Review of mVAM programme: novel application of mobile technologies for food security monitoring

DISI – Development Information Services International

August 2015

Nancy Mock & Nathan Morrow, DISI Review Team Co-Leads

Adam Papendieck, Sofia Curdumi Pendley & Margaret Hudson, Team Members

Table of Contents

List of Acronyms

AMIS	Agriculture Market Information System
CATI	Computer Assisted Telephone Interviews
CO	Country Office
CoP	Communities of Practice
CSI	Coping Strategy Index
DISI	Development Information Services International
DRC	Democratic Republic of the Congo
EVD	Ebola Virus Disease
FCS	Food Consumption Score
FEWSNET	Famine Early Warning System Network
HDX	Humanitarian Data Exchange
HIF	Humanitarian Innovation Fund
ICT	Information and communications technology
IDP	Internally Displaced Person
IGC	International Growth Center
IPC	Integrated Phase Classification
IVR	Interactive Voice Response
L2L	Listening to LAC
LAC	Latin America and the Caribbean
MOOC	Massive Online Open Course
mVAM	mobile Vulnerability Analysis and Mapping
PAPI	Pencil and Paper Interviewing
rCSI	reduced Coping Strategies Index
SES	Socioeconomic status
SIM	Subscriber Identity/Identification Module
SMS	Short Messaging System
ToC	Theory of Change
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
USSD	Unstructured Supplementary Service Data
VAM	Vulnerability Analysis and Mapping
WFP	World Food Programme

Acknowledgements

Executive Summary

Mobile Vulnerability Analysis and Mapping, or mVAM, is an evolving and expanding programme of the World Food Programme (WFP) that leverages cell phone technologies to conduct high frequency data collection and monitoring. mVAM represents a proof of concept for the application of mobile technologies to food security monitoring, particularly in contexts characterized by limited access for face-to-face interviews. mVAM has great potential to impact the way in which monitoring data is collected in humanitarian and development contexts. This review was commissioned by the Vulnerability Analysis and Mapping (VAM) unit of WFP to articulate strategic learning that may inform future deployments and scale-up of activities, as well as to document the experiences of applying the mVAM approach in two pilots and a full-scale deployment.

What is the value added of mVAM to VAM face-to-face food security and vulnerability assessment and monitoring work and existing data products?

mVAM is a very successful prototype. mVAM was able to provide data that was used in a variety of applications even as the pilot initiative was limited in scope and focused on a modest number of indicators. In the rapidly evolving context of Ebola-affected countries, the mVAM deployment was able to rapidly gather relevant data that informed WFP's strategy for food assistance. mVAM is a dynamic and flexible programme that contributes to the food security monitoring Community of Practice (CoP) knowledge and learning; providing timely, open data; producing timely information products in difficult settings; and builds upon best practices. mVAM also demonstrates that inexpensive data collection in difficult contexts is possible and that high frequency monitoring can be efficiently achieved.

What might be the future direction of mVAM, including scale-up and adaptation to WFP information needs?

mVAM has innovative potential that it should leverage, document and expand upon by developing a Learning Lab. The research, curation and capacity development based on programme knowledge and learning is perhaps the greatest potential role of mVAM at the global level. Through strategic research, experimentation and prioritization, mVAM can greatly expand and document best practice approaches to the application of mobile technology to food security monitoring.

mVAM must develop a clear mission and vision as well as a robust Theory of Change (ToC). Additional recommendations include that mVAM:

- Should engage in a serious and ongoing dialogue with information users;
- Should balance technology advantages with true beneficiary participation and include safeguard triggers for face-to-face assessments or conditions under which face-to-face interviews are necessary when possible;
- Should continue to increase the relevance of monitoring information as information needs change in the dynamic contexts where WFP works;
- Must mainstream privacy and protection commitments;
- Needs to continue to grow and evolve its strategic partnerships;
- Should continue to pursue its organic development approach during its expansion phase;
- Should embrace now its role in capacity building for mobile monitoring both within and outside of WFP; and
- Should consider features of study design that lead to high-quality data.

Introduction

History of the Programme

The World Food Programme (WFP) has long had an interest in and commitment to measuring and using information on the food security and vulnerability status of populations at risk of food insecurity. WFP's Vulnerability Analysis and Mapping (VAM) unit was established in 1994. Like the United States Agency for International Development (USAID) Famine Early Warning System Network (FEWS NET), from its inception VAM has been committed to leveraging information and communications technologies (ICT) to enhance the coverage and timeliness of information intended to prevent and mitigate food related emergencies. Recently, the explosive global growth of cellular and Internet technologies has served as an impetus for experimentation with these technologies for food security monitoring, especially in contexts with restricted humanitarian access.

In 2013, VAM established a programme called mobile Vulnerability Analysis and Mapping, or mVAM, as an umbrella programme through which to explore the use of cell phone technologies in the context of monitoring in unstable contexts. The increasing susceptibility of some chronically vulnerable populations, and the emergence of new global health threats such as Ebola Virus Disease (EVD) and conflicts (instability in the Middle East, North Africa and parts of the Sahel), provided incentives as well as opportunities to test new modalities for data collection and monitoring. Initial pilot trials took place in eastern Democratic Republic of the Congo (DRC) and Somalia; while mVAM launched a full deployment of cellular monitoring in EVD-affected countries and more recently in Iraq. In January of 2015, VAM contracted Development Information Services International (DISI) to undertake a review of the mVAM programme for the purposes of organizational learning and accountability to its funders.
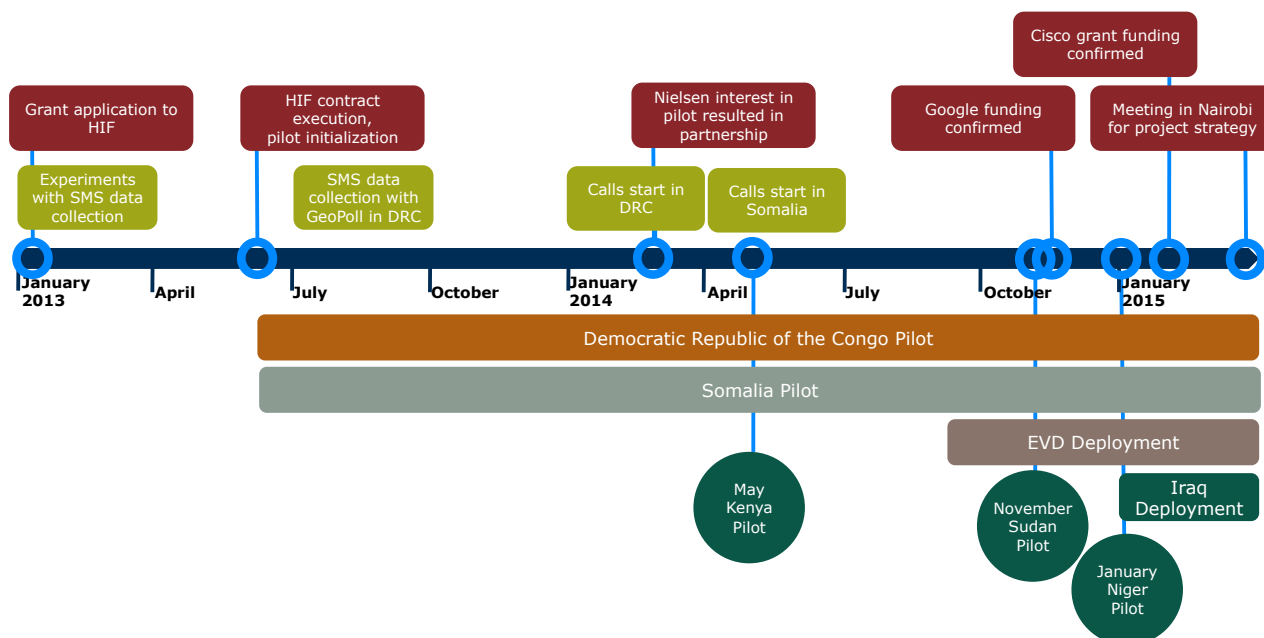

Figure 1: mVAM timeline

Objectives of the Review

The objectives of this review include to:

- Present research that further documents and illustrates the need as well as relevance of the mVAM programme and analyzes mVAM's role in promoting/applying emerging data collection modalities;
- Assess the added value of remote data collection to WFP's programming and field activities;[1]
- Identify opportunities to adjust and improve mVAM's methods;
- Provide guidance to the mVAM team to apply the learning immediately to on-going and future assessment and monitoring activities;[2] and
- Investigate the desirability of, and possibilities for, scaling up the programme while highlighting the challenges this may bring.

The review team was guided by two overarching research questions:

1. What is the value mVAM adds to VAM face-to-face food security and vulnerability assessment and monitoring work and existing data products?
2. What might be the future direction of mVAM, including scale-up and adaptation to WFP information needs?

This systematic review of the mVAM pilots and the first full mVAM deployment in EVD-affected countries intends to: (1) document these experiences, and (2) articulate strategic learning that may inform future deployments and scale-up of activities. The findings and recommendations are intended to inform immediate mVAM priorities, potential scope of activities and strategic direction for the mVAM team and partners. The review is not a formal evaluation, nor will we dwell on detailed methodological debates that would distract from the larger lessons of strategic importance. Technology partners such as the Nielsen Corporation in addition to internal analysts have done an adequate job of pointing out issues such as recall period or comparing the response rates of different modalities. References to this research are made where appropriate in this review. The quality of the data, the scope of field experiments and the potential value-added from relatively small incremental learning about techniques do not merit addition reprocessing as part of this review. This review is focused on the need for more rigorous experimental designs and a robust research agenda for the development of mVAM.

The structure of the document is to first present a reconstructed Theory of Change (ToC) based on interviews, a review of literature and project documentation. The ToC is meant to provide clarity in a visual form on the purpose and key relationships in the mVAM pilots. The ToC was refined through discussion with the mVAM team and can guide the design of future research, evaluation and strategic planning. The aim of the following section is to summarize what is known about error and bias related to panel studies, cell phone surveys and high frequency cell-phone based monitoring via a review of the literature. Narrative case studies largely based on mVAM blog posts are included as an annex and serve

---

[1] The report will specifically focus on the work related to the support received from the Humanitarian Innovation Fund (HIF) as well as in-kind contributions from other collaborators.

[2] To this end specifically, the DISI assessment team will develop a business case that compares different approaches/technologies available with specific recommendations for incremental improvements in tools, frameworks, and analysis techniques.

as the basis for constructing timelines and analyses of these cases referenced in the section on 'Findings'. The 'Findings' section summarizes conclusions based on triangulation of three main sources: the expert opinion of the reviewers, the input those interviewed during the review process, and in reference to issues identified in the literature. These conclusions are presented in two sets: 'Topline Findings' that are strategically relevant; and 'Specific Findings' that are meant to be helpful for making small adjustments to current and upcoming deployments.

Methodology

This review utilized multiple methods and the convergence-of-evidence approach to assess the value-added of mVAM and to develop recommendations for an mVAM strategic vision going forward. **Table 1**, below, summarizes the evidence base of this review.

Table 1: Evidence base

Evidence source	Number	Comments
Literature on ICT applications for international development/ humanitarian work, surveys and panel data, and specifically the use of mobile technology for survey implementation	35	Peer reviewed literature evaluating new methods is only emergent; very few rigorous studies exist
Expert key informants	16	General recognition that early experiments are very promising but stronger evidence base needed
WFP mVAM staff	7 mVAM Staff 5 WFP CO Staff	Substantial internal enthusiasm around the mVAM projects and internal demand within
Case studies	DRC, Somalia, EVD countries	mVAM and VAM data sets, bulletins, blog posts and interviews; one site visit (DRC); validity analyses conducted by technology partners
Project document review	84	Project documents were largely internal documents, blogs and presentations
Panel participant community forums and small group interviews	4	These were conducted in DRC only and included a forum meeting with pilot respondents as well as small group interviews with camp management, the volunteer support committee and the cooperative managing the solar panel
Concept mapping and brainstorming	Rome, New Orleans and Goma team meetings	Critical activity with WFP staff to determine boundary, SWOT of mVAM and the ToC model

The review team reviewed literature relevant to assessing the validity and reliability of panel data and of the use of mobile technologies specifically for estimating population parameters and monitoring trends through panel study approaches. We also reviewed the grey literature related to other mobile technology applications in the field of poverty assessment, health monitoring and beneficiary monitoring in multisectoral programmes (UNICEF uREPORT). Through the literature we reviewed sampling error, respondent bias and reliability of projects attempting to collect panel data using mobile technologies. The review team also interviewed technical experts who had been involved in field trials of mobile technologies for monitoring purposes.

The review team assembled three detailed case studies from the Democratic Republic of Congo (DRC) pilot, the Somalia pilot and the implementation of mobile monitoring in the three countries heavily affected by Ebola Virus Disease (EVD): Guinea, Sierra Leone, and Liberia. Case study materials included blog posts provided by mVAM, project documents, bulletins, staff interviews and beneficiary interviews in the case of the DRC. The three case studies illustrated diversity in contexts; scope of population to monitor/sampling frames; sampling strategy; modes of cell phone data collection; and information products. In the case of the EVD case study, a broad geographic area was covered that included both WFP beneficiaries and non-beneficiaries. In the case of the DRC, the target for monitoring was a displacement camp and in the case of Somalia, high frequency monitoring involved Internally Displaced Persons (IDPs) spread across a large urban area. Both pilot sites included WFP beneficiaries and non-beneficiaries. In Somalia, the beneficiary/non-beneficiary distinction was not considered relevant because WFP has very limited programming in the context. The review team synthesized findings to focus on the issues of validity, effectiveness, efficiency, and relevance of information produced by mVAM under the rubrics of "use cases" and "technical adequacy" as well as lessons learned for future mobile monitoring.

