekly) of SMS treatment reminders. Weekly messages (whether short or long) were found after 48 weeks to increase adherence significantly when compared to the control group ( $p = 0.03$ ). Still, 90% adherence was only achieved by 53% of the sample receiving the weekly

reminder (40% for the control). Despite challenges, well-run programmes of SMS-based treatment reminders appear to be an important tool with which to improve medication adherence in resource-limited settings.

### **2.3.5 Comparison with other methods**

Despite the swell first in popular and more recently in scientific support for the use of mobile phones as data-collection tools in health, there are still very few studies that actively seek to compare the quality of data collected on mobile phones with traditional health research survey instruments. In its place, a reasonable body of literature examines the properties of personal digital assistants (PDAs) and other survey modes. PDAs are mobile devices that function as personal information managers. What distinguishes them from modern mobile phones is that they do not generally have the ability to make or receive telephone calls. Two excellent summaries exist that compare PDA and paper-and-pen data-collection modes. The first, a literature review by Lane et al.,<sup>158</sup> uncovered 201 potentially relevant studies. Nine of these studies met the criteria of a randomised control trial, comparing PDA and paper-and-pen data collection on the outcomes of data quality, user preference, adherence and timeliness. The second, more recent, review by Blaya et al.<sup>159</sup> summarises seventeen studies that used PDAs for data collection in health research studies. Both reviews concluded that PDAs are an effective alternative to paper and pen, often producing data faster and meeting with favourable responses from most participants.

The wealth of literature on PDAs as data-collection tools does not yet exist for mobile phones. Patnaik et al.<sup>124</sup> provide the most comprehensive study, conducted in 2009, evaluating the errors associated with mobile-phone data-collection through the interface

of SMS, voice and electronic form. They found that capturing data by voice call produced the lowest error rate (0.45%), followed by electronic form (4.2%) and SMS (4.5%). The usefulness of these results is, however, limited by sample size. Only 13 participants (six community health workers and seven paramedical staff) were enrolled in the study. Further limiting the generalisability of results is the fact that each participant only captured one 22-item form using voice, SMS or electronic form. The authors were therefore unable to compare within-subject error variability to between-subject error variability to establish if some participants were simply more error-prone than others. In the context of evaluating back-pain reports from subjects referred through the Spine Centre of Southern Denmark, Johansen and Wedderkopp<sup>160</sup> contacted 31 of the 260 enrolled patients by telephone at the end of a 53-week period. Comparing levels of back pain reported by SMS to those given telephonically, there was high agreement in reported levels of pain for one-week and one-month recall. Recalled levels of pain varied significantly between SMS and telephone for one-year recall. Eagle et al.<sup>161</sup> found a similar primacy effect when comparing self-report data on social interactions to those obtained automatically using mobile-phone records. Following 96 people with custom software installed on their mobile phone, researchers were able to log the presence of another participant's phone when within a five-metre proximity radius. When people were asked if they had been in the presence of another participant, self-report recall of interactions that had taken place more than a week earlier were poorly correlated with data from the mobile-phone software log ( $R^2 = 0.1$ ). The salience of the relationship was also an important factor that mediated correlation strength. For people working together but not considered friends, the correlation between self-report and mobile-phone proximity data was  $r = 0.155$  ( $p < 0.001$ ). For friends, the correlation was  $0.412$  ( $p < 0.001$ ).

Apart from these studies, not much else is known about how the method of collecting data on mobile phones influences data quality.

Collecting and managing health research data is a challenging undertaking. Not only is considerable planning required to ensure that paper forms are always available, and are routinely collected and accurately captured, but repackaging this information to support operational initiatives is often slow, expensive and vulnerable. At present, mobile phones offer a real opportunity to address some of these challenges in the collection of various types of health information. It is clear from this review that although mHealth applications continue to proliferate, little research has, to date, been conducted to evaluate the implementation of these new tools in health research.

## **2.4 Summary of the Literature Review**

This literature review introduced two fields of research relevant to this thesis: pervasive computing and survey error. Each was introduced with a short history, followed by an appropriate conceptual framework, and concluded with practical applications of the theory. Common threads from the two fields were then woven together to introduce the notion of mobile-phone-based data collection. Finally, a conceptual framework for the evaluation of health research tools was elaborated upon.

When the research for this thesis began, there was little literature specifically evaluating the role of mobile phones as data-collection tools in low and middle-income countries. At the time, most literature still focused on personal digital assistants (PDAs). The field of research has, however, gained huge momentum in the last year or two, quickly filling

gaps in our knowledge. Despite this, the literature review revealed that a number of research gaps persist. Little research has been conducted to determine whether mobile phones are a viable means of collecting ongoing self-report survey data from participants living in low-resource settings. Furthermore, the acceptability and feasibility of mobile phones as a tool for the collection of sensitive data from vulnerable population groups is not well established and very little research exists that explicitly compares mobile phones to traditional pen-and-paper health surveys. Lastly, as noted in a recent systematic review of eHealth<sup>162</sup> literature, more evidence is needed to support the clinical management claims of mHealth projects.

Consequently, this thesis addresses some of the research gaps identified. The overall aim of the research was to implement and evaluate the use of mobile phones as research instruments with which to collect and analyse information for research in low-resource Sub-Saharan African settings. This primary aim was achieved through the completion of three studies, each designed to address one aspect of the use of mobile phones:

1. How does the use of mobile phones for data collection have an impact on the costs, quality and time-to-availability of data?
2. What perceptions do vulnerable communities and field workers in rural South Africa hold towards mobile phones as a data-collection tool?
3. Can mobile phones be used by adolescents to complete self-report surveys unaided?

In addition to the primary aim, the thesis explores the policy and research implications of the study findings.

Together, these general objectives, addressed in chapters 3, 4, 5, 6 and 7, are designed to address different facets of the primary aim. Two main hypotheses are tested during the course of this thesis. They are:

$H_0$ : Given respondent concerns, it will not be feasible to implement mobile-phone-assisted personal interviewing to collect health research data in resource-limited settings.

$H_0$ : Mobile phones provide no data-quality improvements to the survey data-collection process when compared with paper and pen.

---

## **Part 2: Empirical Papers**

---

## Chapter 3

---

### **Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa**

Published in *The Journal of Adolescent Health*, 46(3), 302-4.

### **Chapter 3: Using mobile phones for adolescent research in low and middle-income countries: preliminary findings from the Birth to Twenty cohort, South Africa.**

---

From fixed-line telephones as an interview medium in the 1950s to more recent computer-aided surveys, technological innovation brings new opportunities. Mobile phones offer potential benefits for research data collection, particularly surveys aimed at adolescents. The International Telecommunication Union report that over one billion new subscriptions were added during 2008, bringing the worldwide total to four billion, approximately one third of these in Brazil, Russia, India and China<sup>163</sup>.

SexINFO<sup>164</sup> was the first mobile service in the United States offering sexual health information directly to adolescents. In response to an opt-in text message, a menu of questions is returned with instructions to text, for example, “D2 if u think ur pregnant”. A similar concept allows youth in South Africa to locate the nearest HIV testing centre by texting “HIV” and their town name to a five-digit number<sup>165</sup>. A Cape Town-based study piloted the use of mobile phones as a means for sending daily treatment reminders to tuberculosis patients on a DOTS (Directly Observed Treatment Short-course) programme. This decreased DOTS employees’ regular workload and freed them up to focus on non-compliant patients. Of the 300 patients on the programme, the World Health Organization reported only five treatment failures and now recommends this model as one of best practice<sup>166</sup>. These examples illustrate the potential of mobile devices. We investigated the viability of delivering self-administered surveys to adolescents on their mobile phones.

### **3.1 Method**

BT20 is a birth cohort study based in Soweto-Johannesburg. In 1990, 3 273 newborns and their mothers were enrolled to track the children's health and development over the first twenty years of life<sup>167</sup>. Data has been collected every six months through face-to-face and computer-assisted self-interviews. Plans are underway to introduce a mobile-phone data-collection platform. The rationale includes increasing the ecological validity and improving the reliability and accuracy of sensitive self-report data. A telephonic survey to determine the acceptability and feasibility of using mobile phones was conducted with BT20 participants, 2 023 of whom responded (88% of current participants; 52% female; 79% black). Each interview was approximately five minutes long with the major reason for non-participation being unavailability at the time of the call. Adolescents were mostly living in households that have electricity (79%), indoor running water (69%), and indoor sanitation (54%). Nonetheless, these households are poor relative to households in high-income countries. Questions were designed to assess the current perceptions and attitudes of BT20 youth to mobile-phone ownership and use, in preparation for research use. The survey, an integral part of BT20, received human-subject research approval from the University of the Witwatersrand. Pearson Chi-square tests were used to compare the proportional difference between gender, race, phone choice and feature usage. Frequency tables were prepared to explore phone contract type, mobile social media usage and opinions about mobile data collection.

### **3.2 Results**

Among respondents, 69% reported currently owning a mobile phone. In contrast, only 43% had access to a fixed-line telephone at home. About 66% of the 31% who do not own their own mobile phone reported feeling socially isolated as a result. No gender

difference was found in how adolescents choose their phone; 60% based their selection on functionality, followed by 38% who made the decision based on styling. Only 8% chose a phone primarily on price. There were significant differences ( $\chi^2 = 62.26$ ,  $df = 3$ ,  $p < 0.01$ ) in phone ownership by ethnic group. Significantly more white adolescents (95%) owned a mobile phone, compared with black (67%) and coloured (63%) adolescents. The majority of participants (86.3%) reported being on a prepaid plan. Of those on a prepaid plan, 62.1% pay for their own airtime, with the other third receiving airtime from their parents.

Table 3.1 provides a summary of the most highly valued mobile phone features. There are few clear gender differences, apart from females placing a high value on a mobile phone being able to play music. The most frequently used mobile phone features were listening to music, texting and taking photographs.

**Table 3.1. Importance of mobile-phone feature (including voice calls, SMS, MMS, camera and Internet access) by gender in the Birth to Twenty Cohort (2009)**

|                                                                                                            | <b>Males</b>                | <b>Females</b>  |
|------------------------------------------------------------------------------------------------------------|-----------------------------|-----------------|
| Making voice calls                                                                                         | 284<br>(64.3%)              | 280<br>(60.1%)  |
| Sending SMS's                                                                                              | 294<br>(66.5%)              | 313<br>(67.0%)  |
| Sending MMS's                                                                                              | 228<br>(51.7%)              | 252<br>(54.1%)  |
| Listening to music                                                                                         | 299<br>(67.8%) <sup>#</sup> | 350<br>(74.9%)* |
| Having a camera                                                                                            | 274<br>(62.3%)              | 303<br>(65.0%)  |
| Is able to connect to the Internet                                                                         | 224<br>(50.9%)              | 239<br>(51.3%)  |
| Being able to do mobile banking                                                                            | 50<br>(11.4%)               | 53<br>(11.4%)   |
| * - Significantly more than expected ( $p < 0.05$ )<br># - Significantly less than expected ( $p < 0.05$ ) |                             |                 |

When asked how they would prefer to answer sensitive questions, 73.1% chose face-to-face interviews over mobile-phone surveys. However, Table 3.2 highlights an interesting pattern between mode of interview preference, and knowledge and use of telecommunication technologies. There was a significant likelihood that those who had used Mxit (a popular mobile instant-messaging application), or who knew about or used Skype (a voice-over-IP application), would prefer a mobile-phone questionnaire over the traditional face-to-face interview.

**Table 3.2. Preference for mobile-phone questionnaire or face-to-face interview and technological affinity in the Birth to Twenty Cohort (2009)**

|                                                 | Response | Mobile-phone questionnaire | Face-to-Face interview   |
|-------------------------------------------------|----------|----------------------------|--------------------------|
| Do you know what Mxit is?                       | No       | 43 (13.0%)                 | 119 (13.3%)              |
|                                                 | Yes      | 287 (87.0%)                | 777 (86.7%)              |
| Have you ever used Mxit?                        | No       | 110 (36.9%) <sup>*</sup>   | 358 (43.9%) <sup>#</sup> |
|                                                 | Yes      | 188 (63.1%) <sup>#</sup>   | 457 (56.1%) <sup>*</sup> |
| Do you know what Skype is?                      | No       | 275 (83.8%) <sup>*</sup>   | 801 (89.9%) <sup>#</sup> |
|                                                 | Yes      | 53 (16.2%) <sup>#</sup>    | 90 (10.1%) <sup>*</sup>  |
| Have you ever used Skype?                       | No       | 110 (84.0%) <sup>*</sup>   | 300 (91.7%) <sup>#</sup> |
|                                                 | Yes      | 21 (16.0%) <sup>#</sup>    | 27 (8.3%) <sup>*</sup>   |
| * - Significantly more than expected (p < 0.05) |          |                            |                          |
| # - Significantly more than expected (p < 0.05) |          |                            |                          |

### 3.3 Discussion

Three main concerns emerged about using mobile phones to collect data. These were: a) not having airtime with which to respond, b) lacking confidence about understanding the questions being asked and c) issues of confidentiality about who might see their answers.

Despite the advantages of speed, event proximity and logistical ease, it is necessary to address these issues. Firstly, airtime should be considered whenever surveys are

administered. Secondly, questions must be worded clearly with detailed instructions prefacing each section. Lastly, a methodology that ensures participant privacy and allows for consent to be obtained via the handset is imperative. Although not relevant to this cohort, the practice of handset sharing could have a negative impact on privacy. It is not possible to password-protect an SMS. Therefore, if information is sent by the researcher using this protocol, it may compromise participants' privacy. This challenge may be largely resolved by using mobile survey software that can be password-protected and by encrypting data.

Knowing that adolescents value multimedia content on their mobile phones raises the possibility of using popular songs, ringtones and screensaver downloads as incentives for completing a survey.

Although low literacy rates and language barriers can be a challenge for self-completed questionnaires, providing the survey in colloquial languages is technologically simple on mobile phones. Each participant could receive the survey in his or her language of choice.

Mobile phones provide an important platform from which to engage adolescents. The devices have penetrated even the most rural villages, are highly valued by adolescents<sup>168</sup>, and offer instantaneous bi-directional communication. How to best scale these projects, manage informed consent and maintain privacy are all issues that need urgent attention as technological advancements continue to push the bounds of what is currently possible.

## Chapter 4

---

### **On the move: Mobile phones as a platform for the delivery of questionnaires to young adults**

In review: *The International Journal of Social Research Methodology*

## **Chapter 4: On the move: Mobile phones as a platform for the delivery of questionnaires to young adults.**

---

Longitudinal cohort studies requiring periodic contact with participants to update exposures and/or record outcome events present evolving challenges, in particular in relation to retaining contact with participants, especially if they are highly mobile<sup>169</sup>. Birth to Twenty (BT20), a birth-cohort study based in Soweto, Johannesburg, has tracked the health and development of 2 300 mother-child pairs over the past 21 years, from pregnancy to age 19.5 years<sup>167</sup>. This is 70 % of the original cohort of 3 273. Data have been collected from the cohort 23 times, including twice a year since age 13 years. Three specific challenges posed by research on adolescent risk behaviour, and specifically faced by BT20 are: 1) retaining contact with an increasingly mobile population, the majority of whom have left formal education and are starting to leave home to seek work; 2) collecting reliable information pertaining to sensitive behaviours (e.g. drug and alcohol use); and 3) accurately measuring infrequently or irregularly occurring behaviours (e.g. sexual behaviour)<sup>170</sup>.

### **4.1 Introduction**

#### **4.1.1 Retaining contact with an increasingly mobile population**

No challenge is more pressing than maintaining a cohort and minimising selective attrition<sup>171</sup>. Common strategies used to increase retention include taking care at enrolment to record a variety of contact information and to assess participants' long-term willingness to participate, bonding with participants through the creation of a study identity (e.g. a logo), ensuring frequent contact, hiring well-trained and enthusiastic study

staff, and providing small incentives or tokens of appreciation for participation<sup>172</sup>. For BT20, the past twenty years has shown that the key driver of attrition is mobility and migration out of the study area<sup>170</sup>, often as the result of job seeking. By the end of Year 15, migration accounted for 27% of loss to follow-up cases. During any given data-collection wave, between 4% and 49% of families report a change of address<sup>169</sup>.

#### **4.1.2 Collecting reliable information pertaining to sensitive behaviours**

Adolescent smoking, alcohol, violence, conflict with the law and risky sexual practices are important and active areas of research<sup>173,174</sup>. However, collecting accurate information about these behaviours is not easy. Parental reports are not always informed or accurate<sup>175–177</sup>, and face-to-face interviews produce more socially desirable responses from young adults than do self-report surveys<sup>75,178,179</sup>. To increase the reliability of sensitive data collected, the BT20 cohort has evolved its data-collection methods from face-to-face questionnaires to self-completion with secret ballot submission and, most recently, to audio computer-assisted self-interviewing (ACASI).

#### **4.1.3 Accurately measuring infrequently or irregularly occurring behaviours**

Many risky behaviours are erratic, occurring only infrequently and sporadically<sup>180,181</sup>. Relying on questionnaires that require young adults to recall the number and detail of these risky events over long periods between follow-ups can result in both under- and over-reporting<sup>182,183</sup>. As a result, some studies have evaluated the use of diaries as a way to record ecological momentary assessment data<sup>184</sup>. Given these challenges and knowing the importance of a good communication infrastructure in maintaining long-term follow-up<sup>185</sup>, we investigated the use of mobile phones as a way to remain in contact and

explored the potential for launching a mobile-phone-based data-collection strategy. This would allow for the autonomous, off-site completion of short self-completed surveys to be repeated frequently in private and when convenient, thereby reducing the recall period between the survey and actual behaviour.

#### **4.1.4 Mobile-phone surveys**

Through a mobile survey “app” (such as Mobenzi Researcher<sup>186</sup>), mobile handsets can be used both in a traditional telephonic interviewer-led survey (mobile-phone-assisted personal interviewing – MPAPI) and as self-complete surveys (mobile-phone-assisted self-interviewing – MPASI). As the handsets are portable and carried around on one’s person, surveys can be completed in close temporal proximity to the behaviours of interest. The challenge of loss to follow-up experience when participants migrate out of the study area may also be avoided.

There is a paucity of data reported in the literature that details MPASI survey characteristics. This paper reports on our first effort to collect data from these young adults using an MPASI platform. The objectives of the study were to assess feasibility, calculate response rates and determine item-response characteristics, and monitor response stability across two waves of data collection.

## **4.2 Methods**

### **4.2.1 Study population and sample**

The study was designed as a prospective intervention evaluation. One thousand handsets were purchased and loaded with the Mobenzi Researcher survey software. These

handsets were then distributed between January 2009 and November 2009 during regularly scheduled BT20 participant visits. At the visits participants consented to their involvement in the study and were then guided through a practice survey by trained research assistants and informed that they would be allowed to keep the handset at the conclusion of the study.

A telephonic follow-up survey, using one of the alternative numbers available in the participant's file, was undertaken in December 2009 to validate that participants who had received a donated handset were still able to receive and complete surveys. As this was the first attempt by BT20 to conduct an MPASI survey, and as the software systems were untested, a targeted sampling frame was used to randomly select 435 participants confirmed in the telephonic follow-up to have survey-ready handsets (see Figure 4.1). All 435 participants were born within the Soweto municipality between April and June 1990. This gives the sample a median age of 20 at the time of the mobile survey<sup>167</sup>. The sample was composed of 227 females (52.2%) and 386 (88.7%) respondents were black. Ethical approval for the study was obtained from the Committee for Research on Human Subjects (Medical) Witwatersrand Human Ethics Committee.

#### **4.2.2 Participant training**

Training on the mobile survey was provided on receipt of the phone to all participants through a practice survey. This survey was designed to familiarise participants with the process of downloading, completing and then uploading responses through the survey software. It consisted of six items that contained examples of date, single-choice, multi-choice, free-text and numeric question types. Research assistants demonstrated to

participants the procedure of downloading the practice survey onto the donated handset. Participants then worked through the survey and were shown how to upload their response.

### **4.2.3 Questionnaire**

Two surveys, each consisting of 60 items, were designed from questions drawn from the BT20 Year 18 Young Adult Core Questionnaire and the Year 18 Adolescent Self-completed Questionnaire. Survey one also included the 10-item short-form Marlowe-Crowne Social Desirability Scale (M-C 2) proposed by Strahan and Gerbasi<sup>187</sup>. Each survey consisted of two date questions (e.g. "Please enter your date of birth"), 39 multiple-choice questions, (e.g. "When do you switch your phone off? (please choose all that apply)"), 18 numeric questions (e.g. "How many close friends do you have?") and one free-text field (e.g. "What were you busy doing immediately before you started to answer this survey?"). Although Mobenzi Researcher makes it possible to mark any of these question types as "required", the questionnaire was designed to allow participants to skip questions at their discretion.

### **4.2.4 Mobenzi Researcher software**

The commercially available Mobenzi Researcher survey software was provided by Clyral ([www.mobenzi.com/researcher/](http://www.mobenzi.com/researcher/)). Mobenzi Researcher is a Java 2 Micro Edition (J2Me) application that runs on both Symbian and Android operating systems. It provides full survey functionality, including the ability to create various question types, mark fields as mandatory and intelligently manage survey branching. The mobile-phone app submits survey responses over either a mobile (e.g. GPRS) or wireless network. When a mobile

Internet connection is not available, responses are stored on the handset until such time as a connection is re-established. Response credits cost US\$0.0125 per item. Therefore, to submit one 60-item survey would cost the researcher US\$0.75 worth of credit and the participant US\$0.002 worth of airtime.

#### **4.2.5 Procedures**

During training and handset distribution, users were informed that they would be notified by a short messaging service (SMS) text message when a survey was ready for them to complete. Two notifications were dispatched, two weeks apart, requesting the completion of a 60-item survey. The first test message was sent on April 20, 2010 and the second on May 5, 2010. Both notifications included an airtime voucher (R5; \$0.69) to ensure that participants were able to respond, as this was indicated to us in pilot work to be a potential constraint. After Mobenzi Researcher received an uploaded survey response, a text message was sent back to the participant to thank him or her for their time.

#### **4.2.6 Data analysis and outcome indicators**

Dependant data collected included handset attrition, SMS messages set, delivered and failed. Key indicators of “success” were defined for each step in process. This provided the means by which to evaluate the outcomes of the study. These indenendant variables were: 1) the number of participants possessing donated handsets capable of submitting surveys; 2) the proportion of these donated handsets, included in the study, contactable by text message; 3) the proportion of survey responses received from donated handsets that were contactable by SMS; and 4) item-response rates. The study did not set out to

appraise the quality of responses. These data were then analysed using proportions and frequencies.

### 4.3 Results

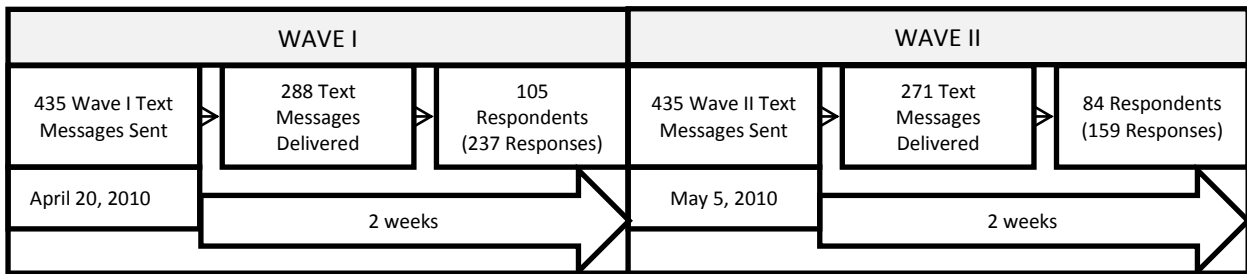
#### 4.3.1 Number of participants possessing handsets capable of submitting surveys

During 2009, 1 000 handsets were purchased, of which 996 were distributed to participants (in the bulk delivery two handsets were found to have faulty batteries and two SIM cards were not adequately registered). Among the 996 participants who received handsets, a telephonic follow-up survey conducted soon after the last handset was distributed revealed that 25 (2.5%) had inserted a new SIM card, 32 (3.2%) had deleted the Mobenzi Researcher (MR) software and 205 (20.6%) reported having lost, stolen, broken or given away their phone.

|                              | Attrition | Total Handsets Available |
|------------------------------|-----------|--------------------------|
| MPs purchased                |           | 1000                     |
| Not distributed              | 4         | 996 (99.6%)              |
| Installed new SIM            | 25        | 971 (97.1%)              |
| Deleted software application | 32        | 939 (93.9%)              |
| Lost and stolen              | 205       | 734 (73.4%)              |
| Random sample                |           | 435                      |

**Figure 4.1. Phones not available for use from the 1 000 donated handsets that were distributed**

#### 4.3.2 Texting the 435 handsets included in the study



**Figure 4.2. Schematic overview of two data-collection waves**

Of the 435 text-message notifications sent out in wave one, 288 (66.2%) were reported by the mobile network as delivered. The remaining 147 messages were undeliverable due to the phone no longer being active on the network. It was not possible to determine whether this was due to the phone having been switched off for an extended period or to the SIM card having been removed from the handset. Similarly, 271 (62.3%) of the 435 second-wave SMS's were delivered (Figure 4.2). Twenty-one (4.8%) participants received SMS one but were not reachable for SMS two; four (0.9%) received SMS two but were not reachable for SMS one; and 267 (61.4%) were reached by both SMS one and two.

In response to the text notification, eight text messages replies were received (Table 4.1). Seven of the eight messages were requests for assistance.

