USING CAPI TO IMPROVE AND EVALUATE THE QUALITY OF SOCIO-ECONOMIC SURVEYS

DRAFT PAPER

PLEASE DO NOT CITE OR CIRCULATE WITHOUT THE AUTHOR'S PERMISSION

Johanna CHOUMERT NKOLO*, Caitlin SPENCE, Callum TAYLOR*
*Economic Development Initiatives (E.D.I.)
High Wycombe, United Kingdom/Bukoba, Tanzania

Contact: j.choumert.nkolo@surveybe.com

September 2016

ABSTRACT

Socio-economic research has been gradually moving from paper questionnaires to phone and electronic surveys. Computer Assisted Personal Interviewing (CAPI) is undoubtedly the most advanced progress as CAPI handles complex, multi-topic and dynamic questionnaires, linking hundreds of variables and dozens of tables. The surge in CAPI, especially in developing countries, has been allowed by the availability of cheaper ultra-mobile PCs and tablets, and great efforts on the part of developers to make a technology at the service of research and focused on the quality of the data collected.

In developing countries, a wealth of data is collected through face-to-face interviews. And a large strand of the socio-economic literature stresses measurement errors arising from these surveys (e.g. implausible responses, poor interviewer performance, etc.). Recent studies illustrate the superiority of CAPI for data collection. Indeed, using CAPI reduces many errors by adding automated routines (skips), validation checks (to detect inconsistencies), etc.

In this paper, we provide insights on how CAPI technology can be used by researchers and survey practitioners to evaluate survey measurement quality. Although, there is now a general consensus that CAPI significantly improves data quality and reduces measurement errors, research on how CAPI can be used to validate survey finding remains sparse.

To address these questions, we build on our experience in developing a CAPI software, Surveybe, and in implementing dozens of large-scale surveys in East Africa. These developments are the results of regular interactions between software developers and researchers and have led to several improvements allowing to evaluate survey measurement quality.

KEYWORDS

CAPI; Survey; Questionnaire design; Paradata; Images; Time stamps; Measurement; Quality Control; Tracking

Table of Contents

1. Introduction

Computer-Assisted Personal Interviewing