The review team conducted numerous brainstorming sessions with mVAM staff, including during a three day meeting convened in New Orleans where the DISI review team and the mVAM programme staff team completed a draft timeline of mVAM activities and synthesized relevant findings from the literature, case studies and interviews.

Nathan Morrow and Nancy Mock co-led the review team. They conducted all stakeholder interviews. Nathan Morrow interviewed WFP staff and case study informants, including a variety of external partners and the WFP beneficiaries that served as mVAM pilot participants in the DRC. Nancy Mock led expert interviews and participated in several interviews with mVAM staff and partners. In several cases, Mock and Morrow interviewed jointly. All literature was entered in to a Zotero bibliographic database and all documents used in the analysis are stored in Google Drive and shared with mVAM staff. Adam Papendieck conducted a review of the literature on mobile monitoring; Sofia Curdumi Pendley and Margaret Hudson drafted the case study summaries and Margaret Hudson served as the knowledge management lead for team materials and led the graphics and final document production efforts. Phil Anglewicz served as a subject matter expert in the area of panel study error.

Towards a Theory of Change for mVAM in the Context of a Literature Review

Very important to the mVAM review, as in any evaluative activity, is the formulation of a mutually agreed upon Theory of Change (ToC) that lays out the pathways through which mVAM may contribute to WFP's core mission of improving food security in vulnerable areas of the world. We build from the literature and brainstorming activities with the mVAM staff to develop a rudimentary ToC. **Figure 2**, below, is the review team's suggested ToC.
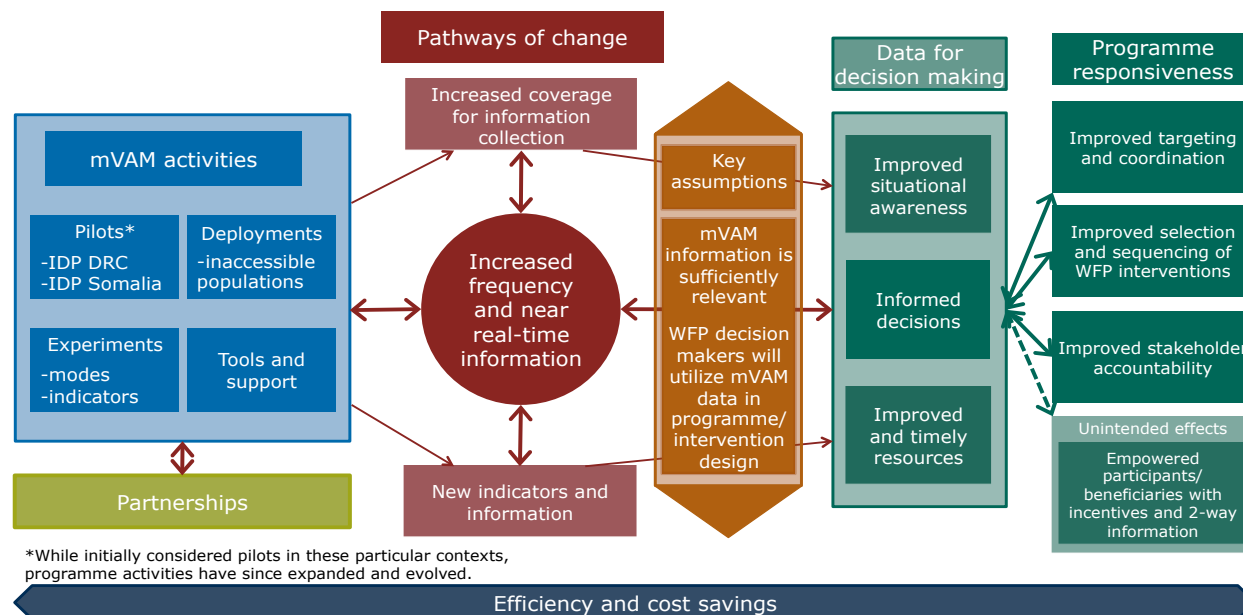
Figure 2: mVAM Theory of Change

mVAM is an innovation center for the application of cellular phone technologies to WFP food security monitoring work. To the left of the diagram, mVAM activities consist of those projects that began as pilots (DRC and Somalia), deployments (EVD countries), and experiments/studies as well as the provision of technology tools and support to field activities. Integral to mVAM is partnership with donors/sponsors and technology organizations (**Figure 3**).

Figure 3: Partnerships

Partnerships fall into the following five categories: (1) Donors and sponsors (HIF, USAID, Cisco, Google, Government of South Korea); (2) Private providers of technology services (GeoPoll, InSTEDD); (3) Community of Practice (CoP) partners (UN Global Pulse, World Bank Listening to Africa project); (4) Corporate voluntary services (Nielsen); and (5) Internal partners. Internal partnerships within WFP include with the VAM Markets team, particularly in the development of EVD information products, the Country Offices and the Monitoring and Evaluation unit of WFP.

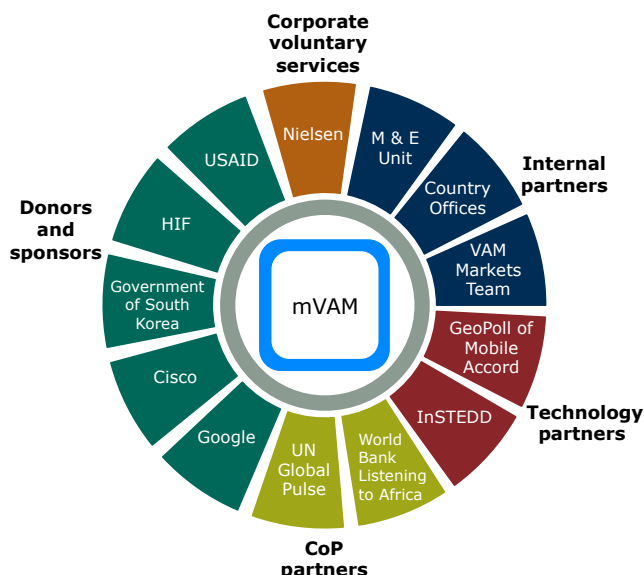mVAM intends to leverage the recent high use and coverage of cell phones in vulnerable areas to improve the coverage, frequency and informational content of food security status reports of these populations, and in this way increasing the temporal and spatial scale of real-time and relevant information. Cell phones can increase coverage in areas where it may be difficult or impossible to employ enumerator teams. Areas affected by EVD and intense conflict are two examples where information on food security status is critical to decision-making, but where conducting face-to-face surveys may expose enumerators and local populations to health and security risks. Similarly, these contexts and other areas where WFP works are often characterized by unstable food security. High frequency monitoring, especially around shocks and stresses, may be particularly valuable to WFP operations, donors and other stakeholders for determining priority and future programme needs. High frequency data provided through cell phones might enable WFP to have near real-time measures of food security status of populations of concern, an important pathway of change.

A final pathway of change is experimentation, lesson learning and building an evidence base around the indicators (e.g. formulations of the Food Consumption Score (FCS)) and modalities of data collection for food security monitoring. mVAM is experimenting with alternative formulations of the FCS, such as the reduced FCS, as well as with different modes of cell phone data capture including computer assisted telephone interviews (CATI, live interviews), Short Messaging System (SMS) and Interactive Voice Response (IVR) in different contexts.

Key assumptions of the mVAM programme relate to the relevance and use of mVAM information products and services. Relevance means that mVAM information is reasonably and sufficiently relevant to influence programmatic decision-making in the contexts where they are applied. Secondly, an assumption is that information provided will influence decision-making. Most decision makers utilize multiple indicators and information sources (and consider a number of other contextual and political factors) when deciding on appropriate interventions (Morrow, Mock and Papendieck, 2010). Factors such as the adequacy of indicators, presentation style, format and timeliness of information products are all key determinants of the relevance of information for decision-making. Another important consideration is the validity and reliability of the data collected by mVAM. The project uses a panel design whereby selected participants are monitored over time. **Box 1** shows the potential threats to validity posed by this type of design.

One interesting unintended effect of mVAM is the empowerment of participants/beneficiaries with unplanned incentives for cell phone use (e.g. reunification, remittances, and safety planning) and bi-directional information flow between cell phone operators and beneficiaries. It may be worth seriously considering such consequences and their potential ramifications.

mVAM Landscape: putting the pilots and early deployments in perspective with a rapid review of key issues for panel and mobile data collection found in the literature

Longitudinal panel data offers many advantages over cross-sectional approaches, such as the ability to examine change at the individual level over time. Longitudinal data also provide opportunities for useful analysis of data quality that cannot be attained with cross-sectional study designs. At the same time, longitudinal studies are threatened by important biases that can affect internal and external validity. Important sources of bias include sample attrition, response accuracy and testing bias.

Box 1: Potential Biases and Approaches to Address

Attrition: All longitudinal data collection projects face the inherent problem of sample attrition: the failure to find or re-interview individuals who were surveyed in an earlier wave of the study. In rural sub-Saharan Africa rates of attrition are particularly high (Anglewicz et al., 2009). Attrition poses two important potential problems for longitudinal research. First, attrition reduces sample size and thus statistical power. More importantly, individuals who leave a study are often different from those who remain in many key characteristics; thus attrition can threaten the external validity of the study sample (Alderman et al., 2001; Fitzgerald, Gottschalk, and Moffitt, 1998; Thomas, Frankenberg, and Smith, 2000).

- The greater the attrition, the greater the likelihood that response bias is also present (see below). Generally, a minimum sample size of 300 is desirable in research; attrition reducing the sample size below 300 is potential cause for concern. The exact sample size necessary, however, depends on the relationships of interest; detecting differences between groups that are relatively small requires a larger sample size.
- There are no "rules of thumb" for the approximate percentage of attrition that affects the generalizability of the sample: this bias depends on the extent to which "attritors" are different from those who remain, which varies across settings. In some cases, attrition bias may be minimal and loss of respondents doesn't threaten external validity. In other cases, even

---

relatively small amounts of attrition can bias the sample if those who leave the study are substantially different from remaining respondents. The extent of attrition bias can, and should be, formally tested.

- Attrition bias can be (and is perhaps best) addressed in advance. Collecting tracking information for respondents in the first wave of data collection can reduce loss to follow-up: collecting cell phone numbers, family contact information, and intentions to move can reduce the number of respondents who aren't found in later waves of data collection.
- Reducing attrition bias can be particularly challenging for displaced populations, who are often more mobile than other populations, with less predictability in their movements (and who may not have reliable contact information). For these groups, maintaining contact with friends and family members can facilitate later tracing. In addition, individuals typically participate if they see some benefit of a study for their interests. Offering incentives for individuals to participate facilitates participation in future wave. Particularly mobile populations may be at greater need for incentives and therefore more responsive.
- In addition, in mobile populations, one can use a hybrid approach that combines (1) tracking respondents leaving the study site, and (2) continuing enrolment of new study participants. In a refugee camp, for example, this approach can be useful to prevent sample size reduction, and test whether new members are similar or different from those leaving.
- If, however, attrition bias appears to be present, statistical methods are available to correct potential bias. Heckman procedures involve a two-step process to identify the extent of attrition bias and then correct for this bias (Heckman, 1979).

Response accuracy: As with cross-sectional studies, longitudinal studies are also vulnerable to inaccurate self-reports. Research has shown that self-reported behaviours are often reported inaccurately, and causes for these invalid reports can often be traced to interviewer effects, social desirability bias, and/or recall bias. Testing for and improving response validity is a challenge for many areas of research.

- Research has long shown that the survey method can affect the validity of responses. Technological developments have permitted surveys to be conducted by computers or over cell phones. But are responses affected by computer or cell phone interviews compared to the traditional face-to-face methods? While some research suggests that computer and cell phone interviews can improve validity, results for the effect of survey methods are not consistent across studies.
- Due to this lack of consistency across studies, testing the effect of survey method on response patterns is an important task for formative research. A common approach is to randomly assign respondents to various survey modes (e.g., phone vs. face-to-face), and the compare responses to sensitive questions between these methods. Significant differences in response patterns show that the mode of interview has an important impact on responses. Similarly, this can be done in longitudinal studies by comparing response patterns between two or more waves of data. Some variables, like age, ethnicity, and other characteristics should be either unchanging, or change predictably over time.
- But if differences between survey modes or responses over time are found, how do you know which response is correct? Measuring response reliability (consistency of responses) is much

easier than validity (the correspondence of self-reports to the "true" value): just because a response is consistent doesn't mean it is correct (both responses could be invalid).

- Measuring validity involves finding an "objective" measure that can be compared to the self-report. A potential approach is to ask an interviewer to observe and report on a household item, like ownership of a bicycle, latrine, etc… that can be compared to the respondent's self-report.

Testing bias: Testing bias refers to changes in what is being measured that are brought about by the reactions to the process of being measured. In a longitudinal study context, respondents may remember the survey and change their answers based on previous survey experience. For example, if individuals receive an incentive for a particular response to a survey, they may continue to respond in that way in order to continue to receive the incentive, even if the response should change (we refer to this as "gaming").

- The challenge with testing bias is to distinguish between natural change over time that is "true" and change over time that is due to respondents changing their responses over time due to learning about incentives. Ideally, one would find a way to validate self-reports of responses to questions that might be vulnerable to this problem.
- Identifying potential testing bias can sometimes be done using longitudinal data. In this approach, one can compare self-reports between waves of data collection to identify outliers, which may suggest misreporting. In other words, if the extent of change in an outcome is substantially greater for some participants than others, this may indicate systematic misreporting. A follow-up with such respondents can inform as to whether this change is valid.
- As with other challenges in data collection, testing bias can be addressed in advance of data collection by informing respondents that their self-reports will be validated. If study participants know that their self-reports will be validated, this might reduce the extent of misreporting among participants.