**Table 4.1. Example of text message responses received from Birth to Twenty study participants in response to survey notification SMS messages**

| <b>Message Received</b>                                                                                                                                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| I have changed phones and cant logg on anymore ...                                                                                                                |
| Please give me the website to download it again coz i cant find mobile researcher.                                                                                |
| Which survey                                                                                                                                                      |
| Hi i sent u an sms before informing u that i lost my cellphone last year.                                                                                         |
| My suvey is deleted,cn u send or gv me the address 2 download it!                                                                                                 |
| I am unable 2 complete survey because da new fone dat im using does nt hav da survey. Thanx                                                                       |
| Hallo BTT can you guys plz help me do my matric?                                                                                                                  |
| Hi. The application was deleted after the phone came back from repairs. Can i pls have the directions on how to download mobile researcher to the phone... Thanks |

#### **4.3.3 Survey response rates**

Of the 288 handsets that received notification one, 105 (36.5%) individuals returned a response. Of the 271 handsets that received notification two, 84 (30.9%) individuals returned a response. Although only one response was expected for each survey, some individuals responded multiple times. During wave one, the 105 responding individuals sent in a total of 237 responses. During wave two, 159 responses were received from 84 individuals. The median number of responses per participant was one, with a maximum of 24. Among the 105 respondents who uploaded a response in wave one, 59 (56.0%) also submitted a response in wave two. There were 25 (of the 84 wave-two respondents) who only submitted a wave-two response (29.8%).

#### **4.3.4 Item response rates**

Using survey data generated by Mobenzi Researcher it was possible to tell that surveys received in wave one were completed in a median time of 18 minutes (IQR 8.11) while surveys in wave two had a median time of 17 minutes (IQR 7.88). This means that roughly three questions were answered every minute. All responses were completed in English. Breakoffs, defined as surveys in which participants chose to quit the survey before the last question, were 4.4% and 3.6% respectively for the two waves.

Of the 105 individuals who responded to survey one, 97.8% completed the single free-text question, 95.6% completed the 39 MCQ-type questions and 91.8% the 18 numeric questions. Date fields were prepopulated by the software, resulting in 100% completion. Very similar results are seen in the 84 survey-two responses, with 97.7% completing the single free-text question, 95.2% the MCQ-type questions and 88.5% the numeric questions. The average length of the free-text responses was three words for both survey one (15 characters, std. dev. 9.1) and two (3 words, 14 characters, std. dev. 8.7).

#### **4.4 Discussion**

Within 11 months of the first donated mobile-phone handsets being issued to the 996 BT20 participants, 734 (73.7%) reported that their phones were still capable of sending and receiving Mobenzi Researcher surveys. Of the 435 randomly selected participants who were enrolled in this study, response rates of 36.5% (105 respondents to the 288 delivered text messages) and 30.9% (84 respondents to the 271 delivered text messages) were achieved during waves one and two, respectively. Item non-response was found to be higher for numeric than free-text, date or multiple-choice items.

Two clear challenges emerged from these results. First, the effort required to provide, support and manage large numbers of mobile phones for the purpose of data collection among young adults should not be underestimated. There was a high initial rate of attrition, with 205 handsets reported lost, broken or stolen within an 11-month period. Technical support for participants having problems with their handset was an everyday activity. In addition, it was challenging to keep track of phones and SIM cards as participant reports indicated that SIM-card swapping and ownership of multiple SIM cards is common. Prior to the start of the study, all 435 phones were validated as being available for participation, but four months later only two thirds of these handsets were contactable by SMS.

The second challenge that emerged was the struggle that participants faced when trying to use the survey software. The authors had previously used Mobenzi Researcher very successfully for MPASI. However, this study highlighted that the process and user-interface design of the software needs to be modified and simplified significantly for MPASI, even with young people who are proficient in the use of mobile phones and associated applications. To complete a survey successfully required participants to receive and read the SMS; know how to find the software application on the phone; load the application; complete the appropriate permissions (for example, allowing the software to access the network); download the correct survey, given that prior surveys, including tutorials, remained on the phone if not deleted; successfully work through the survey; and, finally, upload the completed response. Despite thorough training, text messages sent to the research team by participants suggested that the process was too complicated for a significant number of participants to complete. Participants reported deleting the

application, not being able to find it or the surveys and not having the application available in the phone they had inserted the donated SIM card into.

On the positive side, participants showed huge enthusiasm for the mode of response. At their own expense, and without further prompting, many participants continued to submit responses for over a year after the original data-collection wave had ended. Furthermore, the response rate of 36.5% compares favourably with rates reported for web surveys. In a recent meta-analysis of 39 mode-comparison studies from the past 10 years, Shih and Fan<sup>90</sup> report average response rates of 45.0% for mail surveys, and 34.0% for web surveys. In a second meta-analysis<sup>91</sup> a slightly higher response rate of 53.0% was found for mail surveys, with 33.0% for email surveys.

In order to assess whether the novelty of responding to a mobile survey may have artificially raised the wave-one response rate, we compared participants who sent in a response to wave one with participants who also submitted a wave-two response. Of the 105 respondents who uploaded a response in wave one, 59 (56.2%) also submitted a response in wave two. However, there were 25 (of the 84 wave-two respondents) who only submitted a wave-two response (29.8%). It is not clear whether technical difficulties or a decline in novelty played a role in the submission of data in only one of the two waves.

The item-type analysis revealed significant differences in the type of questions that participants skipped. Numeric fields were left blank more often than other question types. Free-text fields were mainly populated with short responses made up of SMS-

abbreviated words. The question format that showed the most potential was the multiple-choice format, since it was skipped less frequently than numeric fields and provided more useful data than the free-text field responses. Future mobile-survey designers should consider the multiple-choice question as a key component of their survey.

Although the use of “apps” is widespread in high-income countries where smartphones are more common, the “app” approach to MPASI in the youth of low and middle-income countries (at this time) should be considered tentatively. SIM swapping, phone “loss” and app deletion all contributed to the low response rate achieved. Similar challenges were found in a recent review of SMS applications in the prevention of disease in low-resource countries<sup>188</sup>. Barriers included high mobile-phone turnover and system maintenance difficulties. One avenue for further research would be to explore leveraging individuals’ personal handsets while at the same time adjusting the complexity of the survey-delivery software. In 2009, 66.2% of the cohort reported owning their own personal handset<sup>189</sup>. Using individuals’ own handsets would alleviate several challenges faced when managing study-provided handsets, although not knowing the type of phones owned by the cohort would introduce complications in relation to software development. An alternative approach that may better suit young adults in low and middle-income countries, and one that would avoid many of the challenges outlined in this paper, would be to conduct MPASI through an SMS-based survey system or through social network sites<sup>190</sup>. This would be more familiar to participants than having to install and manage sophisticated survey software, and at the same time would ensure that items remained short and easy to comprehend.

In conclusion, although this study found many challenges in effectively collecting data using MPASI, the findings do not exclude the possibility of MPASI being a suitable tool for the population group. If correctly understood and utilised, mobile phones remain a viable way to collect data and stay in contact with the majority of the migrated BT20 cohort, especially if combined with other modes of data collection (e.g. email, web survey and mail-in surveys). The challenge is now to find a more streamlined way of exploiting this opportunity. However, even unchanged, the current system could successfully be used with a significant number of participants, thus lessening the cost and effort of future data-collection waves, where personal contact could be pursued for only those participants who struggle with the current system. The ubiquity and portability of mobile phones make them an enticing tool for the collection of sensitive information from young adults. The results obtained in this paper support this approach. Further research is needed to explore other approaches to MPASI among young adults in low and middle-income countries.

## Chapter 5

---

### **Collecting maternal health information from HIV-positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa**

*In review: Journal of Medical Internet Research*

## Chapter 5: Collecting maternal health information from HIV-positive pregnant women using mobile-phone-assisted face-to-face interviews in Southern Africa

---

### 5.1 Introduction

One half of people living with HIV globally are women and 76% of all HIV-positive women live in Sub-Saharan Africa<sup>191</sup>, where almost half of the world's maternal, newborn, and child deaths occur, with 4.7 million children and 276 000 women dying annually<sup>192,193</sup>.

Mothers living with HIV (MLH) are particularly vulnerable and at risk of adverse maternal outcomes, with at least 20% of maternal deaths being HIV-related<sup>191</sup>. South Africa makes up a large proportion of this disease burden, with 3.2 million women living with HIV<sup>194</sup>, of whom 200 000 annually are pregnant<sup>195</sup>. The province of KwaZulu-Natal (KZN) has the highest HIV prevalence in South Africa<sup>196</sup>. While the national prevalence is around 11%<sup>196</sup>, 40-60% of pregnant women in rural KZN are HIV-positive<sup>197,198</sup>.

In response to this crisis, South Africa has implemented the Prevention of Mother-to-Child Transmission (PMTCT) programme as recommended by the World Health Organization (WHO). The PMTCT programme requires newly pregnant women to complete a series of sequential steps, sometimes known as the PMTCT cascade<sup>199</sup>, that are aimed at first diagnosing and then treating HIV infection<sup>200</sup>. Under ideal circumstances, where no barriers exist to the completion of all tasks, PMTCT has been shown to be highly effective at reducing HIV transmission to below 2% at childbirth<sup>201</sup>. Under the less than ideal circumstances faced by many MLH in low and middle-income countries (LMIC), transmission can occur in approximately one out of every four deliveries<sup>202</sup>. Some of the challenges that make adherence to each of the PMTCT tasks

challenging are being unable to afford transportation to the clinic, fear of stigmatisation, increased household conflict and lack of partner support<sup>199</sup>. To better understand these barriers and their impact on the loss of women through the PMTCT cascade requires accurate, timely and detailed information. Yet gathering data that tracks women's progress through PMTCT to determine how well the system is performing is beset with its own challenges.

One of the major challenges faced in the collection of high-quality data is human-resource constraints. Trained and qualified health staff are in short supply in many resource-constrained settings<sup>203</sup>. Task-shifting strategies used to address these shortages have led to a perceived increased burden for data collection and collation<sup>204</sup>. Minimal support, delayed feedback, little understanding of the usefulness of the data and no interpretation of raw scores mean that the completion and submission of multiple paper-based registers for statistical purposes is often regarded as a low priority for busy and overburdened staff<sup>204</sup>. It is therefore unsurprising that register data has been shown to be fragmented, error prone, incomplete and inaccessible<sup>205,206</sup>.

The usefulness of this register data is further compromised by the fact that it is unlinked, i.e. it is not possible to link individual clinic visits to the same person. This means that although aggregate data may be available, it is difficult to track the path of individuals through the PMTCT cascade. This makes it difficult to identify and understand the bottlenecks in the cascade. These constraints and challenges raise the question of whether there is not a more efficient way to support staff and perform monitoring and

evaluation of the PMTCT programme in geographically remote primary health care facilities<sup>207</sup>.

There is currently a groundswell of interest in the use of mobile phones and information communication technologies in the support of health<sup>133</sup>. The growing body of mHealth literature provides examples of the use of mobile phones as data-collection tools in low and middle-income countries. Although evidence is mixed, on the whole the use of mobile phones as data-collection tools has been found to increase data quality, speed up the turnaround time from collection to analysis, and improve inter-facility communications<sup>124,150,152,208</sup>. Poor follow-up rates could potentially also be improved with mobile phones providing links to participants that are cheaper and more immediate than travel to facilities. These advantages over traditional paper-based clinic registers suggest mobile phones as a potential tool with which to address some of the challenges currently experienced in collecting health information through the PMTCT cascade. The aim of this paper is to describe the feasibility and acceptability of using mobile phones to collect data from pregnant HIV-positive women enrolled into the PMTCT programme in KwaZulu-Natal, South Africa.

## **5.2 Method**

### **5.2.1 Study design**

This study was nested within a larger clinic-based randomised cluster trial known as Project Masihambisane (“let us walk together”)<sup>209</sup>. The primary study aimed to improve mental and physical health outcomes of HIV-positive mothers and their babies by supplementing the PMTCT programme with paraprofessional peer mentors (PM). Using a

mixed-methods design, qualitative data from a single, small focus group was supplemented with two quantitative questionnaires. Paradata, or data on the data-collection process, were gathered in order to examine the feasibility of mobile-phone-assisted personal interviewing (MPAPI).

### **5.2.2 Participants and procedures**

Three groups of participants took part in this study. The first, recruited for participation in the focus group, and consisting of 12 pregnant HIV-positive women living in this region of South Africa, were enrolled through community forums set up for the larger study. Before the focus group was held, each woman was interviewed by a research assistant using a mobile phone. The questions asked ranged from general health questions about the participants' knowledge of HIV to more sensitive questions about recent sexual activity, condom use and disclosure of HIV status to their partners. After completing the questionnaire, the group was brought together to participate in a focus group about their experience. Five questions were designed to facilitate the discussion about general mobile-phone ownership and their reactions to the interview conducted by mobile phone. The focus group lasted for an hour with extensive notes being taken about responses.

The second group of participants was the 16 field research staff recruited and trained on MPAPI for the primary study. Although these participants were all familiar with pen-and-paper questionnaires, none had ever used MPAPI. Before receiving any training on MPAPI, the group was asked to complete a short questionnaire with items from the Technology Acceptance Model proposed by Davis<sup>210</sup>. The scale contains 12 items, six

relating to the perceived usefulness of mobile phones for research purposes and six to the perceived ease of their use (Table 5.1). Each item was rated on a five-point Likert-type scale ranging from “extremely likely” to “extremely unlikely”. Following training, which lasted three days, a post-training assessment was conducted using the same questionnaire. After completing training, field staff were dispatched to eight primary health care clinics in a rural district typical of those found in the region, where they began interviewing pregnant HIV-positive women for the primary study. After a period of three months the questionnaire was administered for a third time to gauge if use of the mobile phone in the field had altered perceptions and attitudes towards the tool.

**Table 5.1. The twelve items of the Technology Acceptance Model Scale**

| <b>Perceived Usefulness</b>  |                                                                                               |
|------------------------------|-----------------------------------------------------------------------------------------------|
| Item 1                       | Using cell phones to collect data in my job would enable me to accomplish tasks more quickly. |
| Item 2                       | Using cell phones to collect data would improve my job performance.                           |
| Item 3                       | Using cell phones to collect data in my job would increase my productivity.                   |
| Item 4                       | Using cell phones to collect data would enhance my effectiveness on the job.                  |
| Item 5                       | Using cell phones to collect data would make it easier to do my job.                          |
| Item 6                       | I would find cell phones useful in my job.                                                    |
| <b>Perceived Ease of Use</b> |                                                                                               |
| Item 1                       | Learning to operate a cell phone would be easy for me.                                        |
| Item 2                       | I would find it easy to get a cell phone to do what I want it to do.                          |
| Item 3                       | My interaction with cell phones would be clear and understandable.                            |
| Item 4                       | I would find cell phones flexible to interact with.                                           |
| Item 5                       | It would be easy for me to become skilful at using a cell phone.                              |
| Item 6                       | I would find a cell phone easy to use.                                                        |

The third and final group of participants were pregnant HIV-positive women recruited through the primary study. Participants were recruited in the Umgungundlovu Health District of KZN. The district, comprised of seven local authorities, has 48 fixed clinics, four community health centres, nine tertiary hospitals and an estimated population of some 995 000 people<sup>204</sup>. From this district, eight clinics were selected through a clinic audit conducted in December 2007. The selection criteria, applied to all potential clinics in the district during the audit, were: presence of other research trials; clinic size; availability of ante-natal and post-natal services; and uptake of ante-natal and post-natal services at the clinic. The clinic audit resulted in the selection of four pairs of clinics for the study, matched on size (small versus large) and geography (rural versus urban). Women who met the eligibility criteria of being 18 years or older, less than 34 weeks pregnant and planning to reside in the study area for the duration of their pregnancy were invited to participate in the primary study. Using a continuous sampling strategy that started approximately halfway through the primary study, 512 of the 1 204 women enrolled into the primary study were recruited to participate in this sub-study. A standard consenting process was used where all potential participants were first introduced to the study using an information sheet. The sheet contained a clear description of the study, the information that would be collected and the procedures that would be used. Written consent was then obtained from each women.

### **5.3 Statistical Analysis**

Data was downloaded as comma-separated values from the online Mobile Researcher database. This CSV file was then imported into SPSS 19. This data was analysed using exploratory data-analysis techniques. Chi-square statistics were calculated difference

between demographic variables and phone ownership along with survey mode preferences. Univariate analysis was used to analyse Perceived Ease of Use and Perceived Usefulness scale data. The quantitative data was supplemented by a thematic analysis<sup>254</sup> performed on the qualitative data generated during focus group discussions.

#### **5.4 Fieldwork Supervision and Ethical Approval**

A team of 16 data collectors administered the questionnaire and two coordinators employed by the primary trial supervised all aspects of the study in the field. The coordinators' role included managing informed consent, supporting field activities and monitoring data quality through quality-assurance checks and ongoing training and supervision of field staff. Ethical approval for the study was obtained from the Witwatersrand Human Ethics Committee. All participants were taken through a study information sheet before signing a form giving written consent to participate.

#### **5.5 Results**

##### **5.5.1 Acceptability**

###### **5.5.1.1 Focus group**

All focus group participants were HIV-positive and pregnant. The majority (9/12) had learnt their HIV status through the PMTCT programme during their first pregnancy. For four of the women this was their first pregnancy and only one of the four had already known that she was HIV-positive. The mean age of the focus group participants was 26 years. Despite an estimated household (not per-capita) income of R1 600 per annum for rural communities in this district<sup>211</sup>, all 12 women in the focus group had access to a

mobile phone at home, with a full two thirds (8/12) personally owning a phone. The four women who did not own a mobile phone had access to a sibling's phone. When asked how important access to a mobile phone was for them, all but one stated that it was extremely important.

Reflecting on the experience of being interviewed using a mobile phone, a number of women reported initial feelings of anxiety about the multimedia capabilities of the phone. In particular, women were concerned about being filmed or photographed by the camera, which had a lens situated on the back of the phone. One woman said:

*I feel it is ok [to be interviewed on the phone] but it is slightly scary. I feel as though you are recording a video of me. But it's alright. It was not nice when the interviewer held the phone too close to me because I was worried about the phone camera... whether I was being recorded.*

Two important factors mediated this experience. Firstly, the model of Nokia phone used influenced how comfortable the interviewees felt. The less familiar looking E61i business phone raised more concerns from participants than the more familiar candy-bar-shaped 2630. Secondly, the interview style adopted by the interviewer impacted significantly on the experience. Although all interviews were one-on-one, some interviewers took up a position directly opposite the woman, while other interviewers adopted a side-by-side position. For those women who experienced the side-by-side interview, being filmed was reported as being less of a concern. Women who were interviewed by an interviewer sitting directly opposite them felt that side-by-side interviewing would have alleviated much of their anxiety about being photographed. The side-by-side approach had the

added benefit of allowing the women who reported being curious about the questionnaire to see what was being asked and captured on the phone. One participant said, *I enjoyed that the interviewer let me see the questions and shared me what was being recorded*. This information was incorporated into the primary trial by training fieldworkers to offer MLH the opportunity to sit alongside them during the interview.

#### **5.5.1.2 Cross-sectional questionnaire**

From October 2010 to March 2011, Project Masihambisane conducted 708 interviews with 520 participants (some participants completed interviews at two different time points). Of these 520 participants, 512 agreed to take part in the mobile-phone questionnaire. The mean age of participants was 27 years, and participants spoke primarily isiZulu (93.4%). A small minority had either no schooling (1.6%) or a tertiary qualification (5.8%). Of the remaining 92.6%, 36.4% had completed primary schooling, 39.9% had started but had not completed secondary schooling, and 16.3% had completed secondary school. Roughly the same percentage of women were married (10%) as were living completely alone (11.9%); the others were either married and living apart, or single and in a co-resident or non-co-resident relationship. There was a 60/40 split between women living in rural versus peri-urban areas; 35.5% had running water on the premises, 59.2% had flush toilets and 82.6% had electricity.

Despite 42.4% of the sample being unemployed, 84% stated that they owned a mobile phone. The majority of these handsets were manufactured by Nokia, ran the Symbian operating system and were able to run Java 2 Micro Edition (J2ME) applications (87.4%), a version of the Java programming language designed to run on resource-constrained devices such as mobile phones. Self-report sets the mean length of ownership at 3.3 years

with a median amount of R30 being spent on airtime per month. Table 5.2 compares the demographic profile of those who own versus those who do not own a mobile phone. Chi-square tests found significant associations between age ( $\chi^2$  6.34, df 2,  $p=0.042$ ), education ( $\chi^2$  33.52, df 2,  $p=0.000$ ), socioeconomic status (SES) ( $\chi^2$  27.65, df 2,  $p=0.000$ ), occupation, housing characteristics (water:  $\chi^2$  9.79, df 1,  $p=0.002$ ; electricity:  $\chi^2$  9.6, df 1,  $p=0.002$ ; flush toilet:  $\chi^2$  12.68, df 1,  $p=0.000$ ) and ownership of a mobile phone. Pregnant women with HIV are less likely to own a phone if they are 20 years old or younger, have a primary school education, are in the lowest SES group and do not have household luxuries such as running water, electricity or flushing toilets.

On a three-point scale, the participants rated the experience of the mobile-phone questionnaire as positive, neutral or negative; 485 (94.7%) of the 512 participants reported a positive experience (Table 5.3). There were 24 (4.7%) neutral responses and three (0.6%) who disliked the mobile-phone questionnaire. With respect to questionnaire mode preference, one third (29.5%) had no preference, two thirds preferred the mobile-phone questionnaire (67.8%) and 2.7% preferred pen-and-paper questionnaires.

A one-way ANOVA revealed a significant increase in positive rating when comparing the number of interviews completed before taking the questionnaire; that is, people who had completed four MPAPI interviews before taking the questionnaire were significantly more likely to rate the questionnaire experience positively than those participants taking an MPAPI questionnaire for the first time ( $F = 6.795$ ,  $df=3$ ,  $p < 0.01$ ).

**Table 5.2. Demographic characteristics of Mothers Living with HIV (n=512) surveyed in 2010**

|                                                         | Do Not Own Phone<br>(n=82) |          |               | Own Phone (n=430) |          |               |
|---------------------------------------------------------|----------------------------|----------|---------------|-------------------|----------|---------------|
| <b>Respondent Characteristics</b>                       | <b>N</b>                   | <b>%</b> | <b>95% CI</b> | <b>N</b>          | <b>%</b> | <b>95% CI</b> |
| <b>Age</b>                                              |                            |          |               |                   |          |               |
| 20 yrs and younger                                      | 19                         | 23.2     | 15.4 – 33.4   | 55                | 12.8     | 10 – 16.3     |
| 21-30 yrs                                               | 39                         | 47.6     | 37.1 – 58.2   | 247               | 57.4     | 52.7 – 62     |
| 31-40 yrs                                               | 24                         | 29.3     | 20.5 – 39.9   | 128               | 29.8     | 25.6 – 34.3   |
| <b>Education</b>                                        |                            |          |               |                   |          |               |
| Primary Schooling*                                      | 28                         | 34.1     | 24.8 – 44.9   | 45                | 10.5     | 7.9 – 13.7    |
| Secondary Schooling*                                    | 53                         | 64.6     | 53.8 – 74.1   | 356               | 82.8     | 78.9 – 86.1   |
| Tertiary Schooling                                      | 1                          | 1.2      | 0.2 – 6.6     | 29                | 6.7      | 4.7 – 9.5     |
| <b>Socioeconomic Status</b>                             |                            |          |               |                   |          |               |
| Lower*                                                  | 25                         | 30.5     | 21.6 – 41.1   | 48                | 11.2     | 8.5 – 14.5    |
| Middle                                                  | 53                         | 64.6     | 53.8 – 74.1   | 293               | 68.1     | 63.6 – 72.4   |
| Upper*                                                  | 4                          | 4.9      | 1.9 – 11.9    | 89                | 20.7     | 17.1 – 24.8   |
| <b>Occupation</b>                                       |                            |          |               |                   |          |               |
| Full-time*                                              | 6                          | 7.3      | 3.4 – 15.1    | 89                | 20.7     | 17.1 – 24.8   |
| Part-Time                                               | 27                         | 32.9     | 23.7 – 43.7   | 121               | 28.1     | 24.1 – 32.6   |
| Unemployed                                              | 44                         | 53.7     | 42.9 – 64     | 174               | 40.5     | 35.9 – 45.2   |
| Other                                                   | 5                          | 6.1      | 2.6 – 13.5    | 46                | 10.7     | 8.1 – 14      |
| <b>Region</b>                                           |                            |          |               |                   |          |               |
| Local Government Administration                         | 36                         | 43.9     | 33.7 – 54.7   | 171               | 39.8     | 35.3 – 44.5   |
| Traditional Administration                              | 46                         | 56.1     | 45.3 – 66.3   | 259               | 60.2     | 55.5 – 64.7   |
| <b>Housing</b>                                          |                            |          |               |                   |          |               |
| Running water on premises*                              | 48                         | 58.5     | 47.7 – 68.6   | 324               | 75.3     | 71.1 – 79.2   |
| Electricity*                                            | 58                         | 70.7     | 60.1 – 79.5   | 365               | 84.9     | 81.2 – 88     |
| Flush toilets*                                          | 34                         | 41.5     | 31.4 – 52.3   | 269               | 62.6     | 57.9 – 67     |
| * Significant different using a 95% confidence interval |                            |          |               |                   |          |               |

**Table 5.3. Questionnaire experience and preference of Mothers Living with HIV (N=512) surveyed in 2010**

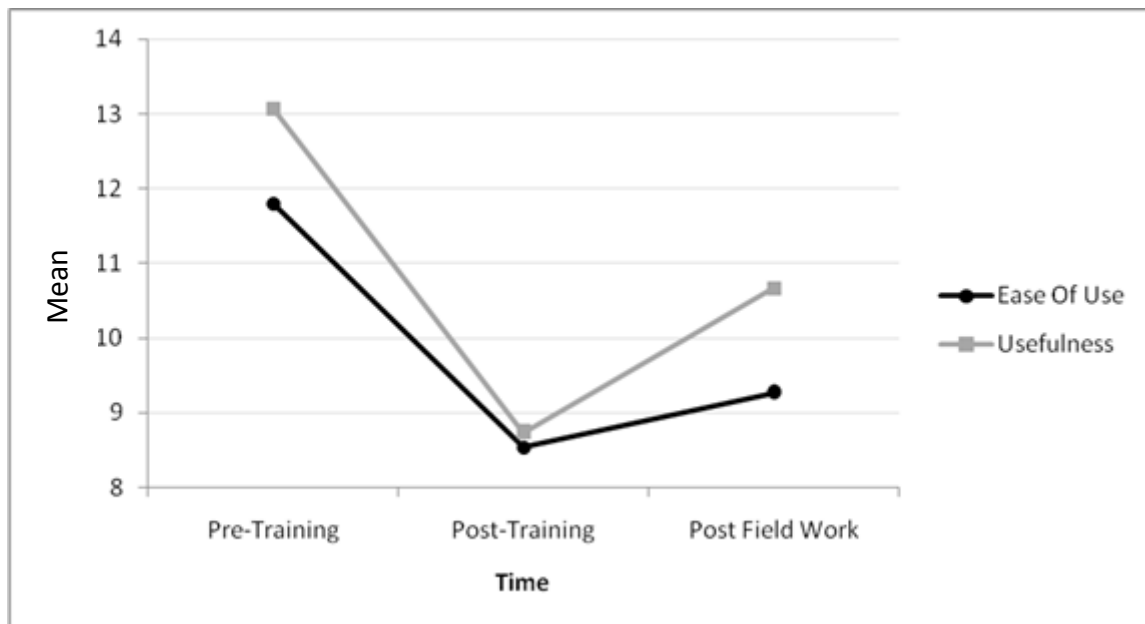
| Respondent Characteristics             | Do Not Own Phone<br>(n=82) |      |             | Own Phone (n=430) |      |             |
|----------------------------------------|----------------------------|------|-------------|-------------------|------|-------------|
|                                        | N                          | %    | 95% CI      | N                 | %    | 95% CI      |
| <b>Questionnaire Experience</b>        |                            |      |             |                   |      |             |
| Positive                               | 76                         | 92.7 | 84.9 – 96.6 | 409               | 95.1 | 92.6 – 96.8 |
| Neutral                                | 6                          | 7.3  | 3.4 – 15.1  | 18                | 4.2  | 2.7 – 6.5   |
| Negative                               | 0                          | 0    | 0 – 4.5     | 3                 | 0.7  | 0.2 – 2     |
| <b>Questionnaire Preference</b>        |                            |      |             |                   |      |             |
| Phone                                  | 50                         | 61   | 50.2 – 70.8 | 297               | 69.1 | 64.5 – 73.3 |
| Paper and Pen                          | 5                          | 6.1  | 2.6 – 13.5  | 9                 | 2.1  | 1.1 – 3.9   |
| No Preference                          | 27                         | 32.9 | 23.7 – 43.7 | 124               | 28.8 | 24.8 – 33.3 |
| <b>Concerned about privacy</b>         | 9                          | 11   | 5.9 – 19.6  | 23                | 4.1  | 3.6 – 7.9   |
| <b>Concerned about confidentiality</b> | 6                          | 7.3  | 3.4 – 15.1  | 21                | 4.9  | 3.2 – 7.4   |

#### 5.5.1.3 Field research staff

Sixteen field research staff were recruited, trained and dispatched to eight health facilities as part of the primary trial. All participants were female, and the average age was 27. Phone ownership was high, with 15 out of the 16 participants owning their own mobile phone. All of the field research staff had completed their secondary education, with 14 having completed a three-year tertiary degree. All had previous experience with paper-and-pen questionnaires but none reported ever having performed a mobile-phone-assisted interview.