Mode and context considerations

Some or all of the threats to validity in **Box 1** may be of concern when collecting data with mobile technologies (Chakraborty et al., 2013; Croke et al., 2012; Dillman et al., 2009; Dillon, 2012; Gallup, 2012a; Gallup, 2012b; Kreuter et al., 2008; Lucci and Bhatkal, 2014; Prydz, 2014; Williams, 2014). Bias and error associated with particular data collection modes are termed "mode effects." Here we will look at considerations which must be made when choosing between three main modes of mobile data collection (CATI, IVR and SMS) in context, making comparisons between them as well as to more traditional face-to-face pencil and paper interviewing (PAPI) modes of data collection. Due to technical and practical similarities, discussion and recommendations related to SMS may generally be considered relevant to Unstructured Supplementary Service Data (USSD). mVAM compared the use of each of these modes except for the computer based CATI that is often used in institutional settings or over the Internet. A key differentiating aspect of mode is whether the survey is self-administered (SMS/USSD, IVR) or enumerator administered as with CATI or PAPI.

Survey design characteristics

Mode effects may result from survey design characteristics such as question length, question order, number of choices, order of choices, and overall length of the survey (Williams, 2014). Recency bias is a general issue with telephone based surveys where respondents must listen to and recall their choice from memory, and such bias can be mitigated by keeping items and choices short and randomizing presentation of questions and responses. The social connection and rapport established between live interviewers and respondents has been cited as a possible factor which contributes to higher item and unit response rates observed for CATI relative to SMS and IVR, and may also enable the use of slightly longer surveys (Ballivian and Azevedo, 2013; Croke et al., 2012). That said, it is recommended that surveys using mobile phones not be longer than 20-30 minutes and probably less for IVR and SMS which have been found to have the highest rates of item non-response (e.g. due to skipping questions) and unit non-response over time (e.g. due to attrition) relative to CATI (Gallup, 2012b). IVR was associated with the highest attrition in the World Bank's Listening to LAC (L2L) project in Honduras, followed by SMS and CATI (Ballivian and Azevedo, 2013; Gallup, 2012b).

For SMS modes, questions and responses typically must be formulated to respect a maximum of 160 character limit on individual text messages which is largely standard among telecoms. Though uncommon, some telecoms enforce a 140 character limit by default, and for non-Latin scripts like Arabic the limit may be as low as 70 characters. In such cases it may be possible to negotiate an increase in the character limit in order to accommodate surveys designed for 160 character communication.

Participation, sampling and attrition
In general, inequalities in phone ownership are an important contextual consideration when deploying mobile data collection strategies (Lucci and Bhatkal, 2014). National registries of mobile subscribers are often unavailable or incomplete (Gallup, 2012a), but more importantly, phone ownership itself is associated with factors such as years of formal education, higher wealth, urban residence and male gender, with disparities being particularly high in Africa, the Middle East and South Asia (GSMA, 2010; Tortora et al., 2011). The fact that such factors are also correlates of food insecurity makes sample selection an extremely important potential source of bias.

Developing a sampling frame for mobile phone surveys can be challenging, and two different approaches have been employed. The first, which has been employed by both the World Bank poverty monitoring initiatives and WFP, is to utilize large scale baseline survey samples as the sampling frame for selecting and monitoring a panel. This method involves either including the entire sample willing to participate in monitoring waves or selecting a subsample of the baseline sample to monitor over time. In order to approximate a representative sample, a common practice is to distribute phones to all individuals/households in the sampling frame or to those that do not have them. There is evidence that while this may incentivize initial participation and help improve the representativeness of the sampled population (Ballivian and Azevedo, 2013), phone distribution does not itself have an effect on the higher level of attrition observed among lower socioeconomic status (SES) households, though modest cash or phone credit incentives may (Gallup, 2012b).

It is therefore important to note that dissemination of phones does not guarantee uniformity of use or reachability within an otherwise representative sample. Communities or groups characterized by relatively low income or rates of mobile penetration may have difficulty servicing and maintaining the devices which are given to them, resulting in relatively higher rates non-response due to unreachability

13

(Croke et al., 2012). Community variability in mobile use and literacy may also be an issue. In the Honduras L2L surveys, household follow-up reporting via SMS tended to "significantly underestimate facts regarding household infrastructure, while over-estimating Internet access and self-perceptions on poverty" relative to the face-to-face baseline. To explain this observed measurement error, it was hypothesized that phones distributed for household reporting via SMS tended to end up in the hands of younger household respondents inclined to report differently on such issues than the older heads of household typically surveyed face-to-face (Gallup, 2012a).

The second methodology employed for sample selection is to utilize cellular subscriber databases, though these are highly variable and difficult to access. Often a reliable database of all mobile subscribers is simply not available for the sampling frame. However, GeoPoll has become an important provider of sampling frames for mVAM and other clients. GeoPoll acquires cellular databases or attempts random digit dialling and provides cellular survey services for self-administered methods. GeoPoll also performs some assessment of the quality of the sampling frame that it provides in terms of coverage of service providers. In general, response rates for surveys that are dispensed to large databases of cellular subscribers are very low. Understanding the nature of bias introduced by the sample of respondents is essential.

Language and literacy

Language and literacy is perhaps one of the most critical and differentiating contextual considerations impacting choice of data collection mode. SMS is particularly vulnerable to response bias due to variation in literacy among respondents, resulting in underrepresentation of groups with low written and digital literacies. Again, this is significant as such groups are also often the most food insecure and vulnerable to shocks. A notable advantage of IVR and CATI over SMS is that they do not necessarily require the ability to read, write or produce text on a mobile device. The ability to leverage multilingual interviewers or a multilingual pool of interviewers to engage populations characterized by high variation in literacy and language is commonly cited as a critical advantage of CATI, justifying its relatively higher per-interview cost relative to IVR and SMS (Croke et al., 2012; Gallup, 2012a; Gallup 2012b).

Social desirability and conflict of interest

CATI has been found to be more susceptible to social desirability bias than self-administered modes like IVR and SMS (Williams, 2014). Respondents may be less likely to reveal sensitive information to live telephone interviewers than they would via self-administered survey (Kreuter et al., 2008). Especially when collecting particularly sensitive information, or working in socio-politically unstable or volatile contexts, CATI modes may be more vulnerable to validity threats resulting from *acquiescence*, where respondents provide answers which align with the perceived "positive" end of a scale, and *satisficing*, where respondents bias their responses to ensure agreement with what is socially desirable. Moreover, CATI holds better promise to safeguard protection in these contexts.

In highly vulnerable populations, deliberate misreporting for the purpose of increasing the likelihood of food distribution may be a particularly significant threat to the validity. This is a fundamental problem in food security assessment and shock monitoring, and one which may be exacerbated through the use of high frequency mobile strategies where the linkage between monitoring efforts and potential food aid may be more clearly perceived by respondents. Mitigation strategies for addressing this occurrence should be considered at the beginning of, as well as throughout, survey administration.

14

Device selection, maintenance and replacement

Contextual considerations must be made when choosing mobile devices for the field, as malfunctioning or lost phones may render respondents unreachable and potentially introduce bias. When mobile devices must be distributed in a population, it is important to ensure there is local capacity not only to use them, but also to repair and maintain them. Selecting devices which are already in circulation, and therefore potentially easier to have serviced locally, may help avoid non-response and attrition resulting from device failure.

Battery quality and battery maintenance has been reported as a particularly important consideration in areas which are not well integrated or prioritized on the electrical grid. In places with low rates of electricity in the home or frequent power outages, a reliable number of affordable charging stations must be available in order to ensure respondents are reachable, or individual means of charging must be ensured. Croke et al. (2012) report that among the panel members they surveyed, those with access to electricity answered on average 18.6 rounds versus 16.4 for those without access to electricity, and suggest that distribution of solar chargers would enhance overall response rates.

Whenever phones are disseminated for data collection, replacement parts and devices should be reserved, especially in areas where repair and maintenance may be complicated by lack of capacity or novelty of the device. Dillon et al. (2012) report that 8 percent of their respondents reported lost, damaged or malfunctioning phones over a nine and a half month survey. They also noted that the prospect of easy replacement might pose a "moral hazard" in the sense that respondents may be more likely to be careless or to sell a phone that is given to them, perhaps claiming that it was lost, if they believe it will be replaced. Respondents in the survey which they conducted were told that lost phones could not be replaced, only malfunctioning phones and batteries could be exchanged for new ones and if a respondent lost a project phone they were asked to continue participating in the survey using a personal phone (Dillon et al., 2012).

Network characteristics and quality

A variety of questions related to mobile networks themselves must be answered in context. Which telecoms work where? Which telecoms allow calls to multiple networks, and at what cost to the caller? How reliable are various networks? Are there rate limits (e.g. messages per minute), usage limits (e.g. characters per text message) or charges (e.g. for unanswered calls) that may affect participation? In cases where populations are migrating across borders, or where frequent border crossings are common, how will network connectivity and respondent reachability be affected?

The use of premium network providers has been found to affect the reachability of respondents; whereby respondents using the premium network tend to respond more often than those using other networks (Croke et al., 2012). In general, SMS is seen as the most reliable mode in places with poor telecom infrastructure or unreliable networks, as SMS communication can work asynchronously and is less affected by periodic service interruption.

Cost and efficiency

A cross cutting theme and assumption of mVAM is that cell phone data collection is highly efficient in terms of technical and financial costs and that data provided is valid and relevant. There are obvious

advantages to mobile modes of data collection in terms of efficiency, frequency and coverage. The World Bank L2L and Listening to Africa projects have reported significantly lower costs associated with mobile modes of data collection. For instance, in Honduras and Peru, mobile phone surveys cost about $8, $17 and $25 per interview via SMS, IVR and voice calls respectively as compared to $40 for a face-to-face survey in the same countries (Ballivian and Azevedo, 2013). Additionally, Croke et al. (2012) have reported that mobile phone surveys in Tanzania cost between $4.10 and $7.30 per interview as opposed to $50 to $150 for face-to-face. Mobile technology affords an increase in frequency and geographic coverage of data collection without necessarily adding to overall project costs. Whereas limited frequency of census and other interview-based assessment data has been a shortcoming of food and nutrition security information systems in many parts of the world, the efficiencies brought by emerging CATI, IVR and SMS techniques are making the collection of new kinds of high frequency panel data more feasible (Dillon, 2012).

The ToC posits that monitoring data from mVAM will substantially enhance decision support for WFP programming by providing critical information in near real time to decision makers on the food security status and changes in status over time, a well-recognized gap in food security monitoring (Mock et al., 2013). Decision makers will be more situationally aware, thus able to improve the timeliness of decisions and resource mobilization. This will ultimately result in improved targeting and sequencing of relevant WFP interventions. Finally, high frequency monitoring of vulnerable populations and beneficiaries offers the opportunity for WFP to more directly engage with client populations regarding their concerns, priorities and assessments of programme relevance.

Findings

Topline Finding: mVAM is a very successful prototype

mVAM represents a strong proof of concept for the use of mobile monitoring of food security status in a variety of settings and is particularly appropriate in contexts characterized by access limitations. With very limited resources, the mVAM team deployed pilot implementations and full-scale deployments/implementations that utilized several best practices. In doing so, mVAM demonstrated the application of novel technologies in relief/development monitoring. mVAM's proof of concept is perhaps best reflected in the demand for mVAM projects in new countries as well as the interest and engagement with mVAM by corporate sponsors and donors. Pilot data and mVAM data collection in EVD-affected countries was used for a variety of purposes that indicate potential impact of the mVAM approach.