A two-way, repeated-measures ANOVA was performed on the two-factor TAM scale (ease of use and usefulness) across three time points (pre-training, post-training, post-use).

There was a significant main effect for the time point at which the data was collected ( $F(1.189, 16.649) = 8.62, P = .007$ ). There was, however, no significant main effect between perceived ease of use and usefulness ( $F(1, 14) = .924, P = 0.35$ ) or any interaction effect between these factors and time ( $F(2,28) = .621, P = 0.54$ ). The findings are presented in Figure 5.1. With a maximum score of 55 implying extreme dislike of the tool, field research staff's attitudes towards ease of use and usefulness of mobile phones before training were already positive.



**Figure 5.1. Estimated marginal means for the main effect of time**

After the three days of training, staff reported perceiving the mobile phones as significantly easier to use than at the baseline measure. Further, the perception of mobile phones as a useful tool with which to collect data underwent an even greater positive shift. The training also increased positive perceptions of the mobile-phone survey as a useful way in which to collect data. The positive perceptions were linked to the insight

that the mobile phones would allow more accurate data to be collected in a timely manner with less administrative burden than paper forms. After three months of use, perceptions of both the usefulness and ease of use of mobile-phone-assisted questionnaires decreased but remained above levels reported pre-training. This decrease in perceived usefulness and ease of use is thought to be related to a more realistic view of the device's potential being realised through practical experience. It is important to note that real-world use did not significantly reduce perceptions of either usefulness or ease of use.

### **5.5.2 Feasibility**

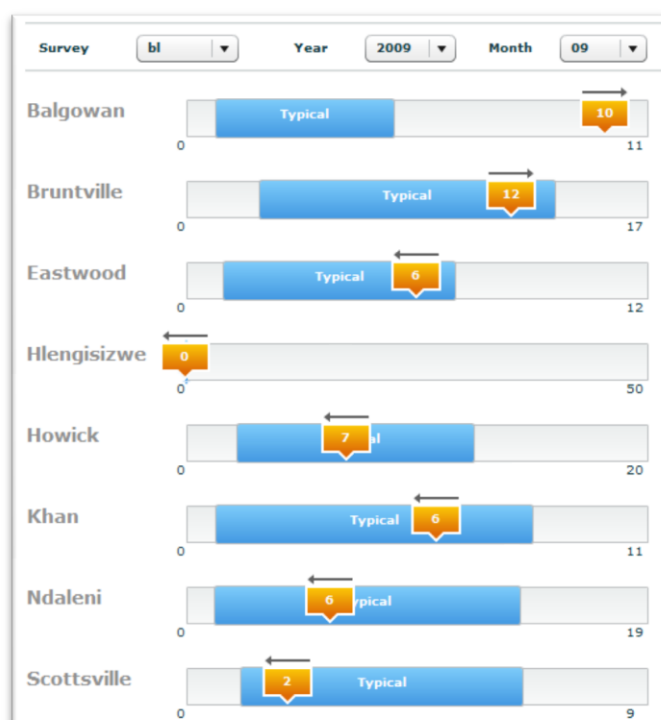
Over the course of the primary study (July 2008 to April 2010), 13 653 mobile-phone-assisted questionnaires were submitted and 13 650 received by the data centre. This number includes the questionnaires that were used to capture three full-length health assessments, the patient-carried maternal health cards and the daily clinic attendance registers. All clinics had good network coverage and questionnaires were received on average 4.2 minutes after being completed.

Participants were identified through the capturing of two unique numeric identifiers. Of the 3 012 full-length questionnaire assessments, 57 (1.9%) were not able to be linked to a participant due to errors in the first 11-digit unique identifier. All 57 cases were resolved by referring to the second five-digit identifier.

Of the 18 mobile phones that were bought for the study, one handset was stolen during a mugging and one became unusable 13 months into the study after a hardware failure.

The remaining handsets are still functional and in use. Over the course of the study there were very few hardware and/or software challenges. To deal with the possibility that phones would be used for private calls and other electronic communications, mobile phones were loaded with a fixed amount of airtime (R200) at the beginning of every month. When this allocation was consumed, staff were responsible for airtime top-ups.

Automatic processing of submitted data enabled field team leaders to monitor the number of interviews, length of interview and quality of interview response being submitted by field staff on a web-based dashboard page. An excerpt from the remote-monitoring web dashboard is given in Figure 5.2. At a glance, field coordinators were able to see the minimum and maximum number of interviews per clinic, typical monthly range, current performance and a trend indicator. These figures could be filtered by questionnaire, year and month. This quick access to reliable data made it possible to respond rapidly to the ever changing challenges faced by staff working in a rural clinic environment. The approach would also be more easily scaled up to the monitoring of many more clinics than would alternative management strategies based on paper and data capturing into a locally stored database.



**Figure 5.2. Webpage dashboard snippet summarising research assistant interview performance for a baseline interview conducted in September 2009**

## 5.6 Discussion

The present study evaluated the acceptability and feasibility of using mobile phones to collect data from HIV-positive pregnant women who were enrolled into the PMTCT programme in KwaZulu-Natal, South Africa. Results suggest that the method was accepted by both affected women and the interviewers trained to collect the mobile-phone questionnaires. The feasibility of collecting data from remote primary health care facilities was found to be high in this particular context.

### **5.6.1 Acceptability**

#### **5.6.1.1 Women living with HIV**

In preparation for interviewing pregnant women living with HIV in a clinic environment, a focus group with 12 women was undertaken. The experience of completing a face-to-face mobile-phone-assisted interview was acceptable to the majority of women. The major concerns raised in the focus group discussion suggest two important recommendations. First, if mobile phones are to be used in health settings as data-collection tools, it is advisable for health staff to adopt an open, side-by-side interviewing style. By placing themselves alongside the women, concerns about data privacy, participant confidentiality and interview anxiety can be minimised. Secondly, using the informed-consent process to explain exactly how data will be collected, how it will be transmitted and who will have access to this private information may reduce participant anxiety about the MPAPI process. Taking these considerations into the clinic-based questionnaire, 512 women were asked to complete a short questionnaire about their experience. Overall, the acceptability of this interview method was confirmed with 94.7% of women, who viewed the experience positively. Only three women reported the experience as being negative for them. Given a choice as to whether future interviews should be conducted using mobile-phone-assisted interviewing or paper and pen, only 2.7% expressed a preference for paper and pen.

Ownership of mobile phones among pregnant MLH, while high at 84%, was significantly associated with age, education, socio-economic status and occupation. This has both positive and negative implications. Positively, the potential of mHealth to reach MLH was confirmed with eight out of every ten women owning a mobile phone. Caution must,

however, be exercised to ensure that mHealth programmes, which often aim to increase access to marginalised and under-resourced individuals, do not inadvertently further disadvantage the most disadvantaged MLH.

#### **5.6.1.2 Field staff**

Perceived ease of use (PEO) and perceived usefulness (PE) of MPAPI were consistently rated highly by clinic-based research assistants. It is not clear whether this rating is in response to a dislike of paper-based data collection or an affinity for mobile phones. Although PEO and PU dropped after three months of field use, they remained above pre-training perceptions. The decrease is likely linked to the fact that real-world use is often more limited than that portrayed under the ideal conditions made possible in a controlled training environment. Nevertheless, the decrease was not significant and remained high for both perceived usefulness and ease of use. This finding suggests that after extended clinic-based use, the mobile-phone-assisted interviews remained useful and easy. Training was found to increase PEO and PU significantly, and highlights the importance of providing a comprehensive training package in order to familiarise staff with both the concepts and practice of MPAPI.

#### **5.6.2 Feasibility**

The feasibility of using mobile-phone-assisted interviews as part of a clinic-based PMTCT programme was monitored for 21 consecutive months. During this time over 13 000 data forms were submitted with only three reported as being sent by research assistants and not received. Very few phones were lost, broken or stolen, and network coverage was good at all eight clinic facilities. A number of strategies were used to control airtime use,

the most effective being a fixed amount of airtime loaded onto the handset at the beginning of the month, and research assistants topping up if they depleted this amount. Initial concerns about charging handset batteries were unfounded as all clinics had electricity.

These limited challenges were supplemented by a number of perceived benefits. The management of geographically dispersed staff was improved through the use of real-time data presented in visually appealing summaries. This allowed problems to be flagged much earlier than if data were captured on paper and collected from the clinics on a monthly basis. Although not used in this study, the ability to geo-tag questionnaire responses could also assist in the management of community health workers providing home-based care. Questionnaire management was also found to be improved with a virtual dashboard giving valuable up-to-date information on key project indicators. Whenever a new questionnaire needed to be introduced or changes needed to be made to an existing questionnaire, updates could be rolled out immediately to each clinic without the need to recall out-of-date paper forms. This problem besets current paper-based registers. Whenever a new indicator needs to be collected or a change in indicator definitions occur, old forms continue to be submitted by facilities whose staff members either do not receive the updated forms or who prevent wastage by continuing to use old form stock. Finally, this study found it feasible to track and link the PMTCT data of participants over a 10 to 14-month period. Although 2% of the records could not be linked due to data-entry errors in the unique identifier, the use of a second identifier resolved these cases.

## **5.7 Relevance to PMTCT Cascade**

Paper-based records that inform the KwaZulu-Natal District Information System have been shown to be of poor quality and in a format that makes the tracking of women through the PMTCT cascade difficult<sup>206</sup>. While data collection using mobile phones is proposed as an alternative in other settings<sup>208,212</sup>, little is known about the acceptability and feasibility of using mobile phones to collect data from MLH in this setting. The findings presented here suggest that mobile-phone-assisted personal interviewing is indeed acceptable to both MLH and interviewers. The use of mobile phones for data collection also proved feasible and provided a number of advantages to the data-collection process over traditional ways of collecting data in a rural clinic setting. One key benefit for the PMTCT programme was the ability to link women's data over a number of visits. This made it possible to track enrolment, uptake and loss of individual women through the PMTCT cascade.

## **5.8 Limitations and Future Work**

The results obtained in this study suggest that MPAPI may be both feasible and acceptable to use within the South African PMTCT programme. South Africa is unique among African countries in its broad mobile network coverage, well-maintained mobile infrastructure and electrification of rural health facilities. How well this approach to the collection of PMTCT data would perform in the context of other low and middle-income countries is therefore a limitation of this work.

The bi-directional communication abilities of mobile phones also hold promise as a way to provide regular and relevant feedback to clinics. With a large proportion of women

owning mobile phones, future studies should explore the acceptability and feasibility of extending PMTCT data collection into the home by using mobile-phone-assisted self-interviewing.

## Chapter 6

---

### **Collecting health research data: Comparing paper-and-pen to mobile-phone-assisted data collection**

Accepted by *Field Methods*

## **6.1 Introduction**

Survey data is routinely collected in health research. Over time the methods used to collect these data have evolved due to the influence of available computer technology. In many parts of the world rapid technological diffusion has resulted in computer-aided interviewing (CAI) becoming the dominant method for completing both self- and interviewer-administered surveys<sup>213</sup>. The major impacts that technological advancements have had on survey methods include: a move from interviewer administration to self-administration; the enhancements of surveys through audio-visual content; survey device mobility; and the ability to measure the behaviour of interest continuously<sup>62,214</sup>. These trends, driven by advancements in the fields of distributed computer networking and mobile computing<sup>29</sup>, have supported the development of computer-aided survey techniques such as CAPI, ACASI, mobile CASI, IVR systems and web and email surveys.

Computer-aided surveys have been shown to have a number of advantages over traditional paper surveys. These include automated skip patterns reducing the cognitive burden placed on the interviewer, the incorporation of logical consistency and range field checks, and the speed at which data can be made available for analysis<sup>215</sup>. Computer-aided interviewing has also been found to have some unique challenges to overcome in the context of low and middle-income countries. Traditional CAI devices are typically expensive, better suited to controlled environments that are electrified and generally require highly skilled individuals to manage<sup>216</sup>. These requirements can make

implementation difficult in resource-constrained contexts where infrastructure such as electricity and wired broadband Internet, expensive computer hardware and skilled human resources are often not widely available. To gain the advantage of electronic data collection while remaining mindful of the contextual constraints, many researchers have piloted the use of PDAs in developing countries<sup>120,217,218</sup>.

More recently, mobile phones have also begun to show promise as data-collection devices in low and middle-income countries. For example, a public-health sentinel surveillance system using mobile phones was shown to increase the detection and response speeds to viral outbreaks of malaria in Madagascar<sup>145</sup>. Although often considered equivalent, important differences exist between typical PDAs and mobile phones. Most importantly, PDAs do not connect wirelessly to the Internet using a mobile network connection. They must be brought into close physical proximity to another computer onto which the data is downloaded<sup>219</sup>. With mobile phones, the cellular network allows data to be uploaded in close proximity to the geographic location of the survey. Second, mobile phones have become much more ubiquitous in the general population than PDAs. Recent estimates by the International Telecommunication Union put global penetration of cellular services at around 87%, meaning that close to six billion people own a mobile phone. The developing world is not far behind this global average, with an estimated 4.5 billion people (79%) having a mobile-phone subscription<sup>97</sup>. Finally, PDAs are often operated using a stylus, whereas modern mobile phones use either a QWERTY keyboard or are touch enabled<sup>220</sup>. These differences make mobile-phone-assisted personal interviewing (MPAPI) sufficiently different from PDA-assisted interviews to warrant further study.

The aim of this paper is to compare data accuracy and costs of health research data collection using MPAPI and paper-and-pen (PAP) survey methods. It is hypothesized that both data accuracy and cost-effectiveness will be greater for the MPAPI survey.

## **6.2 Method**

### **6.2.1 Study design and sample selection**

A randomized two-group post-test design was used. Random assignment accounts for potential baseline differences by distributing potential confounders across the two groups. Instead of randomizing interviewees, interviewers were randomly assigned one of two data-collection methods using a computerised random number generator. Interviewers were recruited from two large rural research offices in the province of KwaZulu-Natal, South Africa. These sites were chosen as both employ a large number of individuals who routinely collect health research survey data from community members in rural and semi-urban areas. Study recruitment posters were placed in both offices, and interested staff members were asked to submit their details. Of the 132 people who expressed interest, 100 (75%) were randomly selected and enrolled into the study. Assuming a medium-effect size of 0.5 for Cohen's  $d$ , this number of participants gives an independent samples  $t$ -test an 80% chance of detecting a significant difference between means. Interviewers varied in age, gender and education (see Table 6.1).

## 6.2.2 Procedure

### 6.2.2.1 Group selection

The 100 participants were randomly assigned to either a mobile (MPAPI) (n=50) or pen-and-paper (PAP) (n=50) survey group. The two groups were similar in age and gender but differed with respect to education and years of experience (see Table 6.1).

**Table 6.1. Comparative demographic profiles of interviewers enrolled to collect survey data using either paper and pen or MPAPI**

|                                                                                               | N = 100 |     |     | Paper Assessment n=50 |     |     | Mobile Assessment n=50 |     |     |
|-----------------------------------------------------------------------------------------------|---------|-----|-----|-----------------------|-----|-----|------------------------|-----|-----|
|                                                                                               | %       | M   | SD  | %                     | M   | SD  | %                      | M   | SD  |
| <b>Age</b>                                                                                    |         |     |     |                       |     |     |                        |     |     |
| < 27 years                                                                                    | 34      |     |     | 30                    |     |     | 38                     |     |     |
| 28 to 32 years                                                                                | 27      |     |     | 32                    |     |     | 22                     |     |     |
| 33 to 37 years                                                                                | 27      |     |     | 26                    |     |     | 28                     |     |     |
| 38+ years                                                                                     | 12      |     |     | 12                    |     |     | 12                     |     |     |
| <b>Gender</b>                                                                                 |         |     |     |                       |     |     |                        |     |     |
| Male                                                                                          | 29      |     |     | 22                    |     |     | 36                     |     |     |
| Female                                                                                        | 71      |     |     | 78                    |     |     | 64                     |     |     |
| <b>Education<sup>a</sup></b>                                                                  |         |     |     |                       |     |     |                        |     |     |
| Incomplete Secondary Schooling                                                                | 19      |     |     | 30                    |     |     | 8                      |     |     |
| Completed Secondary Schooling                                                                 | 37      |     |     | 42                    |     |     | 32                     |     |     |
| Completed Tertiary Education                                                                  | 44      |     |     | 28                    |     |     | 60                     |     |     |
| <b>Interviewing Experience (years)<sup>b</sup></b>                                            |         | 1.6 | 1.9 |                       | 1.1 | 1.5 |                        | 2.2 | 2.2 |
| a. Education levels of the two samples were significantly different $\chi^2(2)=12.86, p<0.01$ |         |     |     |                       |     |     |                        |     |     |
| b. Years of experience in the two samples were significantly different $t(98)=2.88, p<0.01$   |         |     |     |                       |     |     |                        |     |     |

#### **6.2.2.2 Interviewer Training**

All 100 interviewers were provided with one hour of basic training in completing the questionnaire. The MPAPI group received an additional two hours of training on how to use the mobile handset and the survey software (Mobenzi Researcher, <http://www.clyral.com>). It is unlikely that this additional training time affected results as it was aimed at providing only a basic level of familiarity with phone navigation and application use.

#### **6.2.2.3 Measures**

A 43-item survey that included a range of question types was assembled from existing questionnaires designed to explore social and health behaviour. The survey included numeric, date, text, time, and single and multiple-selection fields. A paper version of the survey was prepared in Microsoft Word and translated into isiZulu. Both English and isiZulu versions of the survey were programmed onto the mobile phones using the Mobenzi Researcher web-survey design interface.

#### **6.2.2.4 Mobile survey software**

The commercially available software, Mobenzi Researcher, and its associated data management portal were designed by Clyral (2011). Mobenzi Researcher is a Java 2 Micro Edition (J2Me) application that runs on all MIDP v2 compliant handsets. It provides full survey functionality, including the ability to create various question types, mark fields as mandatory and intelligently manage survey branching. Nokia Mobenzi Researcher was installed on Nokia E63 handsets.

#### **6.2.2.5 Data collection**

Each of the 100 interviewers conducted interviews with the same five interviewees for a total of 500 interviews. The five interviewees were extensively trained on the survey and coached to give exactly the same answers to each of the survey questions. Simulating the interaction for both phone and paper groups made it possible to deduce when and how data-capturing errors were introduced into the survey by interviewers. To ensure that the interviewees were giving the agreed-upon answers every time, 25% of the interviews were observed and recorded by a neutral observer. By collecting five interviews for each interviewer it was possible to a) establish within-subject error variability; b) examine practice effects; and c) more accurately assess the number of errors made by each interviewer. After completing an interview, data from the MPAPI group were uploaded to a secure online database using a local mobile operator. The handwritten questionnaires in the PAP group were captured into an MS Access database by a data-entry clerk. Only a single round of data entry was used due to limited resources.

#### **6.2.2.6 Analytic strategy**

A binary scoring system (0 = correct response given by interviewer, 1 = incorrect response recorded) was used to analyse the number of errors made on each questionnaire. These binary scores were totalled to give a final error score for each interviewer, and error scores from each of the five interviews were averaged to give a mean error rate for each interviewer. An independent samples t-test was calculated to compare mean error rates for the two modes. In addition, a 2x5 Factorial ANOVA was computed to compare MPAPI and PAP data-entry errors between the five interviewees. Factorial models were also run that compared MPAPI and PAP across age, gender, education and interview number

(practice effect). An incremental cost-effectiveness analysis was used to compare the costs of these two approaches<sup>221</sup>.

### **6.3 Results**

In total, 381 data errors (missing data (70) and keying errors (311) introduced during electronic data capture of the PAP surveys and response entry onto the mobile phones) were identified in the 500 surveys (21 500 items) resulting in an overall error rate of 1.8% (1.9% for PAP and 1.7% for MPAPI). Two PAP interviews were found to be incomplete but were retained in the analysis. Comparing PAP to MPAPI, no missing fields were recorded in the MPAPI survey, with all 70 missing responses occurring in PAP questionnaires. Of the 311 keying errors, 134 (43.0%) occurred on the PAP forms and 177 (47.0%) on the MPAPI survey. Table 6.2 presents these figures along with comparisons by gender, age, interview order and interviewee.

Both the means and variance were found to be similar for mode ( $F\ 0.511, df\ 1, p = 0.475$ ) and interaction between mode and interviewer ( $F\ 0.914, df\ 4, p = 0.455$ ), but different for interviewee ( $F\ 2.71, df\ 4, p = 0.03$ ).

**Table 6.2. Comparison of data errors for PAP and MPAPI surveys by gender, age, interview order and interviewee**

|                                     | N = 500 interviews (21 500 items) |      |      |                |      |      |
|-------------------------------------|-----------------------------------|------|------|----------------|------|------|
|                                     | PAP                               |      |      | MPAPI          |      |      |
|                                     | n= 10 750 items                   |      |      | n=10 750 items |      |      |
|                                     | Sum                               | M    | SD   | Sum            | M    | SD   |
| <b>Error Type</b>                   |                                   |      |      |                |      |      |
| Missing                             | 70                                | 0.11 | 2.09 | 0              | 0    | 0    |
| Different Value                     | 134                               | 0.54 | 2.21 | 177            | 0.71 | 1.11 |
| All Errors                          | 204                               | 0.82 | 2.13 | 177            | 0.71 | 1.11 |
| <b>Gender (All Errors)</b>          |                                   |      |      |                |      |      |
| Male                                | 50                                | 0.91 | 2.34 | 66             | 0.73 | 1.22 |
| Female                              | 154                               | 0.79 | 2.07 | 111            | 0.70 | 1.05 |
| <b>Age (All Errors)</b>             |                                   |      |      |                |      |      |
| 18-27yrs                            | 75                                | 1.00 | 3.08 | 52             | 0.55 | 0.91 |
| 28-32yrs                            | 68                                | 0.85 | 2.04 | 51             | 0.91 | 1.12 |
| 33+yrs                              | 61                                | 0.64 | 1.00 | 74             | 0.75 | 1.26 |
| <b>Interview Order (All Errors)</b> |                                   |      |      |                |      |      |
| Interview 1                         | 34                                | 0.68 | 1.08 | 38             | 0.76 | 0.22 |
| Interview 2                         | 33                                | 0.66 | 1.06 | 45             | 0.90 | 1.22 |
| Interview 3                         | 21                                | 0.42 | 0.73 | 24             | 0.48 | 0.86 |
| Interview 4                         | 61                                | 1.22 | 3.69 | 30             | 0.60 | 0.95 |
| Interview 5                         | 55                                | 1.10 | 2.48 | 40             | 0.80 | 1.25 |
| <b>Interviewee (All Errors)</b>     |                                   |      |      |                |      |      |
| Interviewee 1                       | 76                                | 1.52 | 4.29 | 49             | 0.98 | 1.32 |
| Interviewee 2                       | 42                                | 0.84 | 1.11 | 25             | 0.50 | 0.84 |
| Interviewee 3                       | 35                                | 0.70 | 0.89 | 34             | 0.68 | 1.15 |
| Interviewee 4                       | 29                                | 0.58 | 1.07 | 36             | 0.72 | 1.34 |
| Interviewee 5                       | 22                                | 0.44 | 0.86 | 33             | 0.66 | 0.77 |

Errors in each of the five surveys were summed and an average calculated for each of the 100 interviewers. No significant difference ( $t = -0.209$ ,  $df = 98$ ,  $p = 0.835$ ) in the number of data errors was found between PAP and MPAPI surveys (Table 6.3), with error rates for PAP calculated as 1.9% (errors per 100 items) and MPAPI as 1.6%.