Use of mVAM data demonstrates a broad set of applications even in consideration of the relatively modest number of indicators and scope of the pilot. An early use of the mVAM data by the Goma, DRC Area Office was to compare beneficiary and non-beneficiary consumption. WFP Country Office staff also used mVAM monitoring to show "impact of WFPs ration and overall vulnerability situation" following a distribution to the IDP camp. Finally, mVAM analysis was used in an "extrapolation" to other camps including those from another displacement event and outside Goma (IOM 2009 camps). The potential for similar food insecurity in these groups due to seasonal and market issues was a trigger for new round of surveys. In general, WFP staff interviewed in DRC were extremely supportive of the high frequency monitoring and creative in its application to programme decisions.
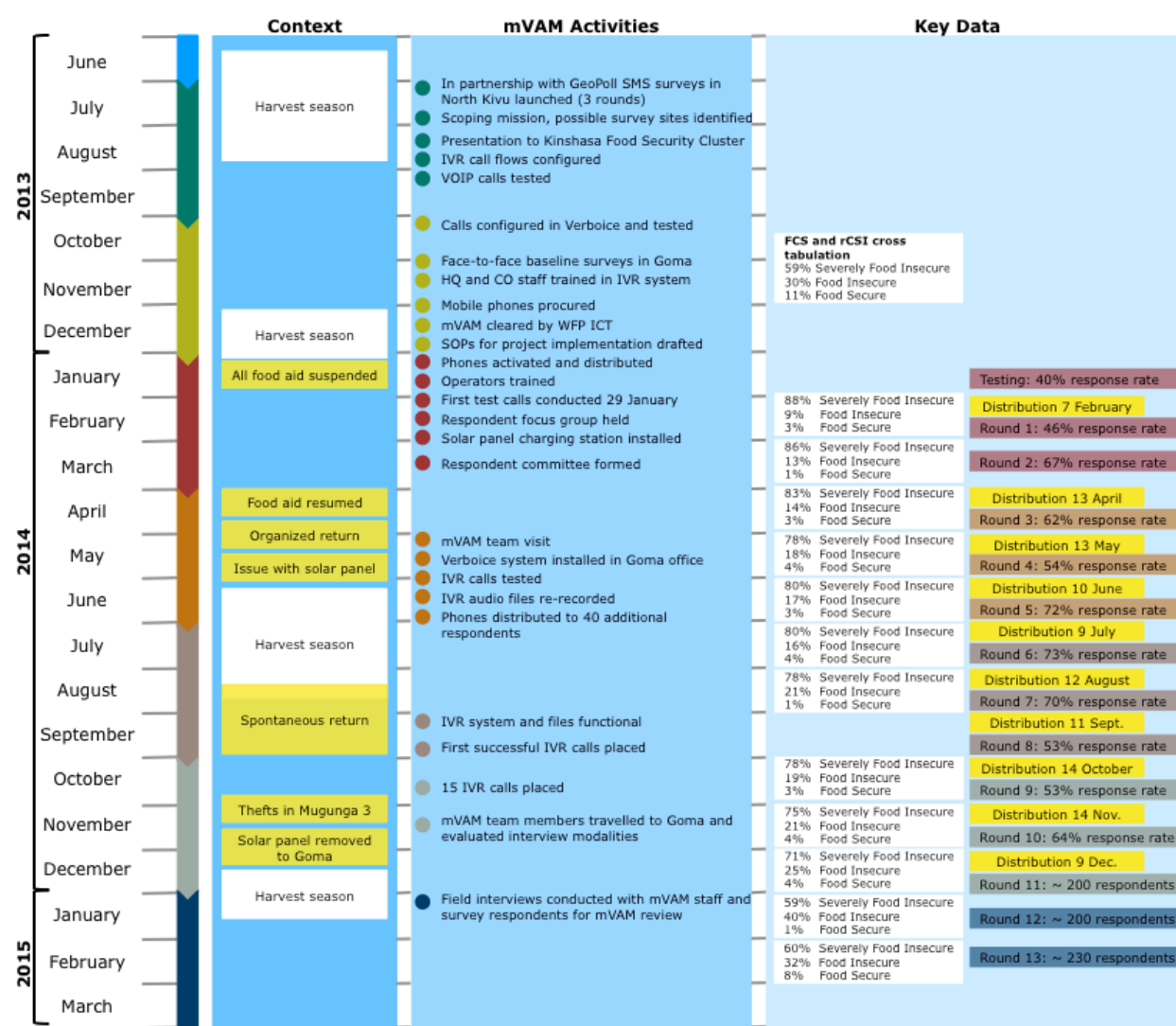
Figure 4: Goma, DRC timeline

Perhaps the most promising use of mVAM data took place in the analysis of prices, wages, food consumption and coping of those affected by the EVD crisis. mVAM data showed that food insecurity was deteriorating in rural areas that had comparatively low number of EVD cases compared to urban centers. Restricted access to work and markets impacted the food security of many rural households that were not in the areas receiving the most attention from the response based on EVD caseloads. This informed the strategy for food assistance during the response. The addition of prices and wages to the mVAM monitoring was used in analysis by a variety of other stakeholders. The EVD deployment was also very quick. mVAM data collection through SMS and IVR was up and running (and reporting) in about a month in three countries. This would not have been possible with WFP VAM legacy systems.
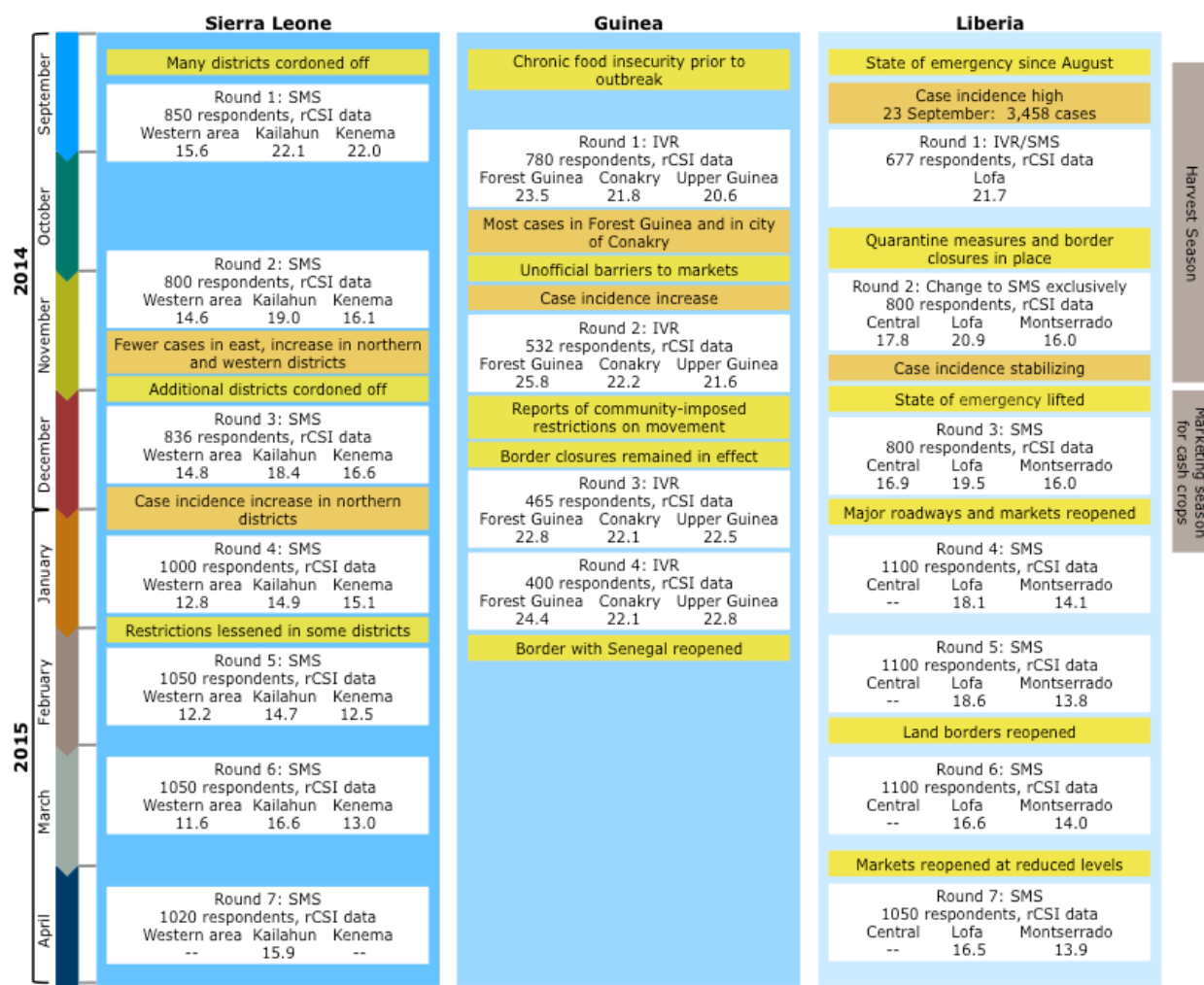
Figure 5: EVD-affected countries timeline

The Joint Research Center funded by the European Union conducted a detailed study of the food security effects of the EVD crisis and referred frequently to the mVAM monitoring data.[3] The International Growth Center that conducts research to inform development policy also used mVAM price and wage data in their analysis of the longer-term effects of the EVD crisis and was shared on the web at www.theigc.org/news-item/the-economic-impact-of-ebola-november-2014-report/.

mVAM is a dynamic and flexible programme within VAM that exemplifies agile design and adaptive management. mVAM has a highly energized, well networked and proactive collaborative team. Field operations also are adaptive, discovering and documenting lessons learned via the mVAM blogs as they occurred with each of the mVAM projects. mVAM, from the start, has demonstrated a commitment to open data (as illustrated by its relationship with the Humanitarian Data Exchange, or HDX, and the

---

[3] See **Joint Research Center.** 2014. *Impact of the West African Ebola Virus Disease Outbreak on Food security*. JRC Science and Policy Reports. Luxembourg.

online availability of data sets). The production of timely information products in difficult settings demonstrates that shock monitoring is now nearly universally possible. mVAM leadership highlighted the important role of the first primary donor to the project, HIF, in promoting continuous and open communication about the pilots. HIF played a large role in ensuring mVAM published as much as it could about the context, progress and results of the projects. Some of what was learned was then written into the proposals and grant agreements.
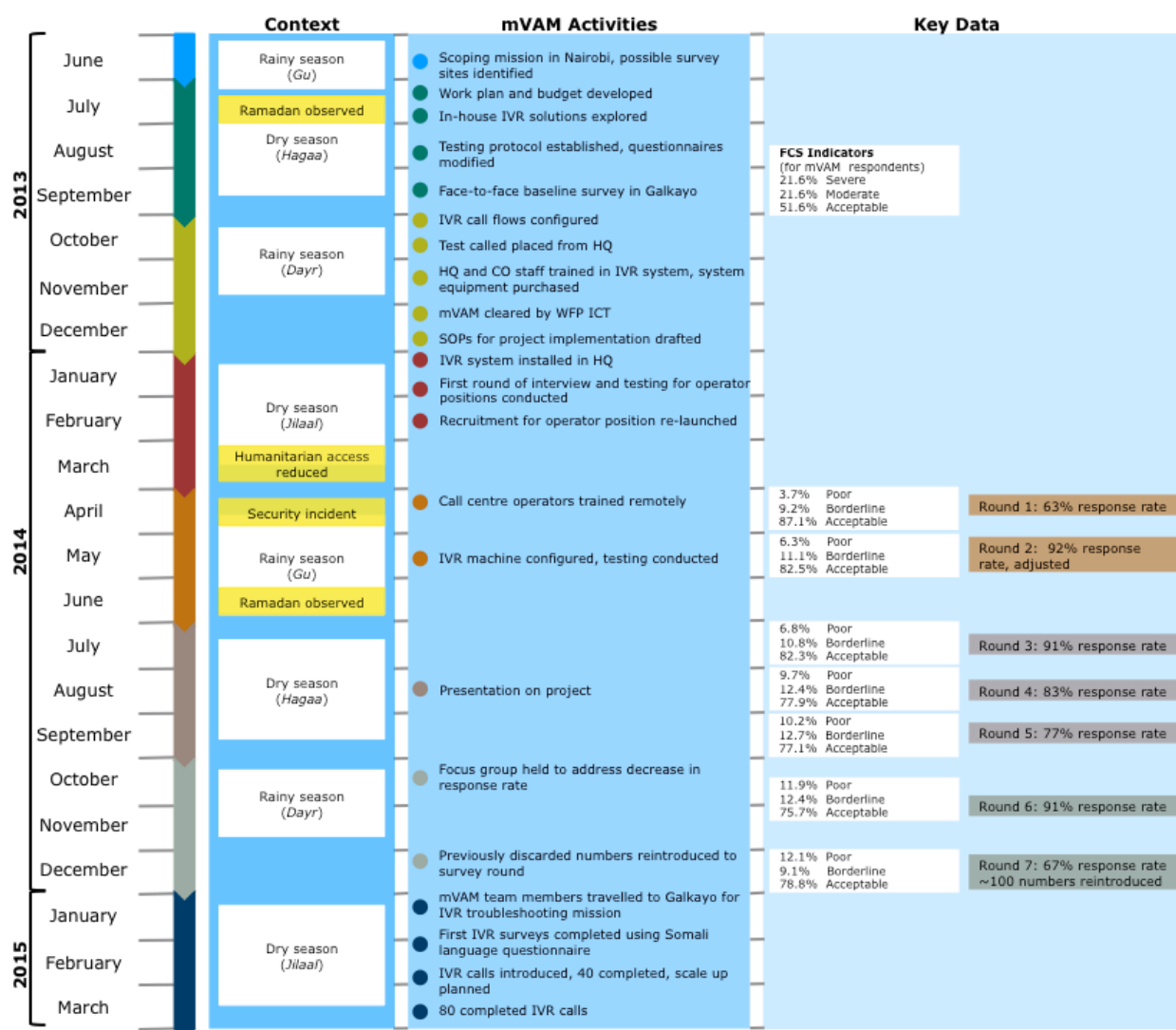

Figure 6: Galkayo, Somalia timeline

mVAM also demonstrates that inexpensive data collection in difficult contexts is possible and that high frequency monitoring can be efficiently achieved. Data streams add value to existing approaches by providing unique information where none might otherwise be gathered.

Topline Finding: the case for mVAM as a 'Learning Lab'

Knowledge gaps will be a critical limiting factor in the successful scale-up of mVAM approaches. The team found that the existence of rigorous evidence-based trials of innovative mobile technology based monitoring strategies is very limited. The pilots and subsequent deployments of mVAM mobile data collections have already shown the importance of understanding the nuances of this innovation in different contexts. mVAM's creative partnering to evaluate the validity of its pilots and implementations is laudable. However, mVAM data is part of a larger information ecosystem that supports decision-making, so the learning dimension of mVAM can be far deeper and more substantive. It should involve formative evaluation of different methodologies and information strategies as well as research on information use and in better understanding the information ecosystems in which it is situated both globally and locally.