**Table 6.3. Comparison of the mean number of data-entry errors made by interviewers assigned to either PAP or MPAPI**

| Mode   | Mean | Median | 95% CI Lower | 95% CI Upper | Std. Dev | T     | df | Sig (2-tail) | Mean Diff. |
|--------|------|--------|--------------|--------------|----------|-------|----|--------------|------------|
| Mobile | 3.54 | 3      | 2.48         | 4.60         | 3.72     |       |    |              |            |
| Paper  | 3.72 | 2      | 2.35         | 5.09         | 4.82     |       |    |              |            |
|        |      |        |              |              |          | -0.21 | 98 | 0.84         | -0.18      |

Age, gender, education and practice effect were considered as possible mediators of error rate. Using an alpha of 0.05, none produced significant main or interaction effects with survey mode. Practice effect, as measured by the data-error rates on questionnaires 1 to 5 by the interviewer was not significant ( $F= 1.360, df=4, p=0.247$ ). The interaction between practise effect and survey mode was also statistically not significant ( $F=1.049, df=4, p=0.381$ ). Gender produced similar results with no significant main ( $F=0.405, df=1, p=0.526$ ) or interaction effect ( $F=0.094, df=1, p=0.760$ ). Although not statistically significant ( $F =1.355, df=2, p=0.245$ ), age varied systematically with data-collection mode. Younger participants produced lower than average error rates when collecting data using MPAPI and higher than average rates when using the PAP mode.

The average PAP interview lasted 5.5 minutes ( $sd = 2.6$ ) while MPAPI interviews lasted on average 14.2 minutes ( $sd = 6.47$ ). A 2x5 Factorial ANOVA comparing mode and interview number indicates both main and interaction effects. The interaction of mode and interview order is one where survey duration remains stable across all five interviews for the PAP group (interview one has a mean of 5.5 minutes and interview five 5.58 minutes), while the MPAPI durations fell from interview one (mean = 18.7 minutes) to interview five (12.76 minutes).

**Table 6.4. Effect of survey mode and the number of interviews already completed (experience) on survey duration**

| Source                        | df | Mean Square | F       | p     |
|-------------------------------|----|-------------|---------|-------|
| Survey mode                   | 1  | 9278.855    | 423.361 | <0.01 |
| Interview Order               | 4  | 165.038     | 7.530   | <0.01 |
| Survey Mode * Interview Order | 4  | 172.787     | 7.884   | <0.01 |

Although duration times were not found to vary by age, age was found to be influential when comparing the rate of change in survey duration; the group younger than 27 years improved by 10.1 minutes from survey one to five, 28-32-year-olds improved by 4.5 minutes, and 33 years and older by 3.7 minutes.

It took on average 14.2 minutes from the start of the interview for MPAPI surveys to be available for cleaning in Excel. Paper surveys were processed and captured as part of the regular workload of a data-entry clerk. It took on average two to three days for this data to be available for cleaning in Excel.

### **6.3.1 Cost-effectiveness**

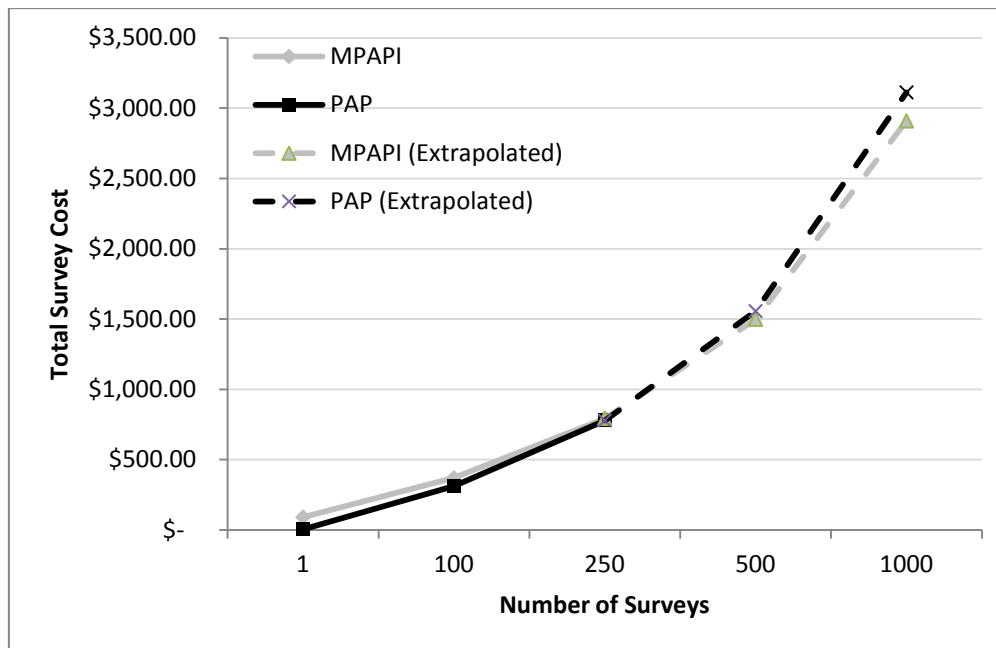
Total costs (including the combined fixed and variable costs) were US\$778.93 for the PAP survey and US\$793.57 for MPAPI (Table 6.5). Fixed costs included the purchase of a mobile-phone handset (US\$88.57) for MPAPI, and pens (US\$1.43) for the PAP surveys. Variable costs for MPAPI included airtime, interviewer salary and survey software costs (\$US2.82 per survey), while the variable costs of PAP included paper, printing, interviewer salary and data-capture clerk's salary, which totalled US\$3.11 per survey. A comparison of

costs and data-capture errors associated with PAP and MPAPI survey modes is shown in Table 6.5.

**Table 6.5. Costs and errors associated with PAP and MPAPI survey modes**

| Mode   | Surveys | Fixed Cost | Variable Cost | Total Cost<br>n=250<br>[C] | Error-free Responses<br>[E] | Cost-Effectiveness Ratio | Incremental cost<br>[ΔC] | Incremental effect<br>[ΔE] |
|--------|---------|------------|---------------|----------------------------|-----------------------------|--------------------------|--------------------------|----------------------------|
| Paper  | 250     | US\$1.43   | US\$3.11      | US\$778.93                 | 10546                       | 0.074                    | US\$778.93               | 10546                      |
| Mobile | 250     | US\$88.57  | US\$2.82      | US\$793.57                 | 10573                       | 0.075                    | US\$14.64                | 27                         |

Both the cost-effectiveness of each mode was calculated as well as a comparison of the two modes through the incremental cost-effectiveness ratio. For the 250 study surveys, each error-free PAP response cost \$0.07, similar to the cost of an error-free MPAPI response (\$0.08). When compared to PAP, an incremental cost-effectiveness ratio (ICER) of 0.54 was calculated for the MPAPI. This indicates that although MPAPI surveys cost US\$14.64 more, the MPAPI survey produced an increased number of error-free responses (27), i.e. each additional error-free response cost US\$0.54. Extrapolating these costs to a larger number of surveys suggests that after 301 surveys, MPAPI becomes more cost-effective than paper (Figure 6.1). This projection should be interpreted with caution given the study limitations mentioned in the following section.



**Figure 6.1. Extrapolation of survey cost and the number of surveys conducted**

#### 6.4 Discussion

We compared data-capturing errors and costs of health research data collected in a rural South African setting using PAP and MPAPI survey methods. Very low error rates of less than 2% were found for both PAP and MPAPI. Errors were, however, found to vary significantly with age. Younger participants made fewer errors in the MPAPI group than older participants, while at the same time making more errors in the PAP group than older participants. Cost-effectiveness was almost identical for the two modes. We found no conclusive evidence to support our hypothesis that MPAPI would improve both data quality and cost-effectiveness in this short survey, small-scale study.

Although limited information exists on MPAPI, electronic survey data collection using PDAs have previously been compared favourably to PAP surveys across a range of outcomes. Both missing values and data-entry errors have been found to be lower for

PDA than paper surveys. Palen et al.<sup>222</sup> report PDA surveys to be missing on average 0.8% (SD= 0.51) of responses, while paper surveys produced an average of 11.5% (SD=6.48) missing responses. By comparison, this study produced no missing responses for MPAPI and 70 (0.7%) for PAP. The ability to mark responses as “required” resulted in no missing values appearing in the MPAPI dataset. Although missing values in the PAP dataset were low, partly due to the interviewees been told to give the same answer to each question, at scale, on longer surveys and under real field conditions the difference might become significant.

Depending on the length of PDA training interviewers received, Patnaik et al.<sup>124</sup> report data-entry error rates of between 0.1% (five days training) to 14% (two to three minutes of training). After an hour of training on the questionnaire and an additional two hours of mobile-phone training, we found error rates of 1.9% for PAP and 1.6% for MPAPI entries. This is similar to the rate of 1.7% found by Missinou et al.<sup>223</sup>, who used PDAs to collect data for clinical trials in Gabon. As noticeable as the similarities in error rate of PAP and MPAPI are, so are the overall low error rates achieved in this study. By hiring well-trained and experienced staff, and by keeping surveys short and clear, very low error rates can be achieved, irrespective of whether PAP or MPAPIs are used for data entry. The low error rates may also partly be explained by the quality-assurance process embedded in this study. A quarter of all surveys took place under the eye of an external supervisor. The presence of the supervisor may have encouraged interviewers to take more care during the interview. These tightly controlled conditions may well have had an impact on the natural variability that may have arisen under a more typical field setting. While PAP produced slightly fewer errors on the simple questionnaire used in the current study, as

questionnaires become more complicated and include more skip patterns and interviewer-decision points, the risk of error increases for the PAP mode.

Using a crossover design, Yu et al.<sup>120</sup> report that 62% of participants perceived a PDA-based survey to be quicker to complete than a paper survey. Data entry, validation and cleaning took a total of 1.5 hours for the 120 PDA surveys and 20.5 hours for the 120 paper surveys. In the current study, a similarly dramatic decrease in the time from interview to data availability was achieved for the MPAPI group. Data was available in minutes using MPAPI, while PAP surveys, captured by a data clerk employed at one of the field sites as part of their daily workload, took two to three days to return the data in Excel. This time difference would become more pronounced as the number of surveys to be captured increases. Contrary to Yu et al.<sup>120</sup>, we did not find the MPAPI interview (14.2 minutes) to be shorter than the PAP interview (5.5 minutes). A practice effect was, however, observed for mobile participants, with Interview 1 taking 18.7 minutes and Interview 5 lasting 12.7 minutes. Interview length for the five paper interviews remained similar from Interview 1 (5.5 min) to Interview 5 (5.8 min). The effect was even more dramatic for the younger interviewers, who cut their interview length by 50% by the fifth interview. Despite the rapid improvement, their times were still double the average length for paper-based survey.

Using commercial software (Mobenzi Researcher) and entry-level mobile phones, cost per survey was very similar for MPAPI and PAP questionnaires. Calculating the cost for each error-free response yielded values of \$0.07 for paper surveys and \$0.08 for mobile-phone surveys. PAP was found to be more cost-effective than MPAPI when less than 301 surveys

needed to be completed. As the number of surveys rise above this cut-off, MPAPI becomes the more cost-effective approach. When comparing the costs of paper with web-enhanced handheld computers, Weber et al.<sup>71</sup> concluded that electronic data collection was more cost-effective than paper surveys. A possible explanation for the difference between Weber et al.'s<sup>71</sup> finding of reduced costs and this study's finding of comparable costs between paper and electronic data entry could be due to the different software billing model employed in the two studies. In this study costs were levied by the software developer on a per-item basis and therefore included as a variable cost. Weber et al.<sup>71</sup> were able to purchase their software outright at the start of the study, making it a fixed cost. For mobile data entry to be cost-effective, it would seem more prudent to either purchase the survey software, increasing the upfront investment, or explore the cost associated with setting up one of the available open-source solutions (e.g. Open Data Kit or EpiSurveyor).

In comparing these two methods, consideration should also be given to the experience of the Interviewer. Although the interview portion of MPAPI interviews initially took longer to complete than PAP, the efficiency of the entire interview process was improved for interviewees in the MPAPI group. These efficiencies were gained primarily from the ease with which interviewers could select the appropriately translated version of the survey which was pre-loaded onto the phone. Anecdotal evidence suggests that the MPAPI surveys were less intimidating for participants, as a mobile phone is a much more common object than the formal-looking clipboard with paper form.

Careful consideration needs to be given to the training and practice time required by interviewers new to MPAPI. The two hours of training time participants had to familiarise themselves with the mobile-phone surveys may not have been sufficient. Interviewers using MPAPI took longer per survey than PAP interviewers. As most of the items were selection or short-answer questions, this cannot be put down to typing speed on the mobile phone. Instead, the challenge lay with navigating through the survey and following the correct procedure required to enter text, numbers and dates. Younger participants, with potentially greater exposure to and daily use of mobile phones, appear to have come to grips with these challenges more rapidly and were able to concentrate more carefully on the actual responses they were entering. This finding should not lead to the conclusion that only young field workers should be hired. Rather, care should be given during training and piloting to ensure that all field staff have been given adequate time to familiarise themselves with the mobile phone.

The study had a number of limitations that may limit its real-world applicability. A small sample size, short interview length and a tightly controlled study design may have had the unintended side effect of limiting data variability. This tightly controlled study design may be partly responsible for the small differences found between the two methods. Using a full-length survey with more participants under standard field conditions may produce results similar to studies that have found electronic data collection to be superior to PAP. A second limitation was that PAP data were not double-entered. This may have reduced the number of errors introduced during data capture. A strength of this design was its focus on interviewers rather than the more common approach of studying interviewees. Future research should focus on a comparison of these two methods at scale.

## **6.5 Conclusion**

Across all measured outcomes, none suggest a clear advantage to either survey mode. Yet implementation experience suggests that in some instances MPAPI holds significant advantages over PAP. The following recommendations may assist when deciding between implementing PAP or MPAPI surveys. Short, once-off surveys favour PAP over MPAPI. Large, complex studies lasting several years across multiple sites with poor infrastructure would likely benefit from MPAPI surveys. When start-up time is limited, paper surveys may be preferred to mobile surveys due to the increased time required to train on MPAPI. Finally, if time-sensitive data are required either for management or surveillance, MPAPI will come into its own. If data are not urgently needed for analysis, a paper system may cost less and be easier to set up and run. Both PAP and MPAPI have strengths and weaknesses. Consideration needs to be given to ensuring that project requirements are carefully matched to these factors.

## Chapter 7

---

**Point of care in your pocket:**

**A research agenda for the field of mHealth**

Accepted by *WHO Bulletin*.

Mobile health, or mHealth<sup>133</sup> has been defined by the Global Observatory for eHealth (GOe) as *medical and public health practice supported by mobile devices, such as mobile phones, patient-monitoring devices, personal digital assistants (PDAs), and other wireless devices*<sup>133 (p14)</sup>. With global penetration of mobile services topping 87%, the opportunity for connecting with six billion people has brought great enthusiasm, energy and media attention to the field<sup>97</sup>. This momentum has seen evidence for the health benefits of mHealth being built not by traditional large-scale and time-intensive randomised control trials, but instead by the proliferation of small pilot projects, particularly in low and middle-income countries<sup>128</sup>. Rapid technological innovation has led to an enthusiastic proliferation of untested methods. Many of these, however, have been small-scale with little policy co-ordination<sup>131</sup>. To reduce wasted time, effort and money, we propose seven key research agenda items required to maximise the potential of mHealth through evidence for a programme's utility and its cost-effectiveness.

### 7.1 mHealth Needs to Develop an Evidence Base

A frontier mindset fosters real and rapid innovation in mHealth. This fast iterative approach should not be lost, especially given how quickly mobile and wireless technologies evolve, but a certain amount of caution is warranted. Although well on target to attaining its goal of reaching one million mothers by 2012, text4baby, a free mobile text service in the United States that provides free health information to pregnant women and new mothers, is a case in point<sup>224</sup>. Launched in February 2010, text4baby has gained wide media attention, winning awards from the United States Office of Science

and Technology in 2010 and the Public Relations Society of America (PRSA) in 2011. Yet formal results are still to be made available from a belated evaluation that was recently completed<sup>225</sup>. Whether the 3.5 million messages sent in 2010 were read, acted upon or in any way able to change behaviour is not known. Only 12% of countries surveyed in the 2011 World Health Organization's Global Observatory for eHealth reported evaluating their mHealth services<sup>133</sup>. Formal and rigorous cost-effectiveness evaluation of mHealth programmes need to be undertaken to ensure that time and money are not wasted on ineffective programmes, and that best practices are disseminated from successful programmes.

## **7.2 mHealth Systems Should Be Interoperable with Existing eHealth Initiatives**

Estrin has argued that mHealth is emerging as a patchwork of incompatible applications<sup>131</sup>. IT and health-information systems are often based on technologies that cannot communicate and share data with one another and have traditionally been characterised by a silo approach<sup>131</sup>. In much the same way, this issue of interoperability parallels the public-health experience of vertical disease-oriented programmes<sup>226</sup>. There is a compelling argument that better outcomes can be achieved by tackling disease through an integrated approach that strengthens health systems<sup>227</sup>. Enabling mHealth systems to share information with one another as well as with broader eHealth systems will increase efficiency, reduce costs associated with the collection of that data and result in better care for patients. In addition, the benefits of such a system may in fact be felt in the broader care system<sup>131</sup>. To achieve this will require strong cooperation, meaningful collaboration and adherence to international standards.

### **7.3 mHealth Should Adopt and Implement the Same Standards Already Present in eHealth**

There are a number of eHealth standards in common use today. These standards were designed to facilitate interoperability of health data between health devices and IT systems. In order for different systems to share data, they must contain comparable information (data elements), employ comparable words (terminology) and use an agreed upon way to communicate (messaging).

mHealth projects such as cellScope<sup>228</sup>, which enables light microscopy to be performed on mobile phones, make mention of the Digital Imaging and Communications in Medicine (DICOM) standard for the exchange of medical images. By upholding and implementing the DICOM standard they provide an excellent example of how mHealth can build on currently existing eHealth standards. Most projects currently do not use this standard. Similarly they do not use the HL7 Clinical Document Architecture (CDA) for the exchange of electronic clinical documents (such as discharge summaries and progress notes). Adherence to privacy, security and confidentiality standards are also critical to the ongoing success of any mHealth intervention. If mHealth and eHealth systems are interoperable and share common standards, integration costs could be better controlled.

### **7.4 mHealth Should Take a Participatory Approach**

Contemporary thinking is that health care should be personalised, predictive, preventive and participatory<sup>12</sup>. mHealth is well suited to enable this participation by providing a simple means by which patients can engage in their own health care. Programmes need to go further than simply connecting patients with health care providers. Many current

mHealth initiatives focus on outdated, unidirectional models of patient communication, for example by exclusively collecting data, providing information or sending reminders. With mHealth technology, patients can easily transmit data and commentary on their health status and health care to their electronic health records, for example, allowing for an ongoing and active health dialogue<sup>229</sup>. Research that evaluates the utility and cost-effectiveness of the participatory care approach should be prioritised.

### **7.5 mHealth Should Promote Equity in Health**

In high-income countries, mHealth programmes have tended to focus largely on supporting ageing populations through delivering prevention programmes and reducing health-care costs through home-based monitoring<sup>230</sup>. In low and middle-income countries, by contrast, mHealth has been conceptualised much more broadly, as a way to promote access to health services for the socioeconomically and geographically disadvantaged. Irrespective of country, all mHealth programmes should not unfairly discriminate against, marginalise or exclude those at the fringes of society. The very high mobile penetration reach of 79%<sup>97</sup> in low and middle-income countries may nevertheless obscure patterns of exclusion. The authors of this brief have recent data from HIV-positive women in South Africa that suggests that the 20% who do not have mobile phones are significantly poorer, live in more remote areas and are older than those who have access to a mobile phone. Consolidated evidence for how best programmes can cost-effectively support and provide care to this marginalised group is needed.

## **7.6 mHealth Programmes Need to Move to a Situation of Sustainability**

The most common mHealth initiatives amongst WHO member states in 2011 are health call centres (59%), emergency toll-free telephone services (55%), managing emergencies and disasters (54%), and mobile telemedicine (49%)<sup>133</sup>. Unfortunately many of these programmes are small, unsustainable pilot programmes. Although a little dated, data from a 2008 review of mHealth literature revealed that 84% of the published programmes were prototypes, pilots or tests<sup>231</sup>. The figure has likely improved over the past four years but nonetheless points to an urgent need to find sustainable and cost-effective models for supporting and developing early exploratory mHealth ideas and for integrating them with local healthcare financing systems.

## **7.7 mHealth Needs to Focus on Health, Not on the Technology**

Debate exists as to whether mHealth services should be designed to have the best coverage by using the lowest common denominator technology, for example text messages, or benefit from the advantages that can be gained by targeting the sensors and processing power of the latest generation smartphone. This discussion may be practical but slightly misleading. Technology changes so swiftly that spending undue amounts of time, energy and effort in this domain may be a waste. Behaviour change is driven by people, not technology. Technology can support, enable and trigger change but it is only through the will and skill of dedicated people that solutions can be found to the health problems we wish to address. A strong collaborative evidence base that is coordinated at an inter-country level by key stakeholders in the mHealth community would reduce wasteful expenditure on programmes that lack strong evidence and promote best practice models.

mHealth technology is exciting and attractive, but it is in the establishment of country-level best practices that are both cost-effective and supported by rigorous research and evaluation that the real challenges lie. Policy makers and funders need to promote, legislate and fund work that integrates and builds upon a common mHealth framework. By fostering a synchronised, strategic global research agenda that is focused more heavily on evaluating evidence related to the impact and cost-effectiveness of mHealth services, the WHO could facilitate the ability of programmes to address the real-world complexity not only of providing access, but, more crucially, of contributing to high-quality health care for all.

---

## **Part 3: Discussion**

---

## Chapter 8

---

### Discussion

This chapter presents a critical discussion of the three studies undertaken in this thesis. The chapter begins by first presenting a consolidated overview of the findings. This is followed by a discussion of two emergent themes that cut across all three studies. The theoretical relevance of the research is then considered and a revised theoretical model proposed. The significance of the findings to the South African context, and the field of survey research more generally, are then elaborated upon. Finally, study limitations and future research directions are discussed and concluding remarks offered.

### 8.1 Consolidated Findings

The aim of this thesis was to explore three research questions pertaining to mobile phones and survey data collection. These objectives were to establish the effect mobile-phone-assisted personal interviewing (MPAPI) had on data quality and cost; to explore the acceptability of MPAPI methods in health research with vulnerable populations; and to explore the feasibility of mobile-phone-assisted self-interviewing (MPASI) in a longitudinal birth cohort. These objectives were met through a series of empirical studies presented in chapters 3, 4, 5, 6 and a comment on mHealth in chapter 7. The central findings linked to each research objective are summarised in Table 8.1. These findings are reorganised diagrammatically in Figure 8.1. This figure outlines three types of mobile-phone-assisted data collection. Two of these, MPAPI and MPASI, are explored in relation to De Leeuw's conceptual model of how data collection mode affects data quality. The third type, mobile-phone-assisted remote sensing (MPARS), was not evaluated as part of this thesis. To recap, De Leeuw defines three characteristics that influence data collection

and differ by mode. The first, media-related factors, accounts for the familiarity of and conventions surrounding different communication channels and media used to collect data. Rather than social conventions associated with the communication channel used, information transmission focuses on the impact of specific technical features of the mode in question. For example, some media allow the participant to control the order in which items are presented, and in others the interviewer controls this process. The third and final component relates specifically to the presence or absence of an interviewer.