In each of the programme activities, there has been a great deal of *ad hoc* innovation. One interviewee commented, "mVAM should explicitly acknowledge, encourage and cultivate innovation from the field." The research, curation and then capacity development based on this knowledge and learning is perhaps the greatest potential role of mVAM at the global level. The idea of an mVAM 'Learning Lab' to fulfil this function may also be attractive for resource mobilization based on the demonstrated need for more systematic research and learning as a foundation for an mVAM research agenda for different contexts. mVAM demonstrated attention to capturing and documenting lessons learned, an activity that might be enhanced in its future development.

Learning labs will be essential to the function of WFP as the landscape of technology and population use of technology is changing quickly, as is our collective understanding of how to understand changing trends in behaviour. Dynamic organizations such as WFP will require spaces committed to focused learning on these issues. Learning labs will need the authority to operate with sufficient flexibility to respond to situational awareness and emergent information, placing an emphasis on constant innovation and experimentation. mVAM has demonstrated its leadership as an innovation center within WFP with an appropriate culture of experimentation and learning. To date, an opportunistic tactic has been used in piloting approaches, focusing on places where staff have good relations or a where compelling need arises. While the dynamic nature of mVAM applications will remain an important component of its strategic approach, mVAM's next phase may also want to systematically learn in order to improve procedures and inform future deployments.

The review team would not want to assume to represent completely the views of the WFP COs and Area Offices or the beneficiaries, but there seemed to be a qualitative difference in the focus of their priorities for better evidence and understanding. Some research priorities that were highlighted in the limited interviews were:

- Generalizable learning on modalities in the different scenarios discussed below with reference to five possible scenarios;
- Levels of credit granted to participants and food and nutrition security (FNS)/protection/livelihood effects;
- Impact of access to mobile phones for people of different ages;
- Controls on gaming and bias in response;

- Closer integration with other WFP programme streams such as Cash and Vouchers (C&V) and monitoring;
- Links to big data as new source of secondary data. Comparative value of mVAM and simply monitoring trends from big data such as done by UN Global Pulse, an extension of mVAM's current work with UN Global Pulse;
- Best practices in controlling fraud and abuse under the pretense of a WFP/mVAM project;
- On-going monitoring and research on protection challenges and opportunities with respect to mobile technologies; and
- Benefits and costs of panel data compared to repeated cross sections.

External partners also provided excellent suggestions for future research that included spatial analysis of the responses and statistical effects of monitoring frequency. We outline in the following 'Findings' some of the key themes for learning.

Topline Finding: five scenarios that may require different mStrategies and experiments

The review team found that mVAM rightly experimented in very different contexts, which gave rise to a deeper understanding of how context matters and that appropriateness of mobile information monitoring strategy is a key consideration. Considerations include whether panels or repeated cross-sectional surveys might be the better approach. Important to note is that randomized experiments to identify best strategies for data collection, analysis and dissemination can be feasibly implemented in many mVAM contexts. This is a theme that fits into the broader 'Learning Lab' theme. The review team suggests that mVAM should begin to develop a typology of contexts for the application of mobile monitoring strategies. We outline a few of the scenarios that might be developed further into typologies:

1. Situations where face-to-face interviews are not possible or feasible such as Ebola and severe conflict: in these cases, mobile strategies might be the only option for WFP to access information about difficult to reach populations. We believe, however, that mobile strategies in these contexts also might include multiple method techniques such as in-depth interviews with key informants, key indicator surveys and situational awareness monitoring through open-ended questions and an iterative data collection approach. A mix of panel and repeated cross-sectional approaches might be considered. We suggest that randomized experiments would work well in this context at some additional cost.

2. Situations where vulnerable populations are on the move, including cross-border, such as in the important examples of the Horn of Africa and the Sahel: the review team found that mobility introduces challenges to mobile cellular access. mVAM can focus further on developing methods and approaches to better monitor vulnerable populations on the move. More information is needed to determine the mix of design considerations. For example, host populations and IDPs may respond differently to IVR or SMS because of different work schedule or literacy rates. A single data collection effort might include more than one modality to best accommodate the intended participants.

3. Dynamic camps with frequent and/or large scale in and out movement: this is another example of a prominent vulnerable group of interest to WFP. Like the second scenario above, mobility

may present specific challenges. However, mobile technology can be better harnessed to address this challenge. Again, here, a mix of panel and repeated cross-sectional designs should be considered.

4. Stable refugee and IDP camps: these have been traditionally monitored through face-to-face interviews. Some combination of face-to-face and mobile techniques might be most appropriate for providing better information for these groups. Formative research here might focus on the effects of testing and/or gaming bias.

5. Stable vulnerable geographies: mobile technologies can provide better access to a broader geographic area. The geography of the responses that can be automatically collected in a random call cross-section has enormous potential to contribute to a more refined analysis. For example, the responses could be compared to a recent health survey in a specific district based on location or seasonal migration could be included based on where the responses come from over time. Because of the efficiencies of SMS or IVR data collection, it may also be possible to sample a wider demographic as a comparison group for assessment of particularly vulnerable places or for WFP beneficiaries.

mVAM is continuously improving its approach through a healthy process of learning by doing. Ongoing operations research would benefit from some consistency in research design, a more systematic approach to capturing learning that includes more quantification where appropriate, and a strategic prioritization of learning agenda around key research questions. It is time for mVAM to move from proof of concept to a full-fledged 'Learning Lab' that applies systematic formative research to food security monitoring.

Disaggregation of data by sex and age is important because of the technological component of the approach. Where possible, face-to-face baseline, mid-term/seasonal and post intervention face-to-face surveys should be included in operational research sites. The FCS and Coping Indexes are quite sensitive to context, as is the case with the majority of proxy indicators. Anthropometrics such as wasting, stunting and body mass index (BMI) would allow better baseline measures that could be interpreted against well known standards to assess nutritional status of research populations. Similarly, a rigorous consumption model should be included in the baseline survey that would allow the FCS to be compared to caloric and micronutrient standards. Qualitative analysis can be employed to better understand the potential differences in populations with and without mobile phones as well as to better understand local coping and livelihoods that promote more nuanced interpretation of any Coping Index.

mVAM's learning has been documented in the blog and a number of other communications and documents were shared with the review team. The blog, in particular, is undeniably an excellent source for capturing and sharing learning from the innovative work. In Goma, the team kept records about the mobile phones' functioning and loss. Meetings with the project participants were captured in some project monitoring notes. mVAM, internally and in partnership with Nielsen and GeoPoll, carried out quantitative comparisons of response rates and other relevant research questions such as recall period. A broader range of quantitative metrics, however, was more difficult to find. A more systematic approach to capturing the learning from operational research would start with developing indicators based on an agreed ToC. For each element and relationship on the ToC, an appropriate number of quantitative and qualitative indicators would be selected, measured and monitored. By capturing the

same indicators across different projects, meaningful comparisons could be made. For example if 'Partnerships' is included on the ToC, then mVAM should define partners and collect information about the type of partner in standard categories and measure the quality of partnership, such as assessing the frequency of information sharing or number of innovations/synergetic ideas. Undoubtedly, standard measures of efficiency, timeliness and use of mVAM data in different contexts would allow comparison and promote sharing of better practices.

Mobile data collection through voice interviews, IVR and SMS all should be incorporated into future VAM analysis as demonstrated in this successful proof of concept. Formative and on-going research might identify the best modes/approaches at any given point in space and time. Going forward, a strategic prioritization of a research agenda will be more about including the right stakeholders and developing a common vision of purpose than comparing technologies in an absolute sense. After a few deployments, mVAM needs to decentralize, broaden the questions to which it responds and bring more stakeholders on board, while at the same time developing a small list of core indicators that it promotes across mVAM implementations. Based on a ToC with a strong vision of the future and aligned with the organization's mandate, mVAM should engage in the development of an applied research agenda in collaboration with relevant stakeholders. mVAM may consider extending its partnerships to include 'Learning Lab' conveners from academic and quasi-academic partners as well as community leaders.

Topline Finding: operational decisions could be informed by high frequency monitoring

Operations and programme staff in the DRC face a resource constrained environment where decisions on when and whom to target for assistance are made on a monthly basis. Monitoring data was seen to be most potentially useful if it was tied to triggers for increasing or decreasing aid. Also, monitoring of different programmatically important groups such as host populations, vulnerable groups or displaced groups would support prioritization of limited resources. Management in the Goma office saw value in regular, reliable and frequent monitoring that could become a routine source of information for operational decisions and for discussions with partners; for example, as a major input to the DRC Integrated Phase Classification (IPC).

Box 2: Use Cases of mVAM data in the Goma, DRC pilot

(1) An early use of mVAM monitoring was analysis that showed non-beneficiaries reported poor consumption and high levels of coping. This information triggered a rapid assessment and eventually another distribution. The distribution was for both current beneficiaries and other non-beneficiaries in the camp.

(2) mVAM was used as part of justification for door-to-door rapid assessment to verify beneficiary needs. Area Office staff also used mVAM monitoring to show "impact of WFP's ration and overall vulnerability situation" following a distribution to the IDP camp.

(3) mVAM analysis was used in an "extrapolation" to other camps including those from another displacement event and outside Goma (IOM 2009 camps). The potential for similar food insecurity in these groups due to seasonal and market issues was a trigger for a new round of surveys.

> (4) Some of the IDP camp residents have started returning home but mVAM continues to monitor their food security situation through the monthly voice surveys.

The use-cases identified in the review are both striking and novel in that they often provided information that went against the conventional thinking (see **Box 3**). In the EVD deployment, monitoring data suggested an increase in food security in rural areas that may have had lower numbers of cases but higher vulnerability to food insecurity. In Goma, DRC, monitoring showed increased food insecurity of non-beneficiaries after a targeted distribution scheme was put in place. It is unlikely that either of these important trends would have been noticed in a timely fashion by a traditional cross-sectional survey. The reviewer proposed that it was the high frequency information that repeatedly showed a more and more worrying trend, and it was this that triggered the management decision and programmatic action.

Box 3: Use cases of mVAM data in the EVD-affected countries deployment

(1) mVAM data showed that food insecurity was deteriorating in rural areas that had comparatively low number of EVD cases compared to urban centers. This informed the strategy for food assistance during the response.

(2) The Joint Research Center funded by the European Union conducted a detailed study of the food security effects of the EVD crisis and referred frequently to the mVAM monitoring data.

(3) The International Growth Center that conducts research to inform development policy also used mVAM price and wage data in their analysis of the longer-term effects of the EVD crisis and was shared on the web at www.theigc.org/news-item/the-economic-impact-of-ebola-november-2014-report/.

Mobile technology, the quick setup of mVAM deployment, and the possibility to rapidly customise indicator collection to fit response manager's information needs makes mVAM a truly innovative tool for WFP with great potential to improve programmatic response and impact. High frequency monitoring data of this sort has not been available for programme decision-making before and based on early experience WFP managers are capable of interpreting the data and putting it to operational use. Future research and development of mVAM should focus on this interface between high frequency data and decision-making.

Topline Finding: the need for a wider vision of mobile monitoring

The highly focused nature of mVAM work presents a double-edged sword. On one hand, it is easy to articulate a project vision and to explain what mVAM does to potential donors. On the other hand, the narrow focus of the experimentation of mVAM on pre-selected indicators and tendency to favour more efficient but restricted modalities like SMS limits the potential relevance and impact of the work. The focus on two basic indicators has been a trend in VAM monitoring over the last several years, and mVAM has in fact shown the ability to quickly incorporate other indicators at the request of the in-country programme staff and leadership. This is where the mVAM pilot really had a chance to demonstrate its potential for near real-time context relevant and programme specific information. WFP VAM made a significant research investment over the years to select and then develop guidance on Coping Strategies Index (CSI) and FCS as primary indicators of food security across contexts. These two

indicators are mandated to be measured as part of WFP's strategic framework and it was natural for mVAM to focus on the most efficient way to collect this policy-mandated data. In the field, mVAM data is taken into consideration when making decisions in conjunction with a wide variety of information brought from partners and shared through the cluster. Those in charge of operations and programme at an Area Office or CO who were interviewed universally acknowledged that mVAM data was largely peripheral to their normal decision-making process, perhaps in part due to the pilot status of the effort, and was used only when helpful to support a decision. The two particular indicators did not specifically provide the information they needed. One of the interviewees sensibly expressed that, with all of the other contextually relevant information and other security and operational concerns that affect a programme area, they "cannot make decisions based on mVAM alone – that is impossible." However, the potential for producing timely, targeted and programmatically relevant information was signalled by mVAM's provision of contextually meaningful information, such as assets or prices, in response to data requests from the Somalia office and the EVD response management

During the pilot phase, it was fully appropriate to focus on a limited number of issues, modalities and indicators. The expansion of the approach in the EVD deployment really showed great potential. Given that, in some cases, mVAM may be the only source of social behavioural data, mVAM should now consider a wider mission. mVAM might consider strengthening its focus on some areas it is already working on such as:

- Situational awareness by using directed but open-ended questions.
- Mobile surveys to triangulate and 'cross-check' alerts raised in cluster meetings. Clusters often have requests to issue alerts but there is often little situational awareness. This would be a more flexible use of mobile than just panel surveys (e.g. compare to the cost of calling mobile numbers to visiting sites by helicopter).
- Work with other units to apply learning for mobile phones and smartphone applications for serving clusters far away from capitals, specifically to share presentations, support data collection, and provide maps. One mVAM staff commented this is already happening in a project called mobile data collection and analysis (MDCA).
- Work with other units on the integration of mVAM monitoring with Cash and Voucher programming to enhance quality. Monitor with same phones used for mobile money. This may also be a more appropriate use of IVR.
- Work with other units to integrate with post distribution monitoring.
- For IDP situations, monitor food security as people return home.
- Monitor people on the move during displacement.
- Monitor host populations.
- Monitor Food for Work output and outcomes.
- Integrate mobile communication in various community-based approaches. These may include food for training, food for assets and food for work activities. Great potential exists to involve community based organizations and include volunteer committees. During stakeholder interviews, a number of complementary WFP efforts were mentioned, including the market information team, SCOPE, WINGS and COMET. The ability to link cell phone based surveys with project log frames, beneficiary cash and voucher data bases provides a powerful approach to beneficiary monitoring and impact evaluation.
- Youth learn to use all the functions on the phone very quickly and mobile integration of youth focused programme may be a natural synergy. Enabled by mobile phones, youth may be able to

25

      play larger supporting roles in building capacity of older or vulnerable participants and also in contributing to food security monitoring.

- High frequency monitoring could meet policy requirements for better understanding of risk, context and programme outputs.
- Spot checks of programme activities as has begun in WFP Somalia.

Topline Finding: mVAM has a profound potential to empower participants/beneficiaries

Mobile technologies provide a mechanism for stronger connectivity between WFP and its beneficiaries, which may result in more effective and beneficiary relevant programmes. In the DRC, beneficiaries interviewed did not consider the short voice interviews as the primary purpose of the pilot and identified a number of other uses of the phones and credit as the key project impacts. The most common response when participants in Goma were asked the purpose of the pilot was "communication." mVAM can also make better use of its strategy of incorporating open-ended questions on survey monitoring tools through successive iterations with participants and more purposeful open ended questions. Additional formative research is required to develop approaches to high frequency measurement among non-beneficiaries.

Open-ended questions require different analysis and often manual coding of answers. Analysts can be given training on coding sheets that ask for metrics about the responses such as deciding if a response was positive or negative. Software for qualitative analysis such as ATLAS, NVIVO or an extension for SPSS can be 'trained' to extract specific words and phrases that can be then counted and analyzed. This may be appropriate for large samples. Micro-tasking may be another option for crowdsourcing the translation and coding of open-ended questions.

An interactive system that solicits feedback on programmes, but also on WFP analysis or operational decisions, could go a long way toward involving beneficiaries in the programmes that affect their lives. For example, VAM findings could be verified with a reference panel of mobile users, asking if the identified trends seemed to be true and asking for contextual information about why or why not food and nutrition security may be changing in a given context. Deliberative polling is a technique where participants are asked about an issue and then provided with relevant information on the topic. New polls are conducted and results shared with participants. This may be an empowering model to inspire future use and design of mVAM analysis.

In addition, the mobile phones and credits acquired by mobile survey respondents can enable participants to be better decision makers for their family food security and safety of their family. For example, in the DRC, some of the identified beneficiary uses of the phone and credit include:

- Participants arranging small loans and gifts from family. In Goma, the gifts often came as top-ups on credit or through a mobile money mechanism.
- Planning returns home and accessing security.
- Reuniting lost family members.
- Radio during the day and a torch at night.
- Games on the phone (potential in Food for Training activities of other health and nutrition messaging).

26

- Contacts in preinstalled in SIM could link to messages or services and hotlines.
- Alarms to encourage survey participation or support behaviour change programming.
- Protection, safety and camp security.
- Location services.

mVAM has not emphasized this very important potential but should do so. mVAM and VAM should not advance too narrow of a role for itself in contributing to Zero Hunger and other stated goals of WFP. VAM has always recruited staff with unique qualitative and quantitative competencies as well as spatial analysis skills. mVAM is well placed to play a lead role in research, capacity development and knowledge management for the collection, analysis and communication of data collected with mobile phones. Indeed, the reviewers encourage mVAM to consider just that, given mVAM's demonstrated applied research approach.

Topline Finding: participant engagement and volunteer community-based approaches may be key to sustained response rates and reaching the most vulnerable

In the DRC, the excellent application of community based approaches and staff facilitation accounted for pilot success. Local staff organized a volunteer committee to support and monitor participants in the camp, track down and solve issues, and provide communication links with participants. Operators also were physically present particularly during kick-off activities and periodically participated in meetings with the community. To some extent, this connectivity with beneficiaries compensated for lack of a comprehensive participatory assessment of the context to inform project design.

In the local context, phones were often not in possession of the intended participant but held by fathers, husbands, extended family and in-laws. This caused minor inconveniences such as operators needing to ask the person with the phone to find the panel participant. In other cases, this caused instances of injury and harm due to violence related to control over the phone. Greater participation of mVAM participants in activity assessment, design and implementation has the potential to mitigate challenges, ensure continuing engagement and could perhaps be instrumental in reaching the most vulnerable and those without phones.

WFP management recognized the role of participant engagement in the pilot success, stating, "the ground work is necessary. The project worked because of the operators and the participant mobilization." Operators concurred that nearly all participants wanted more information about WFP distributions, the project and other issues during the calls. The operators also noted, "the people must understand the project - if they understand they will participate and protect the equipment."

Topline Finding: data credibility issues, social desirability and the concern with 'gaming' responses

Credibility has always been and will always be the prime driver of information use. Different approaches to data collection, refined statistical techniques, and response rates are not the concerns that the *primary users* of mVAM information identified. Instead, data users interviewed expressed concern with validity issues associated with gaming or desirability bias.

mVAM, on the other hand, is very much on the right track in trying to establish the validity of mVAM data with the scientific and practice community as well as external partners. Credibility with these

27

stakeholders may also result in credibility with resource partners. mVAM's work to improve key indicators through testing modes and comparing recall periods is a very positive development. The FCS has been used for several years and mVAM's attention to indicators of family dietary consumption is much needed. We believe that the development of a rapid and comparable dietary intake indicator for monitoring is important not only for mVAM but for WFP and the food security community more broadly. The same can be true for the reduced Coping Strategies Index (rCSI). Clearly the correlation between the FCS and rCSI has not been carefully evaluated and should be in order to interpret the findings of mVAM work.

With the assistance of its survey methods corporate partner, mVAM undertook research to establish some measures of criteria and face validity. However, the finding that mobile strategies/methodology validity is highly contextualized means that mVAM must strengthen its capacity to adapt and customize its approach to the context in order to establish and maintain credibility of mVAM data and information. Credibility tends to be a primary driver of information use and perhaps this testing rigour will improve credibility with resource partners. However, if the primary intended users of the mVAM products are managers with programmatic and operational decision-making responsibilities, those interviewed had one particular issue with the credibility of mVAM data, namely the perceived incentive to exaggerate issues of food insecurity.

The potential of participants to give intentional false responses or try 'gaming' of survey responses was a concern expressed repeatedly by stakeholder information users. This is not a concern that is unique to mVAM, but the capability for high frequency contact with participants and the remote nature of the interviews expose mVAM to a higher risk of this type of error. Although this can be the case in face-to-face interviews, there was the sense that collecting responses over the phone would be less credible because of the lack of contextual data and observation, even when preceded by face-to-face interviews. mVAM must focus on the credibility of their data products with their intended audience. Monitoring food security data, by definition, is collected for the purpose of adjusting programme and operations. Through engagement, participation in the monitoring and additional research, mVAM must win over the confidence of operations and programme staff in the accuracy and validity of their information products.. Building the credibility and facilitating the use of monitoring data is perhaps the most important challenge for mVAM to face and recommendations such as the learning lab are intended to address.

An astute commenter on the review helped explain in detail how the context in Goma, DRC presents a high risk for bias: "The eastern DRC is an area that knows the situation of conflicts and population movements for over two decades. Some people in this area have already been displaced more than once. In addition, the vulnerability survey conducted in North Kivu in the IDP camps has shown that about 30% of IDPs have been in these camps for over three years. These people are accustomed to the frequently asked questions in the food security survey and they know how to manage their answers knowing that there is behind the possibility of assistance if the answers are given in certain way. In the context of IDPs in the Goma camps where many households have been there for more than three years, the results of mVAM surveys can be biased by the fact that the households interviewed are familiar with these questionnaires and they know how to respond to attract the assistance."

Although it may be redundant to raise in this review, the potential of participants to give intentional false responses or try 'gaming' of WFP survey responses that were related to food security was a repeated concern. One external stakeholder simply stated, "Just because it is a call from WFP - there are all sorts of incentives to bias your answer." The implication is that WFP is well known for distributing food to those that have the greatest need so there is a perceived incentive to exaggerate a respondent's household food needs. This concern about gaming was not shared by some of the analysts that produced information products with mVAM or similar data. mVAM actually presents a unique opportunity to test this perception. With mobile data collection, the operator or SMS does not need to have the same prominently displayed WFP logo that one might find on a Land Cruiser or an enumerators t-shirt during a face-to-face survey. With a sufficiently large sample size, an operator could give an elaborate introduction about the role of WFP in the collection and use of the survey data and a comparison group could have sufficient but perhaps more modest explanation of the WFP's role. These types of research questions include ethical dimensions and should only be done with the oversight of an Institutional Review Board (IRB) but could potentially address these persistent perceptions and concerns about bias.

Contrary to this general perception, common sense review of the food security assessments and monitoring indicate that in fact most people are not typically food insecure and likely answer surveys truthfully that in turn allows analysts to make reasonable determinations of severity and magnitude of food security problems. As one commenter observed, "If we were [consistently] being gamed, all our data would be bright red [indicating food insecurity]." *One external expert that was consulted said they had done four rounds of a panel of the CSI in rural Ethiopia and there was **no evidence of gaming** the responses.* This dispute is something that could easily be tested, as described in the previous paragraph. This is where the footprint of mVAM and WFP can be leveraged to make major contributions to understanding the utility of mobile engagement with vulnerable populations. Gaming may be more of an issue in certain context than in others and WFP mVAM has the potential span of operations to systematically collect the information necessary to answer that question.

> **Specific finding: operators use contextual knowledge and probing to elicit accurate responses from participants**
> Operators in the DRC said they used knowledge about different seasonal food availability and other contextual information to correct and probe for more accurate responses. This probing is probably the primary reason for the differences between responses collected by the varying modes. As one operator described the process of getting accurate information during a call, "the time of year is important. You need to know the seasons. Now beans are everywhere. Seek the reality." WFP Somalia staff highlighted knowledge of context and highly skilled operators using a variety of probing techniques as key to obtaining accurate responses. This has implications for capacity building, further development of SOPs, and future research.

Like the information consumers, WFP staff collecting information closest to the field were universally convinced that a significant proportion of respondents did change their answers based on the anticipation of the 'right answer' - either an answer that the interviewer was seeking or an answer that might bring the respondent some benefit. Although this is perhaps weak evidence or pretense to not acknowledge information that might indicate inadequacy of programming, it is the *perception* of validity

rather than the *evidence* that drives information *use*. See the discussion in **Box 1** on bias earlier in this document for links to academic literature that would completely support the conviction of VAM staff in the field and confirm this is not a problem with data collection that is unique to WFP. A variety of suggestions were made to counter the tendency of some respondents to give false information. Triangulation with other indicators or even completely independent sources of information was a commonly suggested technique. Closely monitoring trends from the baseline measurements and identifying respondents that are deviating in unexpected ways from other respondents with similar profiles is another suggested technique. See **Box 1** for other specific statistical and methodological techniques for identifying and correcting gaming biases.

A very experienced field VAM officer suggested a number of ways to handle gaming that are not technical or statistical fixes. Academic experts working on large panel studies in high stakes research with sensitive topics agreed that non-statistical approaches to increasing data quality are particularly effective in context where issues of social desirability and gaming may be a concern. Most of the interviewees were convinced that gaming is very likely with repeated calls. This did correlate with observations by the review team during the focus group in Goma of elites coaching focus group participants on their answers and trying to influence responses in ways that could potentially bring a benefit to the group or the elite themselves. To triangulate, operators making calls did provide specific examples of some respondents trying to guess the answers that would bring benefits and the requirement to 'probe' for the truth in a significant number of interviews. mVAM management has discussed this with operators and offices implementing the mobile data collection. This may warrant a focus on capacity building and retention of highly experienced operators. Good practices could also be integrated into Standard Operating Procedures (SOPs), guidance and capacity development.

A commenter provided a good suggestion in the "proxy means testing" approach; "In vulnerability surveys, the food security cluster has already an experience of using the proxy means testing whereby the FCS is calculated by proxy without asking the classic question of what the household consumed during the past 7 days. This experience could be improved with further analysis of the questionnaire already developed."

This is a list of non-statistical methods suggested by a variety of experts to reduce problems with social desirability and gaming include:

- Probe respondents for more complete or justified answers and periodically conduct face-to-face meetings.
- Ask a set of questions to identify social desirability bias, such as "I never regret my decision," or "I never get sad," which can potentially help to identify individuals who generally don't answer truthfully.
- Include questions that test, via a proxy indicator, information a respondent may have provided. For example, asking respondents about their favourite programme to test if they have a TV and therefore fall in a certain wealth category.
- Gather qualitative data to identify the general response of participants to the study, potential issues involved in data collection, and the extent of misreporting.
- While individuals are not going to admit lying on a survey, survey could ask participants if they're aware of any friends/acquaintances reporting inaccurately (without asking their names). This assumes that respondents know one another.

30

- Ask the same survey to various populations, but offer differing levels of incentives to identify the effect of incentives on reporting patterns.
- Monitor variables that should be correlated, such as income and consumption, to compare indicators against expected relationships. Data that contradicts the predicted correlation may be a sign of increased bias. This could be done with other ratios or correlated variables.