**Table 8.1. Summary of thesis findings**

| Objective                                                                                                                                        |                                                                                                                                                                     | Chapter | Thesis Findings                                                                                                                                                                                      |
|--------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1) <i>How does use of mobile phones for data collection impact on the costs, quality and time-to-availability of data?</i>                       |                                                                                                                                                                     |         |                                                                                                                                                                                                      |
| I                                                                                                                                                | Compare the accuracy of mobile-phone data entry and upload vs. the administration of a paper-based questionnaire and subsequent data capture on a desktop computer. | Six     | <ul style="list-style-type: none"> <li>Compared to PAPI, accuracy was not found to be improved through the use of MPAPI.</li> <li>MPAPI reduced administrative load compared to PAPI.</li> </ul>     |
| II                                                                                                                                               | Compare the time taken to move from survey to actionable data across the domains of mobile and paper-based surveys.                                                 | Six     | <ul style="list-style-type: none"> <li>MPAPI surveys took longer to complete than PAPI.</li> <li>Overall time to availability of data for analysis was reduced considerably using MPAPI.</li> </ul>  |
| III                                                                                                                                              | Compare the cost-effectiveness of mobile-phone vs. paper-based survey data collection at a research field site in KwaZulu-Natal, South Africa.                      | Six     | <ul style="list-style-type: none"> <li>MPAPI costs are associated with survey size.</li> <li>For small surveys MPAPI is equivalent in cost to PAPI.</li> <li>Larger surveys favour MPAPI.</li> </ul> |
| 2) <i>What perceptions do vulnerable communities and field workers in rural South Africa hold about mobile phones as a data-collection tool?</i> |                                                                                                                                                                     |         |                                                                                                                                                                                                      |
| IV                                                                                                                                               | Describe the perceptions of rural HIV-positive pregnant women regarding the capture of personal health information on a mobile phone.                               | Five    | <ul style="list-style-type: none"> <li>Relationship with interviewer positively mediated participant perceptions.</li> </ul>                                                                         |
| V                                                                                                                                                | Assess the perceptions of South African field workers towards the acceptability of mobile phones as a tool for data capture.                                        | Five    | <ul style="list-style-type: none"> <li>Interviewers perceived MPAPI as both easy to use and useful, two factors known to influence the acceptance of new technologies.</li> </ul>                    |

| <i>3) Can mobile phones be used by adolescents to complete self-report surveys?</i> |                                                                                                                     |              |                                                                                                                                                                                                                 |
|-------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|--------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| VI                                                                                  | What is the response rate of a mobile-phone survey completed by adolescents?                                        | Three & Four | <ul style="list-style-type: none"> <li>• Low response rates were obtained – 36.5% and 30.9% respectively.</li> </ul>                                                                                            |
| VII                                                                                 | What are the response characteristics (e.g. item response rates) of a mobile-phone survey completed by adolescents? | Three & Four | <ul style="list-style-type: none"> <li>• Survey break-off rates were low.</li> <li>• Free-text answers were short while multiple-choice questions elicited a better response than numeric questions.</li> </ul> |

| Mobile-phone-assisted Data Collection |                                                                                                                                                                                                                     |                                                                                                                                                                                                                 |                                                |
|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------|
|                                       | MPAPI<br>MP-assisted personal interviewing                                                                                                                                                                          | MPASI<br>MP-assisted self-interviewing                                                                                                                                                                          | MPARS<br>MP-assisted remote sensing            |
| Media-related Factors                 | <ul style="list-style-type: none"> <li>• Familiarity with MPs high</li> <li>• Respondent privacy concerns</li> <li>• Differential interviewer comfort and perceived ease of use</li> </ul>                          | <ul style="list-style-type: none"> <li>• Handset turnover high</li> <li>• Real-time responses</li> <li>• SMS more viable than app for communication</li> </ul>                                                  | <p>Not investigated as part of this thesis</p> |
| Information Transmission              | <ul style="list-style-type: none"> <li>• Network coverage</li> <li>• Unidirectional vs. bidirectional communication</li> </ul>                                                                                      | <ul style="list-style-type: none"> <li>• Need a system for fielding queries</li> <li>• Large administrative burden</li> </ul>                                                                                   |                                                |
| Interviewer Impact                    | <ul style="list-style-type: none"> <li>• Interviewer mediates initial acceptability</li> <li>• Interviewer characteristics do not influence data accuracy</li> </ul>                                                | <ul style="list-style-type: none"> <li>• Presence of others</li> </ul>                                                                                                                                          |                                                |
| Data Quality                          | <ul style="list-style-type: none"> <li>• Improved timeliness</li> <li>• Low accuracy improvement over PAPI</li> <li>• Equivalent costs to PAPI in small sample surveys</li> <li>• Improved accessibility</li> </ul> | <ul style="list-style-type: none"> <li>• Moderate survey response rate</li> <li>• Low item non-response</li> <li>• Low survey break-off rate</li> <li>• Potential privacy and data security concerns</li> </ul> |                                                |

**Figure 8.1. Thesis findings: The impact of mobile-phone-assisted data collection on data quality within De Leeuw's conceptual model of how data-collection mode affects data quality**

### **8.1.1 Media-related factors**

In De Leeuw's conceptual model<sup>72</sup>, media-related factors refer to the social conventions and customs associated with the 'normal' use of the mode. These factors also cover the 'locus of control' introduced by the mode. In self-administered modes the respondent is in control of when, where and how fast the questionnaire is completed. In interviewer-led surveys the interviewer has more control over these factors. In terms of media-related factors, MPAPI and MPASI interviewing were found to differ on several counts. A brief summary of these findings is presented in Table 8.2.

#### **8.1.1.1 Mobile-phone-assisted personal interviewing (MPAPI)**

Mobile-phone familiarity and ownership were found to be high in the low-resource rural health context in which most of this work was completed. Overall mobile-phone ownership amongst the 512 women interviewed as part of the study presented in Chapter 5 was 83.9% and average length of ownership was 3.3 years. Despite the women being very familiar with their phones, and the phones themselves having adequate functionality, the current social patterns of use are limited in terms of Internet (5.8%), email (2.8%) and mobile-phone applications (9.5%), while entertainment through music (42.6%), photography (41.9%) and gaming (46.5%) were more common activities. This led to some initial concerns from participants about the use of phones for the collection of sensitive data. The common use of phones for photography may be partially responsible for concerns women raised about the possibility of being photographed by the phone in the process of or as part of data collection.

Interviewers were also found to be highly familiar with mobile phones. Only one of the sixteen interviewers trained to collect data using MPAPI did not own a phone. Although data was not collected on how interviewers used their phone, practical experience gained in the training of MPAPI interviewers over the past five years has shown that people have very different levels of comfort when learning the concepts and practice of mobile-phone interviewing. One possible explanation, as suggested by the findings in Chapter 5, is that ownership may instil an initial pre-training sense of confidence about how easy the mobile-phone surveys will be to use. However, due to the frequency of Internet access and application use being low, training does not proceed as smoothly for interviewers as they anticipate, as they lack skills in the above areas. This finding is reflected in the still high, but significantly lower post-training ease of use scores.

#### **8.1.1.1 Mobile-phone-assisted self-interviewing (MPASI)**

In Chapter 3, 69% of the Birth to Twenty (BT20) cohort, who were later to be enrolled as part of the self-complete survey study, reported owning a mobile phone. Although not as high as the figure found for the HIV-positive women in Chapter 5, the survey was completed two years earlier. As noted in the literature review, mobile-phone ownership rates have increased rapidly in South Africa from an estimated 49.4% in 2006 to 79.5% in 2011. Media-related usage norms were also found to underpin differences in mode preference in the pilot study presented in Chapter 3. Those respondents who used applications such as Mxit (a popular mobile instant-messaging application) or who knew of and/or used Skype (a voice-over-IP application), reported a statistically significant preference for future surveys to be conducted through mobile phone over a conventional face-to-face interview. Although only limited literature exists to support this view,

anecdotal evidence suggests that young adults in South Africa 1) often own more than one SIM card, 2) use these cards across multiple devices, and 3) have a high rate of handset turnover<sup>188</sup>. If accurate, this view may suggest a partial answer as to why the study experienced the marked drop in “contactable” handsets described in Chapter 4. High handset turnover rendered the device-specific software impotent. A better approach in such an environment would be to use a communication channel that can easily follow a participant from one handset to the next, with SMS being an obvious first choice. The final media-related factor that may influence MPASI surveys is the social convention of using phones on the move. This means that the researcher has little control over the environment in which MPASI interviews are likely to take place. Although not reported in the paper presented in Chapter 4, participants were asked to describe where they were immediately before they completed the survey. Typical answers included *watching TV*, *catching a taxi* and *returning from a nightclub*. Survey and participant mobility mean that control of environment is reduced when compared with traditional self-complete surveys. Paradoxically, however, this challenge is simultaneously one of the potential advantages of MPASI research. Being able to deliver surveys directly to participants as they go about their daily lives means that the ecological validity of the data collected can be increased. Questions such as *In the last six months how many times have you...* can be replaced with randomly repeated questions continuously asking *In the last few hours have you....* To achieve this level of monitoring, current studies rely on participants’ retrospective reporting and memories to complete self-report diaries. The drawbacks of diary-based studies are that participants need to remember to take the diary around with them and then to fill it in at the appropriate time. When mobility of participants is increased, poor compliance in diary studies has been noted<sup>232</sup>.

### **8.1.2 Information transmission**

In De Leeuw's model, information transmission refers to the process by which the cognitive stimulus is received by the respondent and is influenced by the 'communication channel' used. Although difficult to fully separate out, only the technical elements of the communication channel are considered and those facets more concerned with social custom are excluded. Major factors influencing information transmission include the way in which stimuli are presented and communication flow is regulated. A brief summary of these findings is presented in Table 8.2.

#### **8.1.2.1 Mobile-phone-assisted personal interviewing (MPAPI)**

Two technical factors limiting the usefulness of MPAPI have been identified in this research. The first, network coverage, needs to be given sufficient consideration in low-resource settings where mobile network coverage may be low. With 90% of South Africa having mobile reception, this was not a challenge faced by any of the studies conducted; however, through the author's work in the field, this has been found to be a challenge in countries such as Uganda. Without mobile coverage, one of the major benefits of MPAPI, time-to-availability of the data, is severely curtailed. Limited network coverage would also make SMS and online MPAPI surveys impossible to conduct. Depending on study design, an MPAPI survey could be conducted using mobile-phone survey software in a context of limited mobile network coverage. The design would need to factor in a way to upload surveys from the phone either to a laptop computer or over a wireless network to the Internet. A second limitation of reduced network coverage is the inability to make use of the bi-directional communication channel the mobile phone offers.

While most MPAPI surveys require only unidirectional uploading of survey results, more advanced MPAPI designs can utilise algorithms stored in an online server to make decisions about which surveys need to be completed in a particular context. A real-world practical example might be that of a health worker who uses an MPAPI survey to collect health information on participants in a household. On submitting the information provided by one of the household members, the health worker may find the survey software asking her to complete a specialised drug-resistant-tuberculosis questionnaire for this participant. This type of bi-directional survey support system can only be implemented in an environment with consistent and widespread mobile coverage.

#### **8.1.2.2 Mobile-phone-assisted self-interviewing (MPASI)**

As with MPAPI, two technical factors were encountered during implementation of the MPASI survey, presented in Chapter 4. Mobile phones are a bi-directional communication device; in other words, people send and receive communications through their device. In Chapter 4 it is reported that in reply to a text message from the researchers informing respondents of the availability of a new survey to download, text messages were returned by participants. Future MPASI studies should ensure that they have systems in place to deal with the two-way flow of information through mobile-phone communication channels. It is important to consider carefully each of the communication channels available to mobile-phone users when implementing a “participant support system”. Depending on the sample frame, email, social network posts and instant-messaging services may all be mobile communication channels through which participants can be engaged.

The second MPASI survey consideration is mobile-phone management. Without strong administrative procedures, handset turnover was found to be a major challenge. Of the 435 handsets confirmed to be in the respondents' possession and available for survey submissions, one third were no longer contactable a few months later. This finding could have been influenced by the fact that the handsets were donated and so may have been treated differently from the respondent's private mobile phone. Nonetheless, a significant challenge for MPASI survey researchers in low-resource-settings is mobile-phone turnover. In a study completed in KwaZulu-Natal, South Africa, by Crankshaw et al.<sup>208</sup>, 68% of respondents reported either having a phone stolen or damaged. Further research is needed to best describe techniques for maintaining an operational communication channel between researcher and respondent.

### **8.1.3 Interviewer impact**

Different modes of data collection impact both positively and negatively on the role the interviewer can play during a survey. The third component of De Leeuw's model deals specifically with the interaction between mode and interviewer. A brief summary of these findings are presented in Table 8.2.

#### **8.1.3.1 Mobile-phone-assisted personal interviewing (MPAPI)**

In the face-to-face MPAPI survey discussed in Chapter 5, the skill and personal warmth of the interviewer were found to mediate any concerns participants may have felt towards the MPAPI interview. Respondents reported feeling that the mobile phone was not a distraction either to themselves or to the interviewer. In Chapter 6 it was found that the age, gender and education of the interviewer had no significant impact on the accuracy of

the data collected when compared to the keying errors these individuals made on PAP surveys. These findings must be interpreted cautiously and in the context of a study design that was very controlled. The design may have been responsible for reducing and masking interviewer variability, making the discovery of any difference unlikely.

#### **8.1.3.2 Mobile-phone-assisted self-interviewing (MPASI)**

As MPASI is a self-complete survey, the potential for an interviewer to elicit socially desirable responses is removed. What is not clear is whether surveys are being completed in private, in the presence of others or with the help of others. For MPASI survey one, reported on in Chapter 4, on average the respondents reported being in the immediate presence of five other people. This idea will be elaborated on as an area for possible further research.

#### **8.1.4 Data quality**

Media-related factors, information transmission and interviewer impact have a combined impact on the survey mode as well as data quality. It is, however, difficult to predict exactly how each component in the model might interface with other components and how these interactions influence data quality. De Leeuw suggests supplementing the theoretical expectations of the model with empirical data. Chapters 4 and 6 provide some evidence for how the factors discussed above affect the data quality of MPAPI and MPASI surveys.

##### **8.1.4.1 Mobile-phone-assisted personal interviewing (MPAPI)**

The current literature suggests that electronic surveys improve data quality. This general pattern was not replicated by the mode-comparison study presented in Chapter 6. MPAPI

provided few clear examples of improved data quality over face-to-face PAPI. Data accuracy, as measured by keying errors, was found to be high for both modes. Only 381 data-entry errors were identified in 21 500 survey items. PAPI accounted for 204 of these errors and MPAPI the remainder (177). This gives a statistically insignificant difference of 0.3% between the error rate of PAPI (1.9%) and MPAPI (1.6%). An age effect was, however, noted with younger interviewers making fewer errors on the MPAPI survey as compared to PAPI. No cost advantage was conferred on either method following the cost-effectiveness analysis. Over time, the upfront hardware costs and ongoing purchase of survey credits from the software provider equalised with the monthly salary bill associated with the data capture required to process paper surveys. As noted by Groves, data quality and cost go hand in hand. One cannot look at cost without considering whether the increase, or decrease, in cost is associated with a rise or fall in data accuracy. It was found that each error-free PAPI response cost US\$0.07. Each MPAPI error-free response cost US\$0.08. Evaluating the incremental advantage MPAPI holds over PAPI, produced an incremental cost-effectiveness ratio (ICER) of 0.54. This indicates that although MPAPI produced more error-free responses, each additional error-free response costs US\$0.54. One possible explanation for the worse than expected data quality improvements is that the scale of the data collection effort may mediate the effect survey mode has on data quality. This idea is explored further in the section on emerging research themes.

Two data-quality dimensions that were found to vary by mode were timeliness and accessibility. Although MPAPI surveys took significantly longer to complete than PAPI surveys (cf. 14.2 minutes and 5.5 minutes), they were immediately available online and in

a format useful for data analysis. Within a few minutes of completing the interview, data could be analysed by researchers working in a different country. The same advantage was not evident with PAPI. Data capturers were slowed and delayed by the influx of paper surveys and other tasks to which they were assigned. This meant that it often took a day or two for captured surveys to become available. Time was, however, made up with PAPI requiring shorter interviewer-training time than MPAPI.

One conclusion that can be drawn from this is that MPAPI places the administrative and financial burden at the beginning of the research project while PAPI shifts the costs and administrative burden to the end. When using MPAPI, handsets need to be purchased, surveys designed and loaded onto the electronic system, and extensive training offered, all before the survey research can begin. The advantages are that, once in place, MPAPI surveys scale well. That is, MPAPI survey volume can be significantly increased without much further training or infrastructure being required. Data are available quickly, interviewers increase their proficiency over time and the administrative burden of sharing data with off-site colleagues is greatly reduced.

#### **8.1.4.2 Mobile-phone-assisted self-interviewing (MPASI)**

The unexpected data-quality feature of the MPASI research presented in Chapter 4 was the low response rate. Only 105 (of 435) respondents submitted a response to survey one. This figure dropped to 85 for the second survey. It must, however, be noted that only 288 (271 in survey two) of the 435 respondents received the survey notification SMS. This means that one out of three respondents who knew about the survey managed to navigate the complex process of submitting a response. It is not known how many of the

respondents who received the SMS attempted unsuccessfully to submit a response. A key area for further research is establishing if response rates can be mediated and hence improved by changing the mobile communication channel used. Of the responses received, item-level non-response was low. It is not clear if this high level of completeness is influenced by the BT20 participants being familiar with surveys.

A further data-quality issue facing MPASI is ensuring the privacy and security of both survey and response. SMS is an unsecure communication medium and, if used for MPASI, may allow participants' confidential answers to be revealed to anyone with physical access to the mobile phone. Privacy may be further compromised by the sharing of mobile phones.

Having outlined the general thesis findings and discussed them in relation to De Leeuw's model, the following key findings emerged:

**Table 8.2. Table of key findings**

|                                 | <b>Comparative findings of MPAPI and MPASI</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                  |
|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Media-related Factors</b>    | The high familiarity people have with mobile phones makes their presence in both data-collection contexts acceptable. Differences exist in how the mobile-phone medium impacts on personal (face-to-face) and self-complete interviews. MPAPI interviews were generally considered acceptable, easy to use and useful. Mobile-phone familiarity introduced challenges for MPASI with young adults as their “personalisation” of the handset meant many handsets could no longer submit surveys. |
| <b>Information Transmission</b> | Network coverage was found to be extensive enough to allow the successful transmission of data in a number of rural South African locations. Survey management for self-complete interviews was found to be logistically challenging. Both MPAPI and MPASI found strong support for the bi-directional nature of information transmission on mobile phones.                                                                                                                                     |
| <b>Interviewer Impact</b>       | Interviewer characteristics were not found to influence mobile-phone-assisted personal interviews, although as noted this may relate as much to the study design as the validity of the result. Self-complete mobile-phone interviews were often found to be undertaken in the presence of others.                                                                                                                                                                                              |
| <b>Data Quality</b>             | Moderate data-quality improvements were found for both personal and self-complete mobile-phone-assisted surveys. Some elements of data quality, such as accessibility, were found to be greatly improved over pen-and-paper. MPASI surveys had low to moderate response rates and low levels of item non-response and survey break-offs.                                                                                                                                                        |

The research findings of this thesis have provided evidence to reject the null hypothesis that it would not be feasible to implement mobile-phone-assisted personal interviewing to collect health research data in resource-limited settings, of relevance to South Africa. Instead, MPAPI was found to be both feasible and acceptable in a number of situations. Evidence to reject the null hypothesis that data quality would be improved in mobile-phone-assisted interviews is less strong. Although elements of data quality, such as

accessibility and timeliness, were improved, survey error and cost were found to be equivalent in short surveys.

## **8.2 Emerging Research Themes**

In 2005 Couper<sup>62</sup> identified four technology trends in survey data collection, namely: interviewer administration to self-administration; audio-visual advancements; mobility; and continuous measurement. This thesis has focused primarily on the impact of mobility on survey data collection. Findings from field research are reported on in the empirical chapters and a summary of these findings has been presented in the previous section. By making surveys mobile, study scale and survey context have emerged as key themes and will each now be considered in more detail.

### **8.2.1 The influence of scale**

The use of pervasive devices in the collection of survey data needs to be considered with reference to project scope and scale. Interview-led electronic surveys conducted on portable devices hold the most promise in large, multi-site, multi-year collaborative research ventures. It is in these environments that the upfront costs in time, financial and human resources, and system design have an opportunity to pay back dividends through support of project management operations. Although still feasible, health research conducted on a smaller scale where there are fewer respondents, shorter time-frames and where the respondents are geographically concentrated, do not experience fully the benefits associated with mobile-phone-assisted personal interviewing. Some of the advantages offered by mobile phones to large-scale project management include the tracking of interviewer productivity trends, reducing data falsification by recording

geolocation coordinates of interviewers at the time of the survey, rapidly identifying problematic survey items through real-time review of survey data, and improved ability to share data with other investigators in a cooperative research venture.

The advantages of scale noted for MPAPI were in contrast to the challenges associated with scale factors for MPASI. Project management complexity of MPASI surveys seem positively correlated to sample size. Hüsler<sup>156</sup>, reporting on the use of SMS reminders in the treatment of tuberculosis in Cape Town, states the challenge as follows:

*Project management issues are so inherently intertwined with the technology that it is difficult to separate them. Project implementation clearly limited the effectiveness of the Compliance Service, but it is not a reflection on the usefulness of the technology itself. To the contrary, the Service has potential to provide more choice in the care of TB and greater convenience for the patient. However, the problems encountered underline the limitations of the Service and imply there are important conditions for its success. (p 3)*

Once scaled up, MPASI surveys appear more vulnerable to technological challenges than MPAPI surveys. Although not discussed as part of this thesis, the future collection of data using mobile-phone-assisted remote sensing (the automatic collection of user behaviour data through the sensors built into the next generation of pervasive computing devices) will bring its own set of challenges related to project size. Through the continuous monitoring of sensor data relating to the respondents' health and behaviour, new techniques will need to be designed to manage the flood of data that will need to be aggregated and analysed.

Together, the strengths and challenges associated with project scale and the collection of data using pervasive computing devices, make scale a significant dimension in the measurement of survey data quality.

### **8.2.2 The influence of context**

Establishing user context is a burgeoning area of research in the field of pervasive computing. The historical freeing of the computer from the desktop has made the environment in which it is able to function an unknown. The unknown nature of computing context due to the mobility of mobile phones is the second theme to emerge from this thesis. As Schmidt suggests, *there is more to context than just location*<sup>34</sup>.

Although immediate geographic location is important, sociocultural, political and information communication technologies are all contextual factors affecting the use of mobile phones in survey research. Geographic location and elements of the immediate environment are most important for researchers interested in MPASI. To date, “snail” mail, email and web self-complete surveys could generally rely on the respondent being at home or in the office when completing the survey. As discussed previously, MPASI surveys can be completed anywhere and at any time. The impact that environmental context has on survey-data quality needs further research.

Sociocultural factors relating to mobile-phone usage may also impact both on MPAPI and MPASI surveys. As reported in this thesis, rural South African women make considerable use of the camera on their phone. This usage pattern shaped the expectations they had of the MPAPI survey, even though photos were not taken as part of data collection.

Similarly, the high turnover of phones among urban young adults limits the usefulness of a mobile-phone survey application installed on a single device.

In the context of varying national health priorities, mobile-phone-assisted data collection is progressing in two clear directions. On the one hand, resource-rich nations are moving more strongly towards the use of mobile phones as a tool for the support of participatory outpatient care<sup>117,121,130,140</sup>. In the context of resource-limited nations, surveillance, health-system strengthening and access to health resources are being prioritised<sup>127,148,233–235</sup>. These health-system strengthening activities are aimed at improving the underlying health systems of a country in a way that allows for the delivery of more equitable and sustainable health care services.

The final context to be considered is the availability of information communication technologies. In countries with poor mobile network coverage, the usefulness of mobile-phone surveys may be limited. The functionality available in commonly owned mobile-phone handsets also sets a specific context for the type of data that may be collected. In a country where people predominantly own mobile phones with limited feature sets and slow Internet connection speeds, surveys containing multimedia content or requesting geolocation coordinates are not feasible. If high-specification “smartphones” are widespread, health researchers have more options available to them in this context.

### **8.3 Theoretical Relevance**

Following from the consolidation of thesis findings and discussion of research themes, the conceptual framework of Tourangeau et al.<sup>76</sup> will now be revisited. The contribution of

this thesis to the understanding and development of the model will then be discussed. In particular, this work has again highlighted the importance of considering the impact of the data-collection mode on both total survey error and the multiple facets of data quality.

### **8.3.1 Tourangeau et al.'s conceptual model of how psychological variables, mediated by mode, affect data quality**

The framework proposed by Tourangeau et al.<sup>76</sup> presents a convenient model for understanding how mode features affect response quality through their interaction with the psychological appraisal (i.e. the emotional and cognitive meaning people extract from their evaluation of an event) of question response. Research findings can be applied in a limited way to the model. Studies performed as part of this thesis explored how mobile-phone-assisted surveying affected response quality.

Tourangeau et al.<sup>76</sup> suggest that self-administration, computerisation and auditory presentation are three key mode features affecting respondents' psychological appraisal of items in the survey.

- **Self-administration**

The findings of this research support the above model regarding the significance of whether or not the mobile-phone-assisted interview was performed face-to-face or through self-administration. Pronounced differences were found in response quality for MPAPI and MPASI surveys.

- **Computerisation**

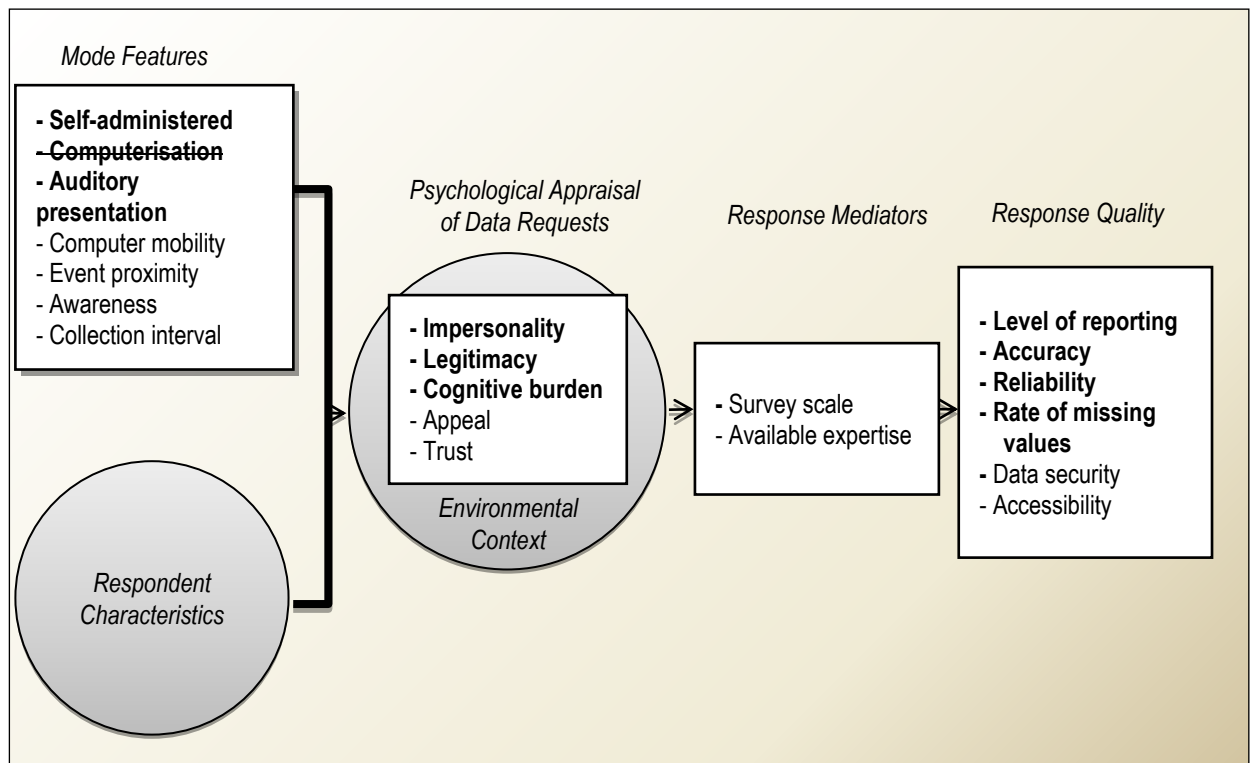
Computerisation, as described by Tourangeau et al.<sup>76</sup> in 2000, is too general a term to be useful in describing the wide range of electronic survey modes available twelve years later. To address this gap, some suggestions to address this situation have been put forward in the following section.

- **Auditory**

The second mode feature in the framework is auditory versus visual presentation of a survey. The thesis did not explore this dimension in any detail and its applicability to mobile-phone-assisted interviewing cannot be reported on.

### **8.3.2 A generalised pervasive computing framework for health research suggested from the thesis**

Building on the foundation of these findings and the model discussed above, an adapted conceptual framework is suggested and presented in Figure 8.2. This framework is particularly geared towards providing a framework for the assessment of novel electronic data-collection techniques. In addition to reworking original model components of mode features, psychological appraisal and response quality, the adaptation suggests the incorporation of three new dimensions that affect the response quality of electronically collected data, which include respondent characteristics, environmental context and response mediators.



**Figure 8.2. A conceptual framework, adapted from Tourangeau et al.,<sup>76</sup> hypothesising the impact of electronic data-collection survey modes on psychological appraisal and response quality**

### 8.3.2.1 Mode features

Adaptions to the mode features listed by Tourangeau include the following:

- The term “computerisation” has been removed as it is inadequate to describe fully the impact of electronic devices on data-collection quality.
- Mobility is a fundamental driver of current technology trends<sup>62</sup> and has been appended to the model. The mobility of a device will impact heavily on the locational context in which interviews take place. Respondents become reachable at any time and anywhere<sup>237</sup>. The impact of this reality needs to be empirically tested in order to establish its effect on psychological appraisal.
- The temporal proximity of data collection to the behaviour of interest has been appended to the model. Formally known as ecological momentary assessment

(EMA), this technique, used in conjunction with a mobile data-collection platform, has the potential to improve data quality over traditional surveys by reducing recall bias. Presenting the question in close proximity to the behaviour in question will reduce the respondent's cognitive burden and thereby increase response quality.

- Awareness by participants of their data being collected is the final mode feature that has been appended to the model. Given current trends, it seems reasonable to conclude that data collection itself may subside from the respondent's awareness. The work of Pentland, Eagle and colleagues<sup>35,161,243</sup> is indicative of the potential to collect data that is not at the forefront of respondents' attention. However, the ethical implications of this type of data collection need to be considered carefully. Issues of privacy, security and confidentiality are all cast in a new light if data collection drifts below awareness<sup>243–245</sup>.

#### **8.3.2.2 Psychological appraisal of data request**

The adapted model is expanded to include two additional appraisal mechanisms that may influence response quality.

- **Appeal**

As mode features allow the data to be more easily collected, regular survey researchers will likely find that they are competing with an array of alternative media and advertising messages for the attention of respondents. The appeal of the survey may become an important driver of participation. User interface patterns that increase engagement and reduce cognitive load will need to be

developed. Creativity will be needed in the research design phase to increase the appeal of the survey if it is not to be lost in the noise of media overload.

- **Trust**

The second addition to the model is a greater emphasis on trust. Future research will need to place a strong emphasis on building up trust with respondents. This will be particularly important for self-administered or autonomous data-collection activities. Trust and rapport will need to be established if data quality is to be maximised.

#### **8.3.2.3 Response mediators**

A theme that emerged across all the studies undertaken as part of this thesis is the importance of survey scale on aggregated response quality. As the extent of interviewer-led electronic data-collection surveys grow, so does its benefit for data quality. Self-administered mobile survey quality was found to diminish with scale due to an increase in technical and project management challenges. These challenges would be mediated by the type of expertise available to a research team. With experience and strong technical skills within a research team, self-administered surveys may also show a positive correlation between scale and quality. Part of the explanation is that, although the upfront costs are higher for electronic modes, they do have an advantage over pen-and-paper in terms of data quality. The longer the timeframe and more complex the data-collection protocol, the more pronounced these issues become.

#### **8.3.2.4 Response quality**

Several response-quality measures are presented in the original framework. The findings from this thesis suggest the addition of two quality measures that are of specific interest to electronic modes of data collection.

- **Data security**

With more and more data moving digitally between multiple devices and ultimately being stored online, data security will increasingly become an important measure and determinant of response quality. Maintenance of confidentiality and trust through securing the collected data will need to be prioritised.

- **Accessibility of data**

A paradox inherent in these new modes of data collection is that the greater demands for data security stand alongside the greater opportunities for data accessibility and collaboration amongst research fraternities. These two concepts need not be in conflict. Carefully prepared data that is made available could increase transparency, reproducibility and scientific discovery. Addressing data accessibility will promote the scientific process and need to be considered an important part of assessing overall data quality.

#### **8.4 Contextual Relevance**

As discussed in Figure 2.7, Heeks<sup>102</sup> shows that the potential of new technologies means little if they do not translate into useful real-world impact. The contextual relevance of mobile phones to South African health research will be discussed with reference to four of the five thematic areas suggested by Mechael et al.<sup>249</sup>

### **8.4.1 Treatment compliance**

Growing evidence suggests that the use of mobile phones may be a valuable tool for treatment adherence<sup>157,208,250,251</sup>. Both randomised control trials and anecdotal evidence suggest that treatment medication reminders, delivered via mobile phone, for HIV/AIDS and TB have been found to be effective in low and middle-income countries (LMICs). In a frequently cited randomised control trial discussed previously<sup>154</sup>, Lester and colleagues found that for a cost of around US\$0.20 cents per 100 patients each month, and with a low impact on health staff requirements (one nurse monitored the adherence of 1 000 patients), ART adherence and viral load were significantly improved for those receiving a text-message reminder. If brought to scale in Kenya, the authors calculate that the intervention would result in an additional 26 354 people with suppressed viral load. Treatment compliance research in high-income countries has focused primarily on the long-term management of chronic diseases such as type 2 diabetes<sup>249</sup>. Most of this impact research has focused on the use of SMS as the communication channel of choice. Studies in high-income settings are also starting to explore the use of mobile health apps and web portals<sup>130</sup> in treatment compliance. The ability to get messages to people easily and affordably as they go about their daily activities, makes this area of research important and relevant for the improved health of populations.

### **8.4.2 Data collection**

This thesis has looked in detail at the role of mobile-phone-assisted data collection in resource-limited health environments. Mobile-phone-assisted data collection, whether applied to disease surveillance, the strengthening of health information or health

research, has the potential to improve data quality. However, evidence remains mixed as to its effectiveness<sup>249</sup>. Lack of clear evidence in the findings may be the unintended result of study scale, as discussed in this thesis. Improved timeliness of data is one element of data quality that is not disputed. This has particular significance and relevance in the context of disease surveillance. Reducing the amount of time it takes to collect disease and syndromic data has great potential to improve low and middle-income countries' sentinel surveillance systems. For maximum value to be obtained from these data-collection tools, it is necessary for standards to be put in place that govern the interoperability of the data that is collected. Current data-collection tools are often custom-built for the task. This limits data accessibility and reduces its value. By building data-collection tools that conform to a standard, such as the Open mHealth Architecture<sup>131</sup>, vertical data silos can be avoided and data usefulness increased.

#### **8.4.3 Health information systems and point-of-care support tools for health workers**

Making electronic health records mobile is an important area of relevance for countries where weak health-information systems make linkage to care difficult. As an example, in 2009 South Africa had around one million people on ART, making it the largest ART programme in the world<sup>252</sup>. This rapid scale-up has created a number of challenges for the monitoring and evaluation of the programme. Cornell et al. (2010) note that loss to follow-up in the ART programme has increased from 1% after 12 months in 2001 to 13% in 2006. After 36 months, a third of all patients were lost. A major challenge facing decentralised community-based services, such as home-based counselling and testing, is how best to ensure people are integrated back into specialised and ongoing care provided

by health facilities and are not lost to follow-up. Mobile-phone-assisted data-collection systems could support health workers in this regard.

#### **8.4.4 Disease prevention and health promotion**

Although mostly discussed in the context of high-income countries, the relevance of mobile phones in health promotion has been established. Smoking cessation, diabetes education and encouragement of individuals to attend their primary care appointments have all achieved positive results<sup>122</sup>. The mobility, high prevalence, bi-directional communication and context-aware sensors of mobile phones all contribute to their value as a tool with which to promote healthy behaviours. Among the most well-known mobile-phone-based behavioural interventions is text4baby. Text4baby is a free mobile text service in the United States that provides free health information to pregnant women and new mothers<sup>224</sup>. A similar mHealth service for pregnant women and their families in rural Ghana is available through the Mobile Technology for Community Health (MOTECH) platform. Launched in February 2010, text4baby has gained wide media attention, winning awards from the United States Office of Science and Technology in 2010 and the Public Relations Society of America (PRSA) in 2011. Yet formal results are still to be made available from a belated evaluation that was recently completed<sup>225</sup>. Whether the 3.5 million messages sent in 2010 were read, acted upon or in any way able to change behaviour is not known. This lack of formal evaluation is one of two problems that beset much of the work in the field of mHealth. The second problem is that much of the research and many of the mHealth programmes are small, unsustainable pilot programmes. Although a little dated, data from a 2008 review of mHealth literature revealed that 84% of the published programmes were prototypes, pilots or tests<sup>231</sup>. These

general limitations, along with specific limitations of the studies in this thesis and directions for future research, will be addressed next.

## **8.5 Limitation and strength**

### **8.5.1 Single mobile data collection method**

A limitation of the research presented in this thesis is the reliance on only one form of mobile-phone-assisted data collection. All the studies in this thesis relied on survey software installed on mobile-phone handsets. A more detailed understanding may have been gained if other communication channels had been used to collect data. These could have included SMS, Internet survey or the use of interactive voice response systems. Whether and how these alternative channels could affect data quality still needs to be explored.

### **8.5.2 Strength in multi-study design**

A strength of this research is that rather than being one large study, it presents a series of related studies that enables different study designs to work together. This can produce multiple perspectives and facilitate a nuanced understanding of the different aspects of mobile-phone-assisted interviewing.

## **8.6 Further research**

The field of mHealth offers a wide range of opportunities for further research. Most pressing are studies that contribute to a stronger evidence base for the effectiveness and sustainability of mHealth programmes.

- **The effectiveness of mHealth programmes**

Formal and rigorous evaluations of the efficacy, effectiveness and cost-effectiveness of mHealth programmes must be undertaken to ensure that time and money are not wasted on ineffective programmes, and to ensure that the best practices of successful programmes are rapidly disseminated.

- **The sustainability of mHealth programmes**

The prevalence of unsustainable pilot programmes supported primarily by external funding needs to be addressed. Research that prioritises the need to find sustainable and cost-effective mHealth models should be given priority. Existing exploratory mHealth research should be supported and developed into sustainable programmes that have been thoroughly evaluated for effectiveness.

Two immediate studies that will follow on from this thesis will focus on efforts to improve the response rates of MPASI surveys, and on Community Health Worker Assistive Technologies (CHAT).

- **Efforts to improve the response rates of MPASI surveys**

The potential of MPASI to collect information from participants who are hard to reach has been established. However, the low response rates obtained in this work need to be improved for the method to produce unbiased data. This line of research could be taken forward by evaluating the use of SMS surveys with the BT20 cohort. A similar research design will be used to the one described in Chapter 4. Rather than installing mobile-phone survey software on donated handsets, this study will send short, single-select SMS questions to the mobile phones listed as the participant's private phone on the BT20 contact sheets. It is hypothesised that this will lead to improved response

rates while retaining the low levels of item non-response found using mobile-phone survey software.

- **Community Health Worker Assistive Technologies (CHAT)**

I will be involved in a three-year study to develop and test an mHealth innovation to support community health care workers in their delivery of home-based HIV education, support and referral. A mobile-phone based tool called Community Health Worker Assistive Technologies (CHAT) will be developed for tablet computers and will leverage available, low-cost devices and open-source software. CHAT will provide community health workers (CHW) with sustained, integrated support, including protocol training, powerful handheld technologies and accessible, media-enhanced resources (e.g., health education videos and animations) that will promote interactive discussions with clients. It will enable electronic data capture that supports high-quality information-gathering using data validation and branching logic, and will support communication with clinic-based health providers and CHW supervisors.

## **8.7 Conclusion**

This thesis reports on work done to better understand the collection of health care data using mobile phones. It aimed to implement and evaluate the use of mobile phones as research instruments with which to collect and analyse information for research in low-resourced Sub-Saharan African settings. In conclusion:

- Mobile-phone-assisted personal interviewing is a feasible and acceptable way to collect data in low-resource, rural settings.

- The benefits of mobile-phone-assisted personal-interviewing over pen and paper are multiplied as the volume of surveys and geographic boundaries of the data collection effort are increased.
- Mobile-phone-assisted self-interviewing with young adults holds promise only if issues of response rate can be adequately addressed.

mHealth is on the cusp of revolutionising the collection of health care data for research and clinical care. In low and middle-income countries this revolution will have the positive effect of increasing access to medical information, improving treatment decisions and providing critical support for chronic disease management. Every effort needs to be made to harness current and emerging technologies for improved community research and its application, to assist with health care delivery for all.

*While on average there are only 14 doctors for every 10 000 people, there are currently 7 620 mobile phones for every 10 000 people!*<sup>254</sup>

## References

---

1. Roberts C. *Mixing modes of data collection in surveys: A methodological review*. London; 2007:1-26.
2. Cui Y, Chipchase J, Ichikawa F. A Cross Culture Study on Phone Carrying and Physical personalization. In: *UI-HCII'07 Proceedings of the 2nd international conference on Usability and internationalization*. Heidelberg: Springer-Verlag Berlin; 2007:483-492.
3. Weiser M. The Computer for the 21st Century. *Scientific American*. 1991;94-104.
4. Ronzani D. The Battle Of Concepts: Ubiquitous Computing, Pervasive Computing and Ambient Intelligence in Mass Media . *Ubiquitous Computing and Communication*. 2000;4(2):9-19.
5. Darwin C. *On the Origin of Species*. London: John Murray; 1859:502.
6. Butler S. *Darwin among the Machines*. The Press. 1863.
7. Richta R. *Man and technology in the revolution of our days*. Prague: Czechoslovak society for the dissemination of political and scientific knowledge; 1963.
8. McLennan. *The Virtual World of Work: How to Gain Competitive Advantage Through the Virtual Workplace*. Charlotte, NC: Information Age Publishing; 2007:364.
9. Gersick CJG. Revolutionary Change Theories: A Multilevel Exploration of the Punctuated Equilibrium Paradigm. *The Academy of Management Review*. 1991;16(1):10.
10. Kuhn T. *The Structure of Scientific Revolutions*. 3rd ed. Chicago, IL: University of Chicago Press; 1996:173.
11. Eldredge N, Gould S. Punctuated equilibria: An alternative to phyletic gradualism,. In: Schopf T, ed. *Models in paleobiology*. San Francisco: Freeman, Cooper & Co. 1972:82-115.
12. Tushman ML, Anderson P. Technological Discontinuities and Organizational Environments. *Administrative Science Quarterly*. 1986;31(3):439-465.
13. Foster R. *Innovation: The Attacker's Advantage*. New York: Summit Books; 1986.
14. Christensen M. Exploring the Limits of the Technology S-Curve, Part I: Component Technologies. *Production and Operations Management*. 1992;1(4).
15. Sood A, Tellis GJ. Technological Evolution and Radical Innovation. *Journal of Marketing*. 2005;69(July):152-168.
16. Basalla G. *The Evolution of Technology*. Cambridge: Cambridge University Press; 1989.

17. Gilfillan S. *Inventing the Ship*. Chicago: Follett; 1935.
18. Arthur WB. *The Nature of Technology: What It Is and How It Evolves*. London: Penguin Group; 2009:1-88.
19. Weiser M. Nomadic issues in ubiquitous computing. In: *Nomadic '96*. San Jose, CA: Technology Transfer Institute; 1996.
20. Kleinrock L. History of the Internet and its flexible future. *Wireless Communications, IEEE*. 2008;15(1):8-18.
21. Weiser M. Open House. *Review, the web magazine of the Interactive Telecommunications Program of New York University*. 1996.
22. Wooldridge M, Jennings NR. Intelligent Agents: Theory and Practice. *Knowledge Engineering Review*. 1995;10(2):115-152.
23. Kay D. Computer interaction: Debugging the problems. In: Sternberg R, Frensch P, eds. *Complex problem solving: Principles and mechanisms*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1991:317-423.
24. Tesler L. Networked Computing in the 1990's. *Scientific American*. 1991:86-93.
25. Weiser M. Some computer science issues in ubiquitous computing. *Communications of the ACM*. 1993;36(7):75-84.
26. Hansmann U, Stober T, Merk L, Nicklous M. *Pervasive Computing*. New York: Springer; 2003.
27. Mossberg W, Swisher K. D: All Things Digital. 2010.
28. Mukherjee A, Saha D. Pervasive Computing: A Paradigm for the 21st Century. *Computer*. 2003;36(3).
29. Satyanarayanan M. Pervasive Computing: Vision and Challenges. *IEEE Personal Communications*. 2001;8(4):10-17.
30. Gupta AK. Challenges of Mobile Computing. In: *Proceedings of 2nd National Conference on Challenges & Opportunities in Information Technology (COIT-2008)*. Mandi Gobindgarh: RIMT-IET; 2008:86-90.
31. Satyanarayanan M. Fundamental Challenges in Mobile Computing. In: *15th ACM Symposium Principles of Dist. Comp*. Philadelphia; 1996.
32. Abowd GD, Mynatt ED. Charting Past , Present , and Future Research in Ubiquitous Computing. *ACM Transactions on Computer-Human Interaction*. 2000;7(1):29-58.

33. Schilit B, Theimer M. Disseminating Active Map Information to Mobile Hosts. *IEEE Network*. 1994;8(5):22-33.
34. Schmidt A, Beigl M, Gellersen H-w. There is more to Context than Location. *Computers & Graphics Journal*. 1999;23(6):893-902.
35. Eagle N. Machine Perception and Learning of Complex Social Systems. 2009:1. Available at: <http://reality.media.mit.edu/>. Accessed January 14, 2012.
36. Pentland A, Lazer D, Brewer D, Heibeck T. Using reality mining to improve public health and medicine. *Studies in Health Technology and Informatics*. 2009;149(February):93-102.
37. Ley D. *Emerging Technologies for Learning - Ubiquitous Computing*. 2nd ed. Coventry: Beca; 2007:64-79.
38. Estrin D, Culler D, Pister K, Sukhatme G. Connecting the Physical World with Pervasive Networks. *IEEE Pervasive Computing*. 2002;1(1):59-69.
39. Frank C, Bolliger P, Mattern F, Kellerer W. The sensor internet at work: Locating everyday items using mobile phones. *Pervasive and Mobile Computing*. 2008;4(3):421-447.
40. Peddemors A, Eertink H, Niemegeers I. Predicting mobility events on personal devices. *Pervasive and Mobile Computing*. 2010;6(4):401-423.
41. Ye J, Dobson S, McKeever S. Situation identification techniques in pervasive computing: A review. *Pervasive and Mobile Computing*. 2011.
42. Couper MP, Baker RP, Bethlehem J, et al. *Computer Assisted Survey Information Collection*. Wiley-IEEE; 1998:653.
43. Likert R. A technique for the measurement of attitudes. *Archives of Psychology*. 1932;22(140):1-55.
44. Groves RM. *Survey Errors and Survey Costs*. (Kalton G, Rao JNK, Schwarz N, Skinner C, eds.). Hoboken, New Jersey: John Wiley & Sons, Ltd. 2004:616.
45. Deming WE. On Errors in Surveys. *American Sociological Review*. 1944;9(4):359 - 369.
46. Anderson R, Kasper J, Frankel M. *Total Survey Error: Applications to Improve Health Surveys*. San Francisco, CA: Jossey-Bass; 1979.
47. Biemer PP. Total Survey Error: Design, Implementation, and Evaluation. *Public Opinion Quarterly*. 2010;74(5):817-848.
48. Groves RM, Lyberg L. Total Survey Error: Past, Present, and Future. *Public Opinion Quarterly*. 2010;74(5):849-879.

49. Weisberg HF. *The Total Survey Error Approach*. Chicago, IL: The University of Chicago Press; 2005:389.
50. Kish L. *Survey Sampling*. New York: Wiley; 1965.
51. Groves RM, Fowler F, Couper MP, et al. *Survey Methodology*. 1st ed. Hoboken, NJ: Wiley–Interscience; 2004:488.
52. Fricker S, Tourangeau R. An Experimental Comparison of Web and Telephone Surveys. *Public Opinion*. 2005;69(3):370-392.
53. Lee J, Meehl PE. Construct validity in psychological tests. *Psychological Bulletin*. 1955;52(4):281-302.
54. Juran J, Gryna F. *Quality Planning and Analysis*. 2nd ed. New York: McGraw- Hill; 1980.
55. Biemer PP, Lyberg LE. *Introduction to survey quality*. Hoboken, NJ: John Wiley & Sons, Ltd. 2003.
56. International Monetary Fund (IMF). *Data quality assessment framework*. 2003.
57. United States Bureau of the Census. *Definition of data quality, (V1.3)*. Washington, DC; 2006.
58. Wang RY, Strong, Diane M. Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*. 1996;12(4):5-34.
59. Force ADQMT. Practice Brief: Data Quality Management Model. *Journal of AHIMA*. 1998;69(6):2-7.
60. Brackstone G. Managing Data Quality in a Statistical Agency. *Statistics Canada, Survey Methodology*. 1999;25(2):1-23.
61. Couper MP, Burt G. Interviewer Attitudes Toward Computer-Assisted Personal Interviewing (CAPI. *Social Science Computer Review*. 1994;12(1):38-54.
62. Couper MP. Technology Trends in Survey Data Collection. *Social Science Computer Review*. 2005;23(4):486-501.
63. Witte JC, Amoroso LM, Howard PEN. Research Methodology: Method and Representation in Internet-Based Survey Tools: Mobility, Community, and Cultural Identity in Survey 2000. *Social Science Computer Review*. 2000;18(2):179-195.
64. Bishop G, Hippler H-J, Schwarz N, Strack F. A Comparison of Response Effects in Self-Administration and Telephone Surveys. In: Groves RM, ed. *Telephone Survey Methodology*. New York: John Wiley & Sons, Ltd. 1988:321-340.

65. Greenlaw C, Brown-Welty S. A comparison of web-based and paper-based survey methods: testing assumptions of survey mode and response cost. *Evaluation Review*. 2009;33(5):464-80.
66. Groves RM, Khan RL. *Surveys by telephone: A national comparison with persona interviews*. New York: Academic Press; 1979.
67. Mcfarlane B, Garland R. A Comparison of Mail and Face-to-Face Survey Methods: New Zealand Health Professionals. *Marketing Bulletin*. 1994;5:71-81.
68. O'Toole B, Battistutta D, Long A, Crouch K. A comparison of costs and data quality of three health survey methods: mail, telephone and personal home interview. *American Journal of Epidemiology*. 1986;124(2):317-328.
69. Roster CA, Rogers RD, Albaum G, Klein D. A comparison of response characteristics from web and telephone surveys. *International Journal of Market Research*. 2004;46:359-373.
70. Shermis MD, Lombard D. A Comparison Of Survey Data Collected By Regular Mail And Electronic Mail Questionnaires. *Journal of Business and Psychology*. 1999;14(2):341-354.
71. Weber B a, Yarandi H, Rowe M a, Weber JP. A comparison study: paper-based versus web-based data collection and management. *Applied Nursing Research*: ANR. 2005;18(3):182-5.
72. De Leeuw E. *Data quality in mail, telephone, and face-to-face surveys*. Amsterdam: TT Publications; 1992:177.
73. De Leeuw ENI. Technological Innovations in *Data Collection: Acceptance, Data Quality and Costs*. 1996. Available at: <http://www.socresonline.org.uk/1/4/leeuw.html#Edwards>. Accessed May 12, 2008.
74. Simões AMA, Bastos FI. Audio Computer-Assisted Interview: a new technology in the assessment of sexually transmitted diseases, HIV, and drug use. *Cadernos de saúde pública / Ministério da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública*. 20(5):1169-81.
75. Waruru A, Nduati R, Tylleskar T. Audio computer-assisted self-interviewing (ACASI) may avert socially desirable responses about infant feeding in the context of HIV. *BMC Medical Informatics and Decision Making*. 2005;5(1):24.
76. Tourangeau R, Rips L, Rasinski K. *Psychology of survey response*. Cambridge: Cambridge University Press; 2000.
77. Jobe JB. Cognitive psychology and self-reports: models and methods. *Quality of Life Research*. 2003;12(3):219-27.