It was also noted that there would always be some trusted informants. The quality and credibility of the answer depends on what one asks, how many questions and to whom at any one time. In the end, high quality and nuanced analysis is required to make use of mobile data in order to pull truth out of noise from gaming and misunderstanding. For example, meta-assessment and analysis of trends in individual responses, group responses and across projects may identify issues with individual respondents. That said, interpretation beyond graphs is required and should focus on putting the responses in seasonal, programmatic, political and other relevant context. This is sometimes referred to as the 'common sense' test for validity and is best done by those nearest to the work and most intimately familiar with the specific context. This type of analysis requires a highly experienced or extremely perceptive officer that understands context and would be very hard to achieve through remote analytical support.

There are arguments for and against the use of panel data. The research agenda for any 'Learning Lab' type of activity should include some form of decision analysis to see if the added information from a panel is worth the risk associated with repeated interviewing. Perhaps a simple repeated cross-section would produce more valid results. This is a question to be answered empirically.

Topline Finding: mVAM data products and information dissemination/communications strategy is evolving but requires considerable improvement

The blog on the HIF project page,[4] which since March 2015 has been cross-posted to an mVAM blog on WordPress,[5] is exemplary as a communications tool for the food security monitoring CoP. The HIF required mVAM to blog about the pilot activities in DRC and Somalia. Blogging up until that point had not been a strategy for sharing VAM or WFP innovations with a broader audience. When the HIF pilots concluded, the blogging continued on the WordPress site and one of the lead innovators also tweeted about news, new blog posts and bulletins. This use of social media represents a commitment to open development and information sharing, and in general could also form the basis for a communication strategy for VAM innovation and information products to a broader group of stakeholders.

In general, the mVAM blogs were not well integrated with the data product dissemination site during and the initial pilot and deployment periods. Navigation between the HIF project blog site and the mVAM site on the WFP platform was difficult and neither site offered easy access to the array of information being produced by mVAM. It is clear that the level of coordination and information dissemination strategy is evolving as the project matures, and significant efforts to unify and simplify the mVAM web presence and its data product dashboards have already been made. The recent mVAM

---

[4] mVAM project blog for HIF: http://www.elrha.org/project/mvam-piloting-mobile-voice-technology-for-household-food-security-data-collection/.

[5] mVAM: the blog, on WordPress: https://mvamblog.wordpress.com/.

monitoring page[6] compiles relevant information for a variety of mVAM settings in an easy to navigate interface. While not all mVAM products are available on this site, and while it is not yet clearly linked to the WFP landing page for mVAM, the user experience is already much improved. From November 2014 to March 2015, mVAM's data product dissemination site on WFP's more formal platform[7] did not have heavy traffic as measured by Google Analytics (**Graph 1**). Most users were based in Rome, and while these statistics are an incomplete measure of information use as reports and data were available from other sources, they do highlight that mVAM may initially be talking to itself more than it hopes to do beyond pilot. It is noted that the analytics available are limited in their scope and further trends in product usage and dissemination may emerge as the mVAM programme, as well product dissemination platform, develops.
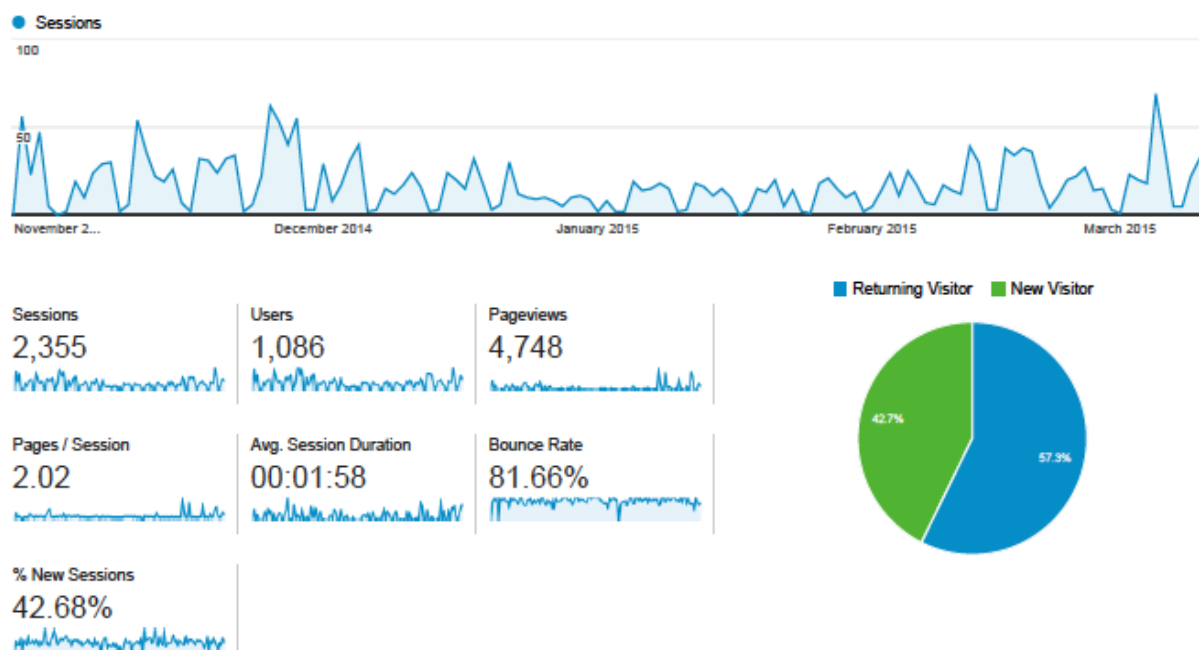
Graph 1: mVAM Monitoring website analytics October 27, 2014-March 11, 2015

Of equal concern is the need to invest in the production of higher quality reports and bulletins. The bulletins suffer from lack of consistency of reporting of indicators, data presentations that are misleading (FCS consumption score graphs or figures with a prevalence scale starting at 40%), and lack of analytical interpretation and integration of other data sources. The EVD bulletins were stronger in these respects, perhaps because of the partnerships that mVAM developed to produce analyses as well as the specific situational context. The quality of bulletins is an area of concern and requires the same focus on innovation that mVAM has placed on data collection and communication of the mVAM process development.

---

[6] mVAM Food Security Monitoring: http://vam.wfp.org/sites/mvam_monitoring/.

[7] mVAM site on WFP: http://www.wfp.org/content/mvam-mobile-voice-technology-food-security-data-collection.

mVAM should place greater emphasis on engaging its stakeholder/users through users surveys and more intensive engagement to better create decision support for its key stakeholders: WFP operations/policy, donors, implementing partners and beneficiaries. mVAM should consider more non-traditional information products such as visualizations, videos and virtual workshops. The VAM Resource Center[8] and the WFP dashboards on the HDX[9] are promising efforts in this direction.

Some of the immediate requirements for context assessment and capacity development include the following specific findings:

Specific Finding: most WFP Area Offices, and potentially Country Offices, do not have capacity for mVAM implementation and data analysis
Pilot data from the DRC was sent to Rome to be analyzed and returned to the Area Office in the form of graphs and some commentary. There was not a VAM officer in the Goma office at the time and managers thought that the skills to process the mVAM data into graphs was not available in the existing staff. Strong VAM capacity may not always be available at the Area Office level. When asked about the qualifications to analyze and promote the use of mVAM data, some interviewed felt strongly that an international officer was required while others felt the skill set and qualifications of the individual was more important. Resources and time were also suggested as constraints in addition to lack of specific skills that could be overcome through capacity development. Interviews with the Somalia CO also identified capacity for data analysis as a key constraint. The VAM officer placed more importance on the ability to interpret the mVAM data in context and knowledge about seasonality, livelihoods and the conflict dynamics as key to credible analysis of the mVAM data. They thought a highly qualified national officer would be best placed to implement and analyze data from mVAM activities. Depending on the context, mVAM activities will need additional support to make the data usable for internal and external stakeholders.

mVAM should evaluate the possibility of automated report generation and training of field staff in modern methods of analysis, including the use of visualization analytics. Before much investment is made in this area, however, a better understanding of decision support needs based upon more careful assessment of use cases and decision makers (donors, WFP, IPs, governments and affected populations) is needed.

Specific finding: some operators complained of fatigue, ear problems and headaches because of perceived low quality headsets and work stations
It is recommended that mVAM assess industry standard for call center equipment and invest in high quality headsets. Good practice on setting up call center workstations should be included in SOPs. Even if this is not universally reported, small investments and care can ensure quality work particularly with repetitive tasks.

Specific finding: implementation issues are key to programme success
Project staff viewed addressing practical issues, such as ensuring phones are in working order and that participants can be reached, as highly important to the success of mVAM. Most of the solutions found to

---

[8] http://resources.vam.wfp.org/.
[9] https://data.hdx.rwlabs.org/organization/wfp.

implementation problems emerged from project staff, the volunteer committee and participants themselves. These learnings need to be recorded and inform future SOPs and should be included in future assessments and planning in support of mVAM activities in new areas or COs. Please note that not all problems were experienced in all pilots. This is not an evaluation but a review to inform future mobile data collection activities. Examples of the implementation issues to be solved in future deployments:

- Participants ability to use phone;
- Battery problems;
- Charging problems;
- Credit bonus did not (or perceived to not) arrive reliably;
- Mobility of participants;
- Network problems;
- Mandatory breaks in contract for operators – no one to call participants;
- Participants – confusion over SIM and different networks – paying for credit with wrong company; and
- Broken, lost and stolen phones.

Specific finding: ICT staff in Goma and mVAM core team members received timely support from InSTEDD

Specific finding: using familiar voices may encourage response rate and accuracy of responses
The operator from the Goma pilot also recorded the questions used in the IVR system. This use of a familiar voice was mentioned in several interviews as a comforting to the participants who received IVR calls. A similar strategy of using local radio personalities or other recognized voices may influence response rates and quality of responses. In Somalia, the mVAM staff plan to partner with a well known local radio host to re-record the IVR audio files. This could be a subject for further research.

Specific finding: field based project staff did not feel there was a clear process for receiving permission to change questionnaire or implementation procedures
The pilots we great success despite the normal challenges of implementing an initiative led from staff based in Rome and Country Offices that had normal staff rotation. In the case of Goma, no VAM officer was present during much of the implementation and Somalia had challenges of staff movement due to security issues and staff turnover. Communication was uneven as a result. Several WFP staff involved in the pilot felt that issues with the questions and procedures for implementation were raised but not necessarily resolved with mVAM programme leaders. mVAM Rome also expressed that it was also difficult to resolve issues remotely because of time and capacity constraints as well as due to the absence of a VAM officer. A regular meeting, discussion, and process for changing questions and implementation procedures is suggested. Some of the questions expressed by Area Office staff were in reference to when it was appropriate to replace a lost or stolen phone. Other examples were given regarding survey questions that respondents did not seem to be able to answer or did not seem correct in the context. As mentioned in a previous finding, the presence of a local focal point, whether a VAM officer or otherwise, may also support the clarification of procedures and processes.

Specific Finding: CO and Area Office priorities should guide mVAM activities

Beginning with the mVAM scoping missions, pilot sites and questionnaires were decided upon in consultation with CO and Area Office staff. In discussion of indicator selection, the implementation sites chose to collect FCS and CSI indicators, and in the case of the EVD survey added a new qualitative element as well as a new price data indicator. All stakeholders agreed that the prioritization of questions, target populations, timing and implementation of the mobile data collection activities should be determined based on field priorities. During the pilot phase, it was appropriate that a particular protocol with fixed questions was followed. Moving forward with mobile data collection or scale-up of mVAM monitoring, standardized procedures should include the intentional assessment and prioritization of field programme and operational information requirements in future mVAM activity design.

Specific Finding: context and level of engagement with community influence survey fatigue in participants
Both operators and respondents noted that being asked the same questions month-after-month without any observed changes in WFP activities was a source of frustration with pilot participants in the DRC. A commenter suggested that the extreme situation in the IDP camps near Goma presents with a difficult but understandable situation. The panel survey took place of months where distributions were given and then suspended for several months. When the distributions resumed, there were targeted and only some respondents received rations. Frustrations with the survey could have also been related to frustrations with the circumstances and associated with the WFP programme. The operators indicated that most participants maintained a good humour during the calls, but began to joke with operators particularly about the food related questions. For example, "Operator: Have you eaten meat in the last week? Respondent: Why? Are you going to give me meat?" Other respondents asked about the reasoning behind asking the same questions repeatedly and why WFP was not using the information.

Operators estimated that about 1% of the respondents became quite hostile when called and that the 'intellectuals' refused to answer the questions after the initial period. It is beyond the scope or ability of the reviewer to interpret the sense of the word 'intellectual' here but it seemed to indicate that there was certain members of the community that were suspicious of the monitoring activity. Even with excellent communication with the community and the best efforts of WFP staff to inform the participants of the purpose of the project, there may be individuals or groups that will not participate or undermine the monitoring project. Perhaps this is an issue for further research and a potential consideration for both participant and staff protection. The rapport of the operators with participants was obvious during focus group discussions and many participants mentioned that they enjoyed speaking with the operators and being visited in the camp. WFP staff also shared the desire for beneficiaries to see and understand the direct benefits of participating in the surveys. They wanted their operations to be seen as dynamic and responsive to beneficiary needs. This supported the overall recommendation for more information sharing and two-way communication back to participants through calls.