78. Sabini J, Garvey B, Hall a. L. Shame and Embarrassment Revisited. *Personality and Social Psychology Bulletin*. 2001;27(1):104-117.
79. Robling MR, Ingledew DK, Greene G, et al. Applying an extended theoretical framework for data collection mode to health services research. *BMC Health Services Research*. 2010;10:180.
80. Krosnick JA. Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology*. 1991;5:213-236.
81. Dillman DA, Smyth JD. Design effects in the transition to web-based surveys. *American Journal of Preventive Medicine*. 2007;32(5 Suppl):S90-6.
82. Schaeffer NC, Presser S. The Science of asking Questions. *Annual Review of Sociology*. 2003;29(1):65-88.
83. Couper MP. Web Surveys: A review of issues and approaches. *Public Opinion Quarterly*. 2000;64:464-494.
84. Hox J, De Leeuw E. A comparison of nonresponse in mail, telephone, and face-to-face surveys. *Quality and Quantity*. 1994;28:329-344.
85. De Leeuw E, van der Zouwen J. Data quality in telephone and face to face surveys: A comparative meta-analysis. In: Groves RM, Biemer PN, Lyberg JT, et al., eds. *Telephone Survey Methodology*. New York: Wiley; 1988:283-299.
86. Kaplan CP, Hilton JF, Park-Tanjasiri S, Pérez-Stable EJ. The effect of data collection mode on smoking attitudes and behavior in young African American and Latina women. Face-to-face interview versus self-administered questionnaires. *Evaluation Review*. 2001;25(4):454-73.
87. Goyder J. Face-to-face interviews and mailed questionnaires: The net difference in response rate. *Public Opinion Quarterly*. 1985;49:234-252.
88. Ekholm O, Gundgaard J, Rasmussen NKR, Hansen EH. The effect of health, socio-economic position, and mode of data collection on non-response in health interview surveys. *Scandinavian Journal of Public Health*. 2010;38(7):699-706.
89. Fowler F. *Survey Research Methods*. 4th ed. Thousand Oaks, CA: SAGE Publications, Inc; 2009:216.
90. Shih T-H, Fan X. Comparing Response Rates from Web and Mail Surveys: A Meta-Analysis. *Field Methods*. 2008;20(3):249-271.
91. Shih T-H, Fan X. Comparing response rates in e-mail and paper surveys: A meta-analysis. *Educational Research Review*. 2009;4:26-40.

92. Cook C, Heath F, Thompson RL. A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys. *Educational and Psychological Measurement*. 2000;60(6):821-836.
93. Heerwegh D, Loosveldt G. Face-To-Face Versus Web Surveying In A High-Internet-Coverage Population Differences In Response Quality. *Public Opinion Quarterly*. 2008;72(5):836-846.
94. Tourangeau R, Rasinski K, Jobe J, Smith T, Pratt W. Sources of error in a survey of sexual behavior. *Journal of Official Statistics*. 1997;13:341-365.
95. Harris L, Weinberger M, Tierney W. Assessing inner-city patients' hospital experiences. A controlled trial of telephone interviews versus mailed surveys. *Medical Care*. 1997;35(1):70-76.
96. Gladwell M. *The tipping point: how little things can make a big difference*. New York: Little, Brown and Company; 2000.
97. International Telecommunication Union. *The world in 2011: ICT facts and figures*. Geneva; 2011:1-8.
98. Castells M, Fernandez-Ardevol M, Qiu JL, Sey A. *Mobile communication and society: A global perspective*. Cambridge: MIT Press; 2007:345.
99. Castells M. *The Rise of the Network Society, The Information Age: Economy, Society and Culture, Vol. I*. 1st ed. Oxford: Blackwell Publishers Ltd; 1996.
100. Townsend AM. Life in the Real: Mobile Telephones and Urban Metabolism. *Journal of Urban Technology*. 2000;7(2):85-104.
101. Aker JC, Mbiti IM. Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives*. 2010;24(3):207–232.
102. Heeks R. Do Information and Communication Technologies ( ICTs ) Contribute to Development? *Journal of International Development*. 2010;22:625-640.
103. Geroski P. Models of technology diffusion. *Research Policy*. 2000;29(4-5):603-625.
104. Rogers E. *Diffusion of innovations*. 5th ed. New York: The Free Press; 2003.
105. Chetley A. Improving Health , Connecting People: The Role of ICTs in the Health Sector of Developing Countries A Framework Paper. *October*. 2006;(7):1-65.
106. International Telecommunication Union. Telecommunications/ICT Markets and Trends in Africa. 2007. Available at: <http://www.cominit.com/en/node/269729>. Accessed June 22, 2008.
107. Flynn D. Cellphones ring new lifestyles for ordinary Africans. *Business Report*. 2007.

108. South Africa Advertising Research. All Media Products Study (AMPS). 2011. Available at: <http://www.saarf.co.za>.
109. Nielsen. *Nielsen Mobile Insights 2011*. Johannesburg; 2011.
110. Anon. GSMA Press Release 2008 - Seventy Million More Africans Now Connected To GSM Networks. 2008. Available at: [http://www.gsmworld.com/news/press\\_2008/press08\\_34.shtml](http://www.gsmworld.com/news/press_2008/press08_34.shtml). Accessed October 2, 2008.
111. Buys P, Dasgupta S, Thomas TS, Wheeler D. Determinants of a Digital Divide in Sub-Saharan Africa: A Spatial Econometric Analysis of Cell Phone Coverage. *World Development*. 2009;37(9):1494-1505.
112. Kleine D, Unwin T. Technological Revolution, Evolution and New Dependencies: what's new about ict4d? *Third World Quarterly*. 2009;30(5):1045-1067.
113. Coughlin JF. Old Age, New Technology, and Future Innovations in Disease Management and Home Health Care. *Home Health Care Management & Practice*. 2006;18(3):196-207.
114. Callahan D. Europe and the United States: contrast and convergence. *The Journal of Medicine and Philosophy*. 2008;33(3):280-93.
115. Healthcare O, Economist T, Unit I, Strait T. Australia Healthcare report. *Intelligence*. 2010;(February):9-16.
116. Kenner AM. Securing the Elderly Body: Dementia , Surveillance , and the Politics of "Aging in Place." *Science and Technology*. 2008;5(3):252-269.
117. Roblin DW. The potential of cellular technology to mediate social networks for support of chronic disease self-management. *Journal of Health Communication*. 2011;16 Suppl 1(January 2012):59-76.
118. Forkner-Dunn J. Internet-based patient self-care: the next generation of health care delivery. *Journal of Medical Internet Research*. 2003;5(2).
119. Cheng KG, Ernesto F, Ovalle-Bahamón RE, Truong KN. Barriers to acceptance of personal digital assistants for HIV/AIDS data collection in Angola. *International Journal of Medical Informatics*. 2011;80(8):579-585.
120. Yu P, de Courten M, Pan E, Galea G, Pryor J. The development and evaluation of a PDA-based method for public health surveillance data collection in developing countries. *International Journal of Medical Informatics*. 2009;78(8):532-42.
121. Boulos MNK, Wheeler S, Tavares C, Jones R. How smartphones are changing the face of mobile and participatory healthcare: an overview, with example from eCAALYX. *Biomedical Engineering Online*. 2011;10(1):24.

122. Klasnja P, Pratt W. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. *Journal of Biomedical Informatics*. 2011;1-15.
123. Park H-il, Min W-K, Lee W, et al. Evaluating the short message service alerting system for critical value notification via PDA telephones. *Annals of Clinical and Laboratory Science*. 2008;38(2):149-56.
124. Patnaik S, Brunskill E, Thies W. 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*. IEEE; 2009.
125. Dabas A, Dabas C. Implementation of Real Time Tracking using Unstructured Supplementary Service Data. *Engineering and Technology*. 2009;241-245.
126. Medhi I, Patnaik S, Brunskill E, et al. Designing mobile interfaces for novice and low-literacy users. *ACM Transactions on Computer-Human Interaction*. 2011;18(1):1-28.
127. Kaplan WA. Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? *Globalization and Health*. 2006;2.
128. Vital Wave Consulting. *mHealth for development: the opportunity for mobile technology for healthcare in the developing world*. Washington, D.C. and Berkshire, UK; 2009:65.
129. Waegemann P. mHealth: the next generation of telemedicine? *Telemedicine and e-health*. 2010;16(1):23-5.
130. Handel MJ. mHealth (mobile health)-using Apps for health and wellness. *The Journal of Science and Healing*. 2011;7(4):256-61.
131. Estrin D, Sim I. Open mHealth Architecture<sup>2</sup>: An Engine for Health Care Innovation. *Science*. 2010;330(6005):759-60.
132. Opportunity T. mHealth for Development. *Technology*. 2009;46(1):1-70.
133. WHO Global Observatory for eHealth. *New horizons for health through mobile technologies*. Geneva: World Health Organization; 2011:112.
134. Haapala I, Barengo NC, Biggs S, Surakka L, Manninen P. Weight loss by mobile phone: a 1-year effectiveness study. *Public Health Nutrition*. 2009;12(12):2382-91.
135. Whittaker R, Borland R, Bullen C, et al. Mobile phone-based interventions for smoking cessation (Review). *Cochrane Database of Systematic Review*. 2009;(4).
136. Liang X, Wang Q, Yang X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. *Diabetic medicine*. 2011;28(4):455-63.

137. Hitman G. Mobile phone intervention for diabetes. *Diabetic Medicine*. 2010;28(4):381.
138. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten year revision of the brief behavioral activation treatment for depression: revised treatment manual. *Behavior Modification*. 2011;35(2):111-61.
139. Bluestein A. As Smartphones Get Smarter, You May Get Healthier: How mHealth Can Bring Cheaper Health Care to All. *Fast Company*. 2012.
140. Freifeld CC, Chunara R, Mekaru SR, et al. Participatory Epidemiology: Use of Mobile Phones for Community-Based Health Reporting. *PLoS Medicine*. 2010;7(12):e1000376.
141. Frenk J, Gómez-Dantés O. Globalisation and the challenges to health systems. *BMJ (Clinical research ed.)*. 2002;325(7355):95-7.
142. malERA Consultative Group. A research agenda for malaria eradication: monitoring, evaluation, and surveillance. *PLoS medicine*. 2011;8(1):e1000400.
143. Muscatello DJ, Churches T, Kaldor J, et al. An automated, broad-based, near real-time public health surveillance system using presentations to hospital Emergency Departments in New South Wales, Australia. *BMC Public Health*. 2005;5:141.
144. McNabb S, Chungong S, Ryan M, et al. Conceptual framework of public health surveillance and action and its application in health sector reform. *BMC Public Health*. 2002;2(1):2.
145. Randrianasolo L, Raoelina Y, Ratsitorahina M, et al. Sentinel surveillance system for early outbreak detection in Madagascar. *BMC Public Health*. 2010;10:31.
146. Thinyane H, Hansen S, Foster G, Wilson L. Using Mobile Phones for Rapid Reporting of Zoonotic Diseases in Rural South Africa. *Studies in Health Technology and Informatics*. 2010;161:179-189.
147. Ashar R, Lewis S, Blazes DL, Chretien JP. Applying information and communications technologies to collect health data from remote settings: a systematic assessment of current technologies. *Journal of Biomedical Informatics*. 2010;43(2):332-41.
148. Curioso WH, Karras BT, Campos PE, et al. Design and implementation of Cell-PREVEN: a real-time surveillance system for adverse events using cell phones in Peru. *AMIA Annual Symposium Proceedings*. 2005:176-80.
149. Clarke M, Schoeman H, Friedman I. *Towards a generic surveillance system to measure the impact of Community Health Worker programmes in South Africa: a comparison of paper-based and mobile/cell phone methods*. Durban; 2007:1-133.

150. Tomlinson M, Solomon W, Singh Y, et al. The use of mobile phones as a data collection tool: a report from a household survey in South Africa. *BMC Medical Informatics and Decision Making*. 2009;9:51.
151. Wouters B, Barjis J, Maponya G, Martiz J, Mashiri M. Supporting Home Based Health Care in South African Rural Communities Using USSD Technology. In: *Proceedings of the Fifteenth Americas Conference on Information Systems*. San Francisco, CA; 2009:1-8.
152. DeRenzi B, Borriello G, Jackson J, et al. Mobile Phone Tools for Field-Based Health care Workers in Low-Income Countries. *Mount Sinai Journal of Medicine*. 2011;78:406-418.
153. Chang LW, Arem H, Ssempijja V, et al. Impact of a mHealth Intervention for Peer Health Workers on AIDS Care in Rural Uganda: A Mixed Methods Evaluation of a Cluster-Randomized Trial. *AIDS and Behavior*. 2011;15(8):1776-1784.
154. Lester RT, Ritvo P, Mills EJ, et al. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): a randomised trial. *The Lancet*. 2010;376(9755):1838-1845.
155. Haberer JE, Kiwanuka J, Nansera D, Wilson IB, Bangsberg DR. Challenges in Using Mobile Phones for Collection of Antiretroviral Therapy Adherence Data in a Resource-Limited Setting. *AIDS and Behavior*. 2010.
156. Bridges.org. *Evaluation of the On Cue Compliance Service Pilot. Testing the use of SMS reminders in the treatment of Tuberculosis in South Africa*. Cape Town; 2005.
157. Pop-Eleches C, Thirumurthy H, Habyarimana J, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: a randomized controlled trial of text message reminders. *AIDS*. 2011;25(6):825-834.
158. Lane S, Heddle N, Arnold E, Walker I. A review of randomized controlled trials comparing the effectiveness of hand held computers with paper methods for data collection. *BMC Medical Informatics and Decision Making*. 2006;6(1):23.
159. Blaya J, Fraser HSF, Holt B. E-health technologies show promise in developing countries. *Health Affairs*. 2010;29(2):244-51.
160. Johansen B, Wedderkopp N. Comparison between data obtained through real-time data capture by SMS and a retrospective telephone interview. *Chiropractic & Osteopathy*. 2010;18:10.
161. Eagle N, Pentland AS, Lazer D. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(36):15274-8.

162. Black AD, Car J, Pagliari C, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. *PLoS medicine*. 2011;8(1):e1000387.
163. Acharya S. *Worldwide mobile cellular subscribers to reach 4 billion mark late 2008*. Geneva; 2008.
164. Levine D, McCright J, Dobkin L, Woodruff AJ, Klausner JD. SEXINFO: A Sexual Health Text Messaging Service for San Francisco Youth. *Am J Public Health*. 2008;98(3):393-395.
165. Turrettini E. SMS to find out your HIV status. 2007. Available at: <http://www.textually.org/textually/archives/2007/11/018155.htm>. Accessed November 28, 2008.
166. World Health Organization. South Africa: a novel approach to improving adherence to TB treatment. *Essential Drugs Monitor*. 2003;33(8).
167. Richter LM, Norris S, Pettifor J, Yach D, Cameron N. Cohort Profile: Mandela's children: The 1990 birth to twenty study in South Africa. *International Journal of Epidemiology*. 2007;504-511.
168. Wei R, Lo V-H. Staying connected while on the move: Cell phone use and social connectedness. *New Media Society*. 2006;8(1):53-72.
169. Norris S, Richter LM, Fleetwood SA. Panel studies in developing countries: case analysis of sample attrition over the past 16 years within the birth to twenty cohort in Johannesburg, South Africa. *Journal of International Development*. 2007;19(8):1143-1150.
170. Richter LM, Panday S, Norris S. Factors influencing enrollment: a case study from Birth to Twenty, the 1990 birth cohort in Soweto-Johannesburg. *Evaluation and Program Planning*. 2009;32(3):197-203.
171. Das Gupta M, Aaby P, Garenne M, Gilles P. *Prospective community studies in developing countries*. Oxford: Clarendon Press; 1997.
172. Hunt JR, White E. Retaining and Tracking Cohort Study Members. *Public Health*. 1998;20(1):57-70.
173. Reyna VF, Farley F. Risk and Rationality in Adolescent Decision Making: Implications for Theory, Practice, and Public Policy. *Psychological Science in the Public Interest*. 2006;7(1):1-44.
174. Galambos NL, Leadbeater BJ. Trends in adolescent research for the new millennium. *Society*. 2000;24(3):289-294.
175. Christian BJ, Pearce PF, Roberson AJ, Rothwell E. It's a small, small world: data collection strategies for research with children and adolescents. *Journal of Pediatric Nursing*. 2010;25(3):202-14.

176. Hadley E, Smith C, Gallo A, Angst D, Knafl K. Parents' perspectives on having their children interviewed for research. *Research in Nursing and Health*. 2008;31(1):4-11.
177. Schilling L, Dixon J, Knafl K, et al. Determining content validity of a self-report instrument for adolescents using a heterogeneous expert panel. *Nursing Research*. 2007;56(5):361-366.
178. O'Reilly JM, Hubbard JT, Lessler PP, Turner C. Audio and Video Computer Assisted Self-Interviewing: Preliminary Tests of New Technologies for Data Collection. *Journal of Official Statistics*. 1994;10(2):197-214.
179. Schwarz N, Strack F, Hippler H-J, Bishop G. The impact of administration mode on response effects in survey measurement. *Applied Cognitive Psychology*. 1991;5(3):193-212.
180. WHO. *The health of young people: a challenge and a promise*. Geneva: WHO; 1993.
181. Oindo ML. Contraception and sexuality among the youth in Kisumu, Kenya. *African Health Sciences*. 2002;2(1):33-9.
182. Slaymaker E. A critique of international indicators of sexual risk behaviour. *Sexually Transmitted Infections*. 2004;80 Suppl 2:ii13-21.
183. Cleland J, Boerma JT, Carael M, Weir SS. Monitoring sexual behaviour in general populations: a synthesis of lessons of the past decade. *Sexually Transmitted Infections*. 2004;80 Suppl 2:ii1-7.
184. Hedeker D, Mermelstein RJ, Berbaum ML, Campbell RT. Modeling mood variation associated with smoking: an application of a heterogeneous mixed-effects model for analysis of ecological momentary assessment (EMA) data. *Addiction*. 2009;104(2):297-307.
185. Marmor J, Oliveria S, Donahue R, et al. Factors encouraging cohort maintenance in a longitudinal study. *Journal of Clinical Epidemiology*. 1991;44(6):531-535.
186. Clyral. Mobenzi Researcher. 2011.
187. Strahan R, Gerbasi K. Short, homogeneous versions of the marlow-crowne social desirability scale. *Journal of Clinical Psychology*. 1972;28(2):191-3.
188. Déglise C, Suggs LS, Odermatt P. Short Message Service (SMS) Applications for Disease Prevention in Developing Countries. *Journal of Medical Internet Research*. 2012;14(1).
189. Van Heerden AC, Norris SA, Richter LM. Using mobile phones for adolescent research in low and middle income countries: preliminary findings from the birth to twenty cohort, South Africa. *The Journal of Adolescent Health*. 2010;46(3):302-4.

190. Wicks P, Vaughan TE, Massagli MP, Heywood J. Accelerated clinical discovery using self-reported patient data collected online and a patient-matching algorithm. *Nature Biotechnology*. 2011;29(5):411-4.
191. UNAIDS. *UNAIDS Report on the Global Aids Epidemic*. 2010:1-364.
192. UNICEF. *State of the World's Children 2009*. New York; 2009.
193. Kinney MV, Kerber KJ, Black RE, et al. Sub-Saharan Africa's mothers, newborns, and children: where and why do they die? *PLoS medicine*. 2010;7(6):e1000294.
194. UN Joint Programme on HIV/AIDS. *Global Report: UNAIDS Report on the Global AIDS Epidemic*.
195. UNAIDS. *Children and AIDS: fourth stocktaking report 2009*. Geneva, Switzerland; 2009.
196. Shisana O, Rehle T, Simbayi L, et al. *South African National HIV Prevalence , Incidence, Behaviour and Communication Survey, 2008 A Turning Tide Among Teenagers?* Cape Town; 2008.
197. Hussain A, Moodley D, Naidoo S, Esterhuizen TM. Pregnant Women's Access to PMTCT and ART Services in South Africa and Implications for Universal Antiretroviral Treatment. *PloS one*. 2011;6(12):e27907.
198. Karim QA, Kharsany ABM, Frohlich J a, et al. Stabilizing HIV prevalence masks high HIV incidence rates amongst rural and urban women in KwaZulu-Natal, South Africa. *International Journal of Epidemiology*. 2011;40(4):922-30.
199. Bwirire LD, Fitzgerald M, Zachariah R, et al. Reasons for loss to follow-up among mothers registered in a prevention-of-mother-to-child transmission program in rural Malawi. *Transactions of the Royal Society of Tropical Medicine and Hygiene*. 2008;102(12):1195-200.
200. Doherty T, Sanders D, Goga A, Jackson D. Implications of the new WHO guidelines on HIV and infant feeding for child survival in South Africa. *Bulletin of the World Health Organization*. 2011;89(1):62-7.
201. WHO. *Towards universal access: Scaling up priority HIV/AIDS interventions in the health sector: progress report 2009*. Geneva; 2009.
202. Rollins N, Little K, Mzolo S, Horwood C, Newell M-L. Surveillance of mother-to-child transmission prevention programmes at immunization clinics: the case for universal screening. *AIDS (London, England)*. 2007;21(10):1341-7.
203. Bärnighausen T, Bloom DE, Humair S. Universal antiretroviral treatment: the challenge of human resources. *Bulletin of the World Health Organization*. 2010;88(12):951-2.

204. Garrib A, Stoops N, McKenzie A, et al. An evaluation of the District Health Information System in rural South Africa. *South African Medical Journal*. 2008;98(7):549-552.
205. Sherman GG, Jones SA, Coovadia AH, Urban MF, Bolton KD. PMTCT from research to reality - results from a routine service. *South African Medical Journal*. 2004;94(4).
206. Mate KS, Bennett B, Mphatswe W, Barker P, Rollins N. Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *PloS one*. 2009;4(5):e5483.
207. Wringe A, Cataldo F, Stevenson N, Fakoya A. Delivering comprehensive home-based care programmes for HIV: a review of lessons learned and challenges ahead in the era of antiretroviral therapy. *Health Policy and Planning*. 2010;25(5):352-62.
208. Crankshaw T, Corless IB, Giddy J, et al. Exploring the patterns of use and the feasibility of using cellular phones for clinic appointment reminders and adherence messages in an antiretroviral treatment clinic, Durban, South Africa. *AIDS Patient Care and STDs*. 2010;24(11):729-34.
209. Rotheram-Borus M-J, Richter L, Van Rooyen H, et al. Project Masihambisane: a cluster randomised controlled trial with peer mentors to improve outcomes for pregnant mothers living with HIV. *Trials*. 2011;12(1):2.
210. Davis F. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*. 1989;(Sept.):319-339.
211. Jacobs E, Punt C, Buthelezi D. *A Profile of the KwaZulu-Natal Province: Demographics, Poverty, Income, Inequality and Unemployment from 2000 till 2007*. Elsenburg, South Africa; 2009:1-46.
212. Alam M, Khanam T, Khan R. Assessing the scope for use of mobile based solution to improve maternal and child health in Bangladesh: A case study. In: *International Conference on Information and Communication Technologies and Development*. London; 2010.
213. Couper MP. Usability Evaluation of Computer-Assisted Survey Instruments. *Social Science Computer Review*. 2000;18(4):384-396.
214. Tourangeau R, Smith TW. Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context. *Public Opinion Quarterly*. 1996;60(2):275.
215. McClamroch KJ. Evaluating the Usability of Personal Digital Assistants to Collect Behavioral Data on Adolescents with Paradata. *Field Methods*. 2011;23(3):219-242.
216. Forster D, Behrens RH, Campbell H, Byass P. Evaluation of a computerized field data collection system for health surveys. *Bulletin of the World Health Organization*. 1991;69(5152):107-111.

217. Seebregts CJ, Zwarenstein M, Mathews C, et al. Handheld computers for survey and trial data collection in resource-poor settings: development and evaluation of PDACT, a Palm Pilot interviewing system. *International Journal of Medical Informatics*. 2009;78(11):721-31.
218. Curioso WH, Kurth AE. Access, use and perceptions regarding Internet, cell phones and PDAs as a means for health promotion for people living with HIV in Peru. *BMC Medical Informatics and Decision Making*. 2007;7:24.
219. Blaya J a, Cohen T, Rodríguez P, Kim J, Fraser HSF. Personal digital assistants to collect tuberculosis bacteriology data in Peru reduce delays, errors, and workload, and are acceptable to users: cluster randomized controlled trial. *International Journal of Infectious Diseases*: *IJID*: official publication of the International Society for Infectious Diseases. 2009;13(3):410-8.
220. Childs J, Landreth A. Analyzing Interviewer/Respondent Interactions while Using a Mobile Computer-Assisted Personal Interview Device. *Field Methods*. 2006;18(3):335-351.
221. Hoch JS, Smith MW. A Guide to Economic Evaluation: Methods for Cost-Effectiveness Analysis of Person-Level Data. *Journal of Traumatic Stress*. 2006;19(6):787-797.
222. Palen L-A, Graham JW, Smith EA, et al. Rates of Missing Responses in Personal Digital Assistant (PDA) Versus Paper Assessments. *Evaluation Review*. 2008;32(3):257-72.
223. Missinou MA, Olola CHO, Issifou S, et al. Short report: Piloting paperless data entry for clinical research in Africa. *The American Journal of Tropical Medicine and Hygiene*. 2005;72(3):301-3.
224. Jordan E, Ray E, Johnson P, Evans D. Text4Baby: Using Text Messaging to Improve Maternal and Newborn Health. *Nursing for Women's Health*. 2011;15(3):206-212.
225. Vaughan C. *First Text4baby Results Available*. San Marcos; 2011.
226. Tomlinson M. Family-centred HIV interventions: lessons from the field of parental depression. *Journal of the International AIDS Society*. 2010;13 Suppl 2:S9.
227. De Maeseneer J, Roberts RG, Demarzo M, et al. Tackling NCDs: a different approach is needed. *The Lancet*. 2011;6736(11):11-12.
228. Breslauer DN, Maamari RN, Switz N a, Lam W a, Fletcher D a. Mobile phone based clinical microscopy for global health applications. *PloS one*. 2009;4(7):e6320.
229. Lejbkowitz I, Paperna T, Stein N, Dishon S, Miller A. Internet Usage by Patients with Multiple Sclerosis: Implications to Participatory Medicine and Personalized Healthcare. *Multiple Sclerosis International*. 2010;2010:1-7.