Crosscutting issues of protection, privacy and gender aspects of mVAM pilots

Safety and protection of mVAM participants was an appropriate concern of a variety of those involved in the mVAM pilots. The unique protection challenges of data collection with mobile phones raised new questions about privacy and safety of the innovative approach with nearly everyone interviewed from pilot participants to those leading the programme. A number of precautions and discussions about how to address protection concerns were documented throughout the mVAM pilots and no doubt made good use of the experience of WFP staff that had worked previously in complex programming environments. When asked about the most valuable potential outcomes of the review, leadership of the Area Office that hosted one of the pilots immediately identified 'credibility of the data' as most important area of focus and then added that additional information on 'protection, privacy and gender aspects' of the pilots were essential to scale-up. During a subsequent focus group with pilot participants, more than half of the discussion focused on jealousy, tensions, theft and fraud experienced during the mVAM pilot. 'Perhaps protection should be a central focus of all the mVAM activities' was the natural reaction of the mVAM staff overseeing the programme when presented with these concerns from the field. The section below presents some of the specific findings of the review and then proposes some simple recommendations based on WFP policy that could help future mVAM activities address protection, privacy and safety concerns more systematically.

Context played a large role in the protection concerns identified in the pilots. Somalia had one of the highest rates of mobile phone access in Africa according to the World Bank[10] and the Food Security Monitoring baseline confirmed that about 70% of the respondents in the pilot area indeed had access to a mobile phone. This was a stark contrast to the pilot in Goma in the DRC where most pilot participants had 'never used a phone before.' Transfers, particularly of high value goods, are usually a red flag for protection concerns and have been documented in complex contexts to increase the probability of theft, fraud and tensions.[11] The review of protection challenges focuses on the pilot in Goma where mVAM distributed phones and a member of the review team was able to personally interview pilot participants and other stakeholders.

WFP staff often work in complex emergency environments and regularly confront protection issues in day-to-day programming where benefits of the programme are regularly weighed against contextual risks and the safety of beneficiaries. No doubt, that WFP and its partners have developed a great experience and variety of techniques to ensure protection against harm for disaster and conflict-affected populations with whom they work. It is clear that discussions and implementation of the mVAM pilots were informed by the experience particularly of the Area Office staff and national staff with long experience working in those specific contexts as well as a number of seasoned team members. In fact, mVAM results were even shared with the Protection Cluster in Goma. The review team was able to identify three specific instances where protection concerns were addressed in the pilots:

---

[10] **United States (US) Broadcasting Board of Governors (BBG) & Gallup.** N.d. Media Use in Somalia 2013.

[11] See **Anderson, M.B.** 1999. *Do no harm: How aid can support peace or war.* London, Lynne Rienner.

1. Before deciding on a strategy for the Goma pilot, protection concerns with regard to the distribution of phones was discussed. Although the camps were considered to be insecure and violent places, the benefits of the project were seen to outweigh the risks;
2. When discussing declining response rates in the Somalia pilot, security concerns were to be included in future focus groups; and
3. Protection of participant privacy was incorporated in two places in the SOP document for Goma.[12]

WFP states its own standards for protection in WFP's Humanitarian Protection Policy (HPP), WFP/EB.1/2012/5-B. The review team will refer to this policy for comparison of the attention to protection in the pilots and as a basis for recommendation for future mVAM activities with regard to protection. The policy specifically calls for 'incorporating protection objectives in programme design and protection concerns in all tools' to bring human rights and the Organization's humanitarian principles to the center of all activities (ibid.). It is clear that the need to focus on protection concerns and the will to make the approach safer for participants is present at all units and offices involved in the pilots. Although there is not enough specific documentation for the review team to speak about the Somalia pilot or other mVAM activities, some of the findings from the Goma pilot can inform future practice.

Distribution of phones both provided new opportunities for participants to keep themselves safe as well as exposed the participants to new and additional risks. The WFP protection policy specifically requires that WFP activities 'do not exacerbate the risks to which people are already exposed.' The distribution of phones increased risk of attacks and theft to participants, particularly for women, and raised tensions within households particularly between couples. In all, 42 phones were stolen during the mVAM pilot in Goma. In one particularly awful event, one of the blind participants was targeted for robbery. Several other reports on file documented violent attacks in the process of stealing the participant's phone. Tensions were also raised between participants and non-participants resulting in reported jealousy that motivated threats, attacks and theft. Participants also reported tensions when visiting other camps that did not have the mVAM pilot project. On a visit to a family member's camp, one participant was held by the camp managers who were suspicious of the participant having a phone. WFP staff was required to intervene to release the participant and explain why the project was not in all camps.

Project participants were not able to correctly identify the 'random selection' of pilot participants and most frequently said that the phones were distributed 'based on need.' The relatively modest number of phones distributed compared to 'the need for phones' identified by the camp leadership and a variety of camp residents was cited as the source of jealousy and tension. The lack of understanding about why one person was given a phone and another denied a phone was a key source of contention about the pilot by camp residents.

---

[12] The Goma mVAM SOP included the following 2 procedures:
- This list (of phone numbers) should under no circumstances be shared beyond WFP Goma. They are private, personal.
- Note that the data sets sent to Kinshasa and Rome will only contain the unique ID of the respondent - no phone number, name, or other identifying data.

In the SOP document, it was stated that 10% additional phones were available for the replacement of a phone loss in circumstance beyond the participant's control. During interviews, it became clear that project staff required more explicit direction on managing phone loss, malfunction and theft. The volunteer committee and other efforts such as discussions with participants were inadequate to protect participants from additional risk associated with participation in the mVAM pilot.

Although the WFP project staff wisely facilitated the formation of a volunteer support group to help build the capacity of participants to use the phones, unfamiliar technology increases risk for participants to manipulation and fraud. A focus group of participants related at least 5 stories of participants becoming victims of fraud by unscrupulous vendors of airtime credit or phone charging services. Older participants and women were specific targets for overcharging and fraud such as taking advantage of illiterate participants by under-charging credit that was purchased.

Although the SOP did address the issues of protection of participants' privacy by not distributing phone lists and removing identifying information when shared outside of the Area Office, phone lists were kept on open bookshelves in unlocked offices. Additional procedures such as keeping lists in locked cabinets, encrypting digital files, and perhaps keeping identifying/personal information and phone numbers separate at all times may be appropriate to ensure privacy protections. One WFP staff member as clearly concerned, "it is just a matter of time until someone uses this in scams." There was trepidation that sooner or later some group will begin to make calls under the pretense of a WFP phone survey for 'fishing', trafficking, taxation and other protection concerns such as intimidation. Due diligence in protecting participants' privacy, as well as proactive communication about the potential for scams, is necessary if mVAM pilots are to be scaled-up. mVAM had already drafted brief guidance on this subject, but continued vigilance and monitoring is necessary as the mobile applications at WFP expand and the reputation risk increases.

The review team would urge WFP to consider carefully the protection concerns associated with distributing phones in future mVAM activities. Although the benefits of the phones and opportunity for participation in programmes is compelling, programme activities must always first seek to do no harm. One could imagine a situation where security concerns were low and a phone could be distributed to everyone. Other than this idealized circumstance, protection concerns should be fully assessed and all identified risks mitigated in a participatory process. Involvement and ownership of the project by participants is likely to be the best way to mitigate identified risks. Addition reflection should be given to distributing phones in insecure environments, where there is low phone ownership, and the number of phones is insufficient to provide general phone access to the target population. These are clear red flags to programme staff that additional assessment and risk mitigation is necessary. If phones are to be distributed, a standard protocol to address protection concerns would include:

- A full risk assessment including a gender analysis;
- Mitigation strategy for all identified risks;
- Full guidance and training for project staff on mitigation measures;
- Monitoring of risk context and participant risk exposure;
- Capacity development of participants and community on purpose of project and safe phone use; and
- Evaluation and lessons learned to strengthen protection measures.

On the other hand, the distribution of phones had great potential to increase the safety of participants. These benefits may outweigh the risks if proper mitigation strategies are in place. The WFP policy mandates that the capacity to understand and assess context and risk should be an objective of WFP activities. There were several reports of mVAM pilot participants using their phones and the credit they received to find out about security conditions in their home communities. The phones and credit were used to find lost family members and to request small loans and gifts from distant family to help the displaced return to their places of origin.

The phones also have functionality in addition to SMS and voice calls that contributed to participant safety. Participants used the phones as torches at night. They also used the radio function to receive information. It is conceivable that the contacts could include numbers for a variety of vital services or hotlines to report abuses. The calculator could help to protect against fraud. The games, apps, and other functionality could further be exploited to increase knowledge and safe behaviours of participants. Imagination on protection functions of mobile phones could further increase their well-demonstrated contributions to participant safety.

Recommendations

(1) It is time for mVAM to fully develop its vision and mission

mVAM now has operated for more than two years and has achieved some very impressive results. It has proven the concept of cellular phone based high frequency monitoring. It is time to develop a longer term mission, vision, and resource mobilization strategy. We outline some of the steps needed below.

(2) Develop a robust ToC

As is the case with many young innovations, a clear ToC has yet to be developed. mVAM had the foresight to commission this review, reflective of its learning orientation. An important step forward is to develop a shared vision of mVAM and ToC framework.

(3) mVAM should engage in a serious dialogue with information users

This could involve an Internet survey tool to reach subscribers to the mVAM news lists as well as a more aggressive survey of potential/intended user groups. Information use is a problematic aspect of most information systems. mVAM should push its innovative approach farther into this domain in order to improve its effectiveness and impact.

(4) mVAM should develop a learning lab

mVAM has demonstrated its potential as a learning center for mobile monitoring in WFP. The development of its strategic plan should consider seriously this mission. Based on an established culture of external consultation, mVAM is well placed to promote cross-learning at interagency level and with communities of practice such as mHealth. Learning agenda could include:

- Developing a scenario typology for food security monitoring;
- Better gauging the validity of various modes of mobile data collection;
- Better gauging and responding to information needs of key stakeholders;
- Improving information products and services based upon user surveys and research;

- Leveraging client population behavioural change through mobile monitoring;
- Catalyzing and leading the CoP for mobile food security monitoring;
- Addressing the difficult question of whether or not WFP can capture valid data (should mVAM monitor or just innovate? under which conditions is mVAM appropriate/inappropriate?);
- Exploring rapid-response action improvements to complement mVAM surge capacity; and
- Capturing information important to understanding and exploring gender issues.

(5) mVAM should take care to leverage technology in favour of true beneficiary participation and include safeguard triggers for face-to-face assessments

WFP's reputation with donors, member nations, and communities is built on nothing if not its 'deep field' presence. The core of WFP's humanitarian emergency response commitment is to be at the frontline of food insecurity and assessment and monitoring are central to that mission. mVAM should take care that a reliance on technology does not further remove WFP from its ability to stay and deliver, and therefore should build in safeguards that trigger face-to-face assessments. The complementarities of mobile and face-to-face surveys should be clearly articulated with a policy that sets a clear preference for beneficiary participation. The possibility for two-way communication and information sharing has enormous potential for accountability to affected populations and beneficiary participation. These ideals should guide the future strategic direction and vision for mVAM.

(6) Increasing relevance of monitoring information

mVAM has great potential to include questions about context and risk in the monitoring activities. In fact, one of the niches that mVAM can best fill is real time shock monitoring. These could be shared with a number of partners including more two-way communication with participants about the changes in risk and context that are occurring. Further support to participants to use the phones to understand and assess the risk environment for themselves has great potential for increasing participant safety.

(7) Mainstreaming privacy and protection commitments in mVAM work

Two general recommendations may further contribute to improved protection in future mVAM activities. It is recommend that all mVAM project staff including those at Area Offices participate in the WFP standard protection training (WFP's Humanitarian Protection Policy (HPP), WFP/EB.1/2012/5-B). Secondly, mVAM research activities should include an academic partner that can provide IRB coverage of research designs and protocols. This will ensure the protection and ethical standards are ensured to the research participants. mVAM should ensure that all mVAM operations have clear guidance and systems for managing protection-related information.

(8) mVAM needs to continue to grow and evolve its strategic partnerships

mVAM has developed a number of robust partnerships within and among UN Agencies, other international public organizations and private corporate partners. This growth is and should continue to be organic. As a growing enterprise, however, mVAM should take stock of its needs for partnerships, identifying gaps as a basis for re-evaluating its current partnerships and identifying a future trajectory. We see at least two compelling areas of need:

(1) Strengthened local partnerships (CSOs). Volunteer communities that manage challenges as they arise key to success (e.g. charging, learning to use phones, drop out tracing, etc.). Local NGOs

and CBOs may be excellent partners for implementing and maintaining mVAM system. CBOs may also be essential for mitigating protection issues.

(2) Strengthened links to sources of cognitive capital such as university students or diaspora. We also believe that a strategic look at role definition in the different scenarios would be useful. For example, under which conditions might a third party be more appropriate to implement surveys? How should mVAM access mobile data subscribers in the long run?

(9) mVAM should continue to pursue its organic development approach during its expansion phase

mVAM has had an innovative organizational life to date. It is still a young and emergent organization. Flexibility, learning, evolving organizational arrangements explain its success and also are needed in its future. mVAM should pursue creative organizational arrangements as it expands.

(10) mVAM should embrace now its role in capacity building for mobile monitoring within and outside of WFP

mVAM is developing a unique set of competencies. Both as its responsibility and as a source of income, mVAM should invest in the development of tools and activities to strengthen food security monitoring. mVAM is already working on guidance documents. However, the strategies that it might consider could be innovative, consistent with its organizational origin. Leveraging information technology tools, MOOCs, dynamic CoPs all are ways that mVAM might move forward.

(11) mVAM should consider features of study design that lead to high-quality data

To address potential threats to validity (due to misreporting, attrition, testing bias), mVAM can use the substantial literature on survey methods to (1) address potential threats to validity before data collection, and (2) use approaches to identify and remedy biases in the data after collection. Because of technical, credibility and potentially ethical issues of doing research with vulnerable populations, mVAM should consider partnering with academic institutions. A combination of study designs (longitudinal approaches, randomization to survey mode), data quality tests (measuring reliability and validity), and general scrutiny of overall data collection approaches is useful for this purpose.