230. Pare G, Jaana M, Sicotte C. Systematic Review of Home Telemonitoring for Chronic Diseases: The Evidence Base. *Journal of the American Medical Informatics Association*. 2007;14(3):269-277.
231. Orwat C, Graefe A, Faulwasser T. Towards pervasive computing in health care - a literature review. *BMC Medical Informatics and Decision Making*. 2008;8:26.
232. Palen L, Salzman M. Voice-Mail Diary Studies for Naturalistic Data Capture Under Mobile Conditions. In: *Proceedings of CSCW: ACM Conference on Computer Supported Cooperative Work*; 2002:87-95.
233. Curioso WH, Mechael PN. Enhancing "M-health" with south-to-south collaborations. *Health Affairs (Project Hope)*. 2010;29(2):264-7.
234. Rashid AT, Elder L. Mobile Phones and Development: An Analysis of IDRC-Supported Projects. *The Electronic Journal on Information Systems in Developing Countries*. 2009;36(2):1-16.
235. Morris K. Mobile phones connecting efforts to tackle infectious disease. *The Lancet Infectious Diseases*. 2009;9(5):274-274.
236. Tourangeau R. Survey research and societal change. *Annual Review of Psychology*. 2004;55:775-801.
237. Ling R, Donner J. *Mobile Phones and Mobile Communication (DMS - Digital Media and Society)*. Cambridge: Polity Press; 2009:191.
238. Courvoisier DS, Eid M, Lischetzke T. Compliance to a cell phone-based ecological momentary assessment study: The effect of time and personality characteristics. *Psychological Assessment*. 2012.
239. Kreuter F, Couper MP, Lyberg LE. The Use of Paradata to Monitor and Manage Survey Data Collection. In: *Proceedings of the Joint Statistical Meetings, American Statistical Association*. Alexandria: ASA; 2010:282-296.
240. Couper MP, Tourangeau R, Marvin T. Taking the Audio Out of Audio-CASI. *Public Opinion Quarterly*. 2009;73(2):281-303.
241. Galesic M, Tourangeau R, Couper MP, Conrad F. Eye-tracking data: New insights on response order effects and other cognitive shortcuts in survey responding. *Public Opinion Quarterly*. 2009;72:892-913.
242. Mueller F, Lockerd A. Cheese: Tracking mouse movement activity on websites, a tool for user modeling. In: *Proceedings from Conference on Human Factors in Computing Systems*. Seattle, WA; 2011.
243. Atallah L, Yang G-Z. The use of pervasive sensing for behaviour profiling — a survey. *Pervasive and Mobile Computing*. 2009;5(5):447-464.

244. Raento M, Oulasvirta a., Eagle N. Smartphones: An Emerging Tool for Social Scientists. *Sociological Methods & Research*. 2009;37(3):426-454.
245. Eagle N, (Sandy) Pentland A. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*. 2006;10(4):255-268.
246. Deterding S, O'Hare K, Sicart M, Dixon D, Nacke L. Gamification<sup>2</sup>: Using Game Design Elements in Non-Gaming Contexts. In: *CHI Conference on Human Factors in Computing Systems*. Vancouver, BC, Canada: ACM; 2011:5-8.
247. Deterding S, Dixon D, Khaled R, Nacke L. Gamification<sup>2</sup>: Toward a Definition. In: *CHI Conference on Human Factors in Computing Systems*. Vancouver, BC, Canada: ACM; 2011:12-15.
248. Molloy JC. The Open Knowledge Foundation: Open Data Means Better Science. *PLoS Biology*. 2011;9(12):e1001195.
249. Mechael P, Batavia H, Kaonga N, et al. *Barriers and Gaps Affecting mHealth in Low and Middle Income Countries<sup>2</sup>: Policy White Paper*. New York; 2010:1-79.
250. Lester RT, Mills EJ, Kariri A, et al. The HAART cell phone adherence trial (WeTel Kenya1): a randomized controlled trial protocol. *Trials*. 2009;10:87.
251. Kunutsor S, Walley J, Katabira E, et al. Improving Clinic Attendance and Adherence to Antiretroviral Therapy Through a Treatment Supporter Intervention in Uganda: A Randomized Controlled Trial. *AIDS and Behavior*. 2011.
252. WHO. *Towards Universal access: Scaling up priority HIV/AIDS interventions in the health sector. Progress Report 2010*. Geneva, Switzerland; 2010.
253. Cornell M, Grimsrud A, Fairall L, et al. Temporal changes in programme outcomes among adult patients initiating antiretroviral therapy across South Africa, 2002-2007. *AIDS*. 2010;24(15):2263-70.
254. Ritchie M. *GSMAmHealth Alliance mHealth Summit. Summit report prepared for NDoH and DFIDSA*. Cape Town; 2012:9.
255. Braun, V., Clarke, V. Using thematic analysis in psychology. *Qualitative Research in Psychology*. 2006;3(2):77-101.

---

# Appendices

---

## Appendix 1 – Approval of thesis title

---



Mr AC Van Heerden  
16 Kings Place  
Buckingham Avenue  
P.M.B.  
3201  
South Africa

Faculty of Health Sciences  
Medical School, 7 York Road, Parktown, 2193  
Fax: (011) 717-2119  
Tel: (011) 717-2745

Reference: Ms Tania Van Leeve  
E-mail: [tania.vanleeve@wits.ac.za](mailto:tania.vanleeve@wits.ac.za)  
26 November 2009  
Person No: 379316  
PAG

Dear Mr Van Heerden

### **Doctor of Philosophy: Approval of Title**

We have pleasure in advising that your proposal entitled "*Pervasive computing and public health research in Africa: Mobile phones in the collection, analysis and dissemination of health research*" has been approved. Please note that any amendments to this title have to be endorsed by the Faculty's higher degrees committee and formally approved.

Yours sincerely

A handwritten signature in black ink, appearing to read 'S Benn'.

Mrs Sandra Benn  
Faculty Registrar  
Faculty of Health Sciences

## Appendix 2 – Ethical approval

---

**UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG**  
Division of the Deputy Registrar (Research)

**HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)**  
R14/49 Mr Alastair van Heerden

**CLEARANCE CERTIFICATE**

**M091035**

**PROJECT**

Pervasive Computing and Public Health  
Research in Africa: Mobile Phones in the  
Collection, Analysis and Dissemination of  
Health Research

**INVESTIGATORS**

Mr Alastair van Heerden.

**DEPARTMENT**

School of Public Health

**DATE CONSIDERED**

2009/10/30

**DECISION OF THE COMMITTEE\***

Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

**DATE** 14/01/2010

**CHAIRPERSON** .....  
(Professor PE Cleaton-Jones)

\*Guidelines for written 'informed consent' attached where applicable  
cc: Supervisor : Dr S Norris

**DECLARATION OF INVESTIGATOR(S)**

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.

I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

## Point of care in your pocket: a research agenda for the field of m-health

Alastair van Heerden,<sup>a</sup> Mark Tomlinson<sup>b</sup> & Leslie Swartz<sup>b</sup>

### Introduction

Mobile health, also known as m-health,<sup>1</sup> has been defined by the Global Observatory for e-health of the World Health Organization (WHO) as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants and other wireless devices”.<sup>2</sup> With global penetration of cellular services topping 87%, the opportunity for connecting with 6 billion people has brought significant enthusiasm, energy and media attention to the field of m-health.<sup>3</sup> Amidst this momentum, evidence for the health benefits of m-health has been built not through traditional, large-scale and time-intensive randomized controlled trials, but by a multitude of small pilot projects, particularly in low- and middle- income countries.<sup>3</sup> Rapid technological innovation has led to an enthusiastic proliferation of untested methods, which are small-scale and, for the most part, have little policy coordination.<sup>3</sup> To lay the groundwork for establishing an evidence base for the cost-effectiveness and programme utility of m-health innovations, we outline seven key recommendations aimed at reducing resource wastage and programme duplication.

### 1. M-health needs to develop an evidence base.

A frontier mindset fosters real and rapid innovation in mobile health. This fast, iterative approach should not be lost, especially given the speed with which mobile and wireless technologies evolve, but caution is warranted. Although well on track towards attaining its goal of reaching 1 million mothers by 2012, Text4Baby, a free mobile text service sponsored by Johnson & Johnson in the United States of America, is a case in point.<sup>4</sup> Launched in February 2010, Text4Baby, which provides free health information to pregnant women and new mothers, has gained wide media

attention and has won awards from the United States Office of Science and Technology in 2010 and from the Public Relations Society of America in 2011. Yet formal results are still to be made available from a belated evaluation that was recently completed.<sup>5</sup> Whether the 3.5 million messages sent in 2010 were read, acted upon or in any way able to change behaviour is not known. Only 12% of the countries surveyed in 2011 by WHO's Global Observatory for e-health reported evaluating their m-health services.<sup>1</sup> Formal and rigorous evaluations of efficacy, effectiveness and cost-effectiveness of m-health programmes must be undertaken to ensure that time and money are not wasted on ineffective programmes and that the best practices of successful programmes are rapidly disseminated.

### 2. M-health systems should be interoperable with existing e-health initiatives.

M-health is emerging as a patchwork of incompatible applications.<sup>3</sup> Information technology (IT) and health information systems are often based on the use of devices that cannot communicate and share data with one another and have traditionally been characterized by a silo approach.<sup>3</sup> This lack of interoperability parallels the public health experience with vertical disease-oriented programmes.<sup>6</sup> There is a compelling argument that better outcomes can be achieved by tackling diseases through an integrated health systems strengthening approach.<sup>7</sup> Similarly, by enabling m-health systems to share information with one another as well as with broader e-health systems, we can improve efficiency, reduce the costs associated with data collection and provide patients with better care. In addition, benefits such as these may in fact be felt within the broader health-care system.<sup>3</sup> To achieve this will require strong cooperation, meaningful collaboration and adherence to international standards.

### 3. M-health should adopt and implement the same standards already present in e-health.

Several e-health standards are in use today. They were designed to facilitate the interoperability of health data between devices and IT systems used in health care. In order for different systems to share data, they must contain comparable information (data elements), employ comparable words (terminology) and use an agreed-upon way to communicate (messaging). M-health projects such as cellScope,<sup>8</sup> which enables light microscopy to be performed on mobile phones, make mention of the Digital Imaging and Communications in Medicine (DICOM) standard for the exchange of medical images. By upholding and implementing the DICOM standard, they provide an excellent example of how m-health can build on existing e-health standards. Most projects are not using this standard, nor the HL7 Clinical Document Architecture (CDA) for the exchange of electronic clinical documents (such as discharge summaries and progress notes). Adherence to privacy, security and confidentiality standards is also critical to the ongoing success of any m-health intervention. If m-health and e-health systems are interoperable and share common standards, integration costs could be better contained.

### 4. M-health should take a participatory approach.

Contemporary thinking in health care is that health care should be personalized, predictive, preventive and participatory.<sup>9</sup> M-health can facilitate this by providing a simple means by which patients can engage in decisions about their own health care. Programmes need to go beyond merely connecting patients with health-care providers. Many current m-health initiatives focus on outdated, unidirectional models of patient communication (e.g. exclusively collecting data, providing information

<sup>a</sup> Human Sciences Research Council, PO Box 90, Msunduzi, Pietermaritzburg, KwaZulu-Natal, South Africa.

<sup>b</sup> Department of Psychology, Stellenbosch University, Stellenbosch, South Africa.

Correspondence to Alastair van Heerden (e-mail: [avanheerden@hsr.ac.za](mailto:avanheerden@hsr.ac.za)).

(Submitted: 19 November 2011 – Revised version received: 9 February 2012 – Accepted: 16 February 2012)

or sending reminders). With m-health technology, patients can easily transmit data and commentary on their health and health care to their electronic health record, which allows for an ongoing and active health dialogue.<sup>10</sup> However, this ease of use introduces additional challenges in connection with the privacy and confidentiality of patient data and requires the development of appropriate legal policies and frameworks.<sup>11</sup> Research that evaluates the utility and cost-effectiveness of such a participatory care approach should be prioritized.

### 5. M-health should promote equity in health.

In high-income countries, m-health programmes have focused largely on supporting ageing populations through delivering prevention programmes and reducing health-care costs through home-based monitoring.<sup>12</sup> In low- and middle-income countries, however, m-health has been conceptualized much more broadly, as a way to promote access to health services for the socioeconomically and geographically disadvantaged. Irrespective of country, no m-health programme should unfairly discriminate against, marginalize or exclude people at the fringes of society. The very high mobile phone penetration rate of 79%<sup>3</sup> in low- and middle-income countries may nevertheless obscure inequities and patterns of exclusion. Equity issues in the field of m-health remain under-researched.

### 6. M-health programmes need to move towards sustainability.

The most common m-health initiatives

among WHO Member States in 2011 were health call centres (59%), emergency toll-free telephone services (55%), management of emergencies and disasters (54%) and mobile telemedicine (49%).<sup>1</sup> Unfortunately, many of these programmes are small, unsustainable pilot projects. Although a little dated, a 2008 review of the m-health literature revealed that 84% of the published programmes were prototypes, pilots or tests.<sup>9</sup> The figure has probably improved since then, but it still points to an urgent need to find sustainable and cost-effective models for supporting and developing early exploratory m-health ideas and integrating them with local health care financing systems.

### 7. M-health needs to focus on health, not on the technology.

Debate exists as to whether m-health services should be designed to have the best coverage by using the lowest common denominator technology, such as text messages, or whether they should benefit from the advantages that can be gained by targeting the sensors and processing power of the latest generation smartphone. This discussion may be practical but slightly misleading. Technology changes so swiftly that spending undue amounts of time, energy and effort in this domain may be a waste. Technology can support, enable and trigger change, but it is only through the will and skill of dedicated people that solutions can be found to the health problems we wish to address. A strong collaborative evidence base that is coordinated at an inter-country level by key stakeholders in the m-health

community would reduce wasteful expenditure on programmes of showing little evidence of effectiveness and would instead promote best-practice models.

## Conclusion

M-health technology is exciting and attractive, but the real challenge lies in establishing country-level best practices that are both cost-effective and supported by rigorous research and evaluation. Policy-makers and funders must promote, legislate and fund programmes and interventions that integrate and build upon a common m-health framework. Fostering a synchronized, strategic global research agenda that is focused more heavily on evaluating the evidence for, and the impact and cost-effectiveness of, m-health services is imperative. International agencies like WHO and major funders of health-care research can be sources of crucial support to international coordination efforts and thereby contribute to high-quality health care for all. ■

**Funding:** Mark Tomlinson is supported by grants from the National Institute of Alcoholism and Alcohol Abuse (USA), the National Institute on Drug Abuse (USA), the National Research Foundation, South Africa, and the Department for International Development (DFID), United Kingdom. The opinions expressed in this article are not necessarily those of the NIAAA, NIDA, NRF or DFID.

**Competing interests:** None declared.

## References

1. WHO Global Observatory for eHealth. *New horizons for health through mobile technologies*. Geneva: World Health Organization; 2011:112.
2. *The world in 2011: ICT facts and figures*. Geneva: International Telecommunications Union; 2011.
3. Estrin D, Sim I. Open mHealth Architecture: An Engine for Health Care Innovation. *Science* 2010;330:759–60. doi:10.1126/science.1196187 PMID:21051617
4. Jordan ET, Ray EM, Johnson P, Evans WD. Text4Baby: using text messaging to improve maternal and newborn health. *Nurs Womens Health* 2011;15:206–12. doi:10.1111/j.1751-486X.2011.01635.x PMID:21672170
5. Vaughan C. San Diego researchers first to report positive impact of Text4Baby program (press release). San Marcos California State University, Office of Communications. 2011. 1 November. Available from: <http://www.text4baby.org/index.php/news/180-sdpressrelease> [accessed 16 February 2012].
6. Tomlinson M. Family-centred HIV interventions: lessons from the field of parental depression. *J Int AIDS Soc* 2010;13(Suppl 2):S9. doi:10.1186/1758-2652-13-S2-S9 PMID:20573291
7. De Maeseneer J, Roberts RG, Demarzo M, Heath I, Sewankambo N, Kidd MR et al. Tackling NCDs: a different approach is needed. *Lancet* 2011;373:11–2. PMID:21899880
8. Breslau DN, Maamari RN, Switz NA, Lam WA, Fletcher DA. Mobile phone based clinical microscopy for global health applications. *PLoS ONE* 2009;4:e6320. doi:10.1371/journal.pone.0006320 PMID:19623251
9. Orwat C, Graefe A, Faulwasser T. Towards pervasive computing in health care – a literature review. *BMC Med Inform Decis Mak* 2008;8:26. doi:10.1186/1472-6947-8-26 PMID:18565221
10. Lejbkiewicz I, Paperna T, Stein N, Dishon S, Miller A. Internet usage by patients with multiple sclerosis: Implications to participatory medicine and personalized healthcare. *Mult Scler Int* 2010;20:640749. PMID:22096625
11. WHO Global Observatory for eHealth. *Legal frameworks eHealth*. Geneva: World Health Organization; 2012.
12. Paré G, Jaana M, Sicotte C. Systematic review of home telemonitoring for chronic diseases: the evidence base. *J Am Med Inform Assoc* 2007;14:269–77. doi:10.1197/jamia.M2270 PMID:17329725

Adolescent health brief

## Using Mobile Phones for Adolescent Research in Low and Middle Income Countries: Preliminary Findings From the Birth to Twenty Cohort, South Africa

Alastair C. van Heerden, M.Sc.<sup>a,\*</sup>, Shane A. Norris, Ph.D.<sup>b,c</sup>, and Linda M. Richter, Ph.D.<sup>a</sup>

<sup>a</sup>Child Family Youth and Social Development, Human Sciences Research Council, Pietermaritzburg, South Africa

<sup>b</sup>Birth to Twenty Research Programme, Department of Paediatrics, University of the Witwatersrand Johannesburg, South Africa

<sup>c</sup>Department of Paediatrics, University of Cambridge, Cambridge, UK

Manuscript received January 8, 2009; manuscript accepted September 30, 2009

**Abstract** Mobile phones enable engagement with adolescents through a familiar medium. Survey data are presented from 2,023 South African adolescents who were asked about phone ownership, usage, and their willingness to divulge sensitive information in short message service surveys. Barriers to participation are addressed as are recommendations for follow-up research. © 2010 Society for Adolescent Medicine. All rights reserved.

**Keywords:** Mobile phone; Data collection

From fixed line telephones as an interview medium in the 1950s to more recent computer-aided surveys, technological innovation brings new opportunities. Mobile phones (MPs) offer potential benefits for research data collection, particularly surveys aimed at adolescents. The International Telecommunications Union reports that <1 billion new subscriptions were added during 2008, bringing the worldwide total to 4 billion; approximately one-third of these in Brazil, Russia, India, and China [1].

SexINFO [2] was the first mobile service in the United States offering sexual health information directly to adolescents. In response to an opt-in text message, a menu of questions is returned with instructions to text, for example, “D2 if u think ur pregnant.” A similar concept allows youth in South Africa to locate the nearest HIV testing centre by texting “HIV” and their town name to a 5-digit number [3]. A Cape Town-based study piloted the use of MPs as a means for sending daily treatment reminders to tuberculosis patients on a DOTS (directly observed treatment, short-course) program. This decreased regular workload of

DOTS employees and gave them sufficient time to focus on noncompliant patients. Of the 300 patients on the program, the World Health Organization reports only five treatment failures and now recommend this model as one of best practice [4]. These examples illustrate the potential of mobile devices. We investigated the viability of delivering self-administered surveys to adolescents on their MPs.

### Methods

Birth to Twenty (BT20) is a birth cohort study based in Soweto-Johannesburg; 3,273 newborns and their mothers were enrolled in 1990 to track health and development over the first 20 years of life [5]. Data are collected every 6 months through face-to-face and computer-assisted self-interviews. Plans are underway to introduce a MP data collection platform. The rationale includes increasing the ecological validity and improving the reliability and accuracy of sensitive self-report data. A telephonic survey to determine acceptability and feasibility of using MPs was conducted with BT20 participants, 2,023 of whom responded (88% of current participants; 52% female; 79% black). Each interview was approximately 5 minutes long, with the major reason for nonparticipation being unavailability at the time of call. The survey was completed by 2,023 participants (88% of current participants, 52% female, 79% black). Despite adolescents were mostly

\*Address correspondence to: Alastair C. van Heerden, M.Sc., Child Family Youth and Social Development, Human Sciences Research Council, Pietermaritzburg, South Africa.

E-mail address: [avanheerden@hsr.ac.za](mailto:avanheerden@hsr.ac.za)

Table 1  
Importance of mobile phone feature by gender

|                                    | Males                    | Females                  |
|------------------------------------|--------------------------|--------------------------|
| Making voice calls                 | 284 (64.3%)              | 280 (60.1%)              |
| Sending SMSs                       | 294 (66.5%)              | 313 (67.0%)              |
| Sending MMS                        | 228 (51.7%)              | 252 (54.1%)              |
| Listening to music                 | 299 (67.8%) <sup>a</sup> | 350 (74.9%) <sup>b</sup> |
| Having a camera                    | 274 (62.3%)              | 303 (65.0%)              |
| Is able to connect to the Internet | 224 (50.9%)              | 239 (51.3%)              |
| Being able to do mobile banking    | 50 (11.4%)               | 53 (11.4%)               |

<sup>a</sup> Significantly less than expected ( $p < .05$ ).

<sup>b</sup> Significantly more than expected ( $p < .05$ ).

living in households that had electricity (79%), indoor running water (69%), and indoor sanitation (54%), these households were considered poor relative to high-income country households. Questions were designed to assess the current perceptions and attitudes of BT20 youth to MP ownership and use in preparation for research use. The survey, an integral part of BT20, received human subject research approval from the University of the Witwatersrand. Pearson chi-square tests, frequency tables, and other descriptive techniques were used to analyze the data.

## Results

Among respondents, 69% reported currently owning a MP. By contrast, only 43% had access to a landline at home. Approximately 66% of the 31% reported feeling socially isolated as a result of not owning their own MP.

No gender difference was found in how adolescents choose their phone; 60% selected on functionality, followed by 38% who made the decision on the basis of styling. Only 8% chose a phone primarily on the basis of price. There were significant differences ( $\chi^2 = 62.26$ ,  $df = 3$ ,  $p < .01$ ) in phone ownership by ethnic group. Significantly, more white adolescents (95%) owned a mobile as compared with black (67%) and colored (63%) adolescents. Most of the participants (86.3%) reported being on a pay-as-you go plan. Of those on pay-as-you go, 62.1% paid for their own airtime and the remaining one-third received airtime from their parents.

Table 1 provides a summary of most valued MP features. There are few clear gender differences; females placed high value on an MP being able to play music. The most used MP features were listening to music, texting, and taking photographs.

When asked which medium they would prefer to answer sensitive questions, 73.1% chose face-to-face interviews over MP surveys. However, Table 2 highlights an interesting pattern between interview preference and knowledge and use of telecommunication technologies. There was a significant likelihood that those who had used Mxit, a popular mobile instant messaging application, or knew or used Skype, a voice-over-IP application, would prefer a MP questionnaire over the conventional face-to-face interview.

## Discussion

Three main concerns emerged regarding using MPs to collect data. These were: (a) not having airtime with which to respond, (b) lacking confidence about understanding the questions being asked, and (c) issues of confidentiality about who might see their answers.

Despite the advantages of speed, event proximity, and logistical ease, it is necessary to address these issues. First, airtime can be considered whenever surveys are administered. Second, questions must be worded clearly with detailed instructions prefacing each section. Finally, a methodology that ensures participant privacy and allows for consent to be obtained using the handset is imperative. Although not relevant to this cohort, the practice of handset sharing could affect negatively on privacy. It is not possible to password protect an SMS. Therefore, if information is sent by the researcher using this protocol it may compromise on the privacy of the participants. This challenge may be largely resolved by mobile survey software that can be password protected and by encrypting data.

Knowing that adolescent's value multimedia content on their MPs raises the possibility of using popular songs, ringtones, and screensaver downloads as an incentive for completing a survey.

Although low literacy rates and language barriers can be a challenge for self-completed questionnaires, providing

Table 2  
Interview preference and technological affinity

|                            | Response | Mobile phone questionnaire | Face-to-face interview   |
|----------------------------|----------|----------------------------|--------------------------|
| Do you know what Mxit is?  | No       | 43 (13.0%)                 | 119 (13.3%)              |
|                            | Yes      | 287 (87.0%)                | 777 (86.7%)              |
| Have you ever used Mxit?   | No       | 110 (36.9%) <sup>a</sup>   | 358 (43.9%) <sup>b</sup> |
|                            | Yes      | 188 (63.1%) <sup>b</sup>   | 457 (56.1%) <sup>a</sup> |
| Do you know what Skype is? | No       | 275 (83.8%) <sup>a</sup>   | 801 (89.9%) <sup>b</sup> |
|                            | Yes      | 53 (16.2%) <sup>b</sup>    | 90 (10.1%) <sup>a</sup>  |
| Have you ever used Skype?  | No       | 110 (84.0%) <sup>a</sup>   | 300 (91.7%) <sup>b</sup> |
|                            | Yes      | 21 (16.0%) <sup>b</sup>    | 27 (8.3%) <sup>a</sup>   |

<sup>a</sup> Significantly more than expected ( $p < .05$ ).

<sup>b</sup> Significantly more than expected ( $p < .05$ ).

the survey in colloquial languages is technologically simple on MPs. Each participant could receive the survey in their language of choice.

MPs provide an important platform by which to engage adolescents. The devices have penetrated even the most rural villages, are highly valued by adolescents [6], and offer instantaneous bidirectional communication. How to best scale these projects, manage informed consent, and maintain privacy are all issues which need urgent attention as technological advancements continue to push the bounds of what is currently possible.

#### Acknowledgments

The main funding used to support this research was a grant from the Wellcome trust.

#### References

- [1] Acharya S. Worldwide mobile cellular subscribers to reach 4 billion mark late 2008. Geneva: International Telecommunication Union, 2008.
- [2] Levine D, McCright J, Dobkin L, Woodruff AJ, Klausner JD. SEX-INFO: A sexual health text messaging service for San Francisco youth. *Am J Public Health* 2008;98:393–5.
- [3] Turrettini E. SMS to find out your HIV status. Available at: <http://www.textually.org/textually/archives/2007/11/018155.htm>. Accessed November 28, 2008.
- [4] World Health Organization. South Africa: A novel approach to improving adherence to TB treatment. *Essential drugs monitor* 2003; 33:8.
- [5] Richter L, Norris S, Pettifor J, Yach D, Cameron N. Cohort Profile: Mandela's children: The 1990 birth to Twenty study in South Africa. *Int J Epidemiol* 2007;36:504–11.
- [6] Wei R, Lo V. Staying connected while on the move: Cell phone use and social connectedness. *New Media Society* 2006;8:53–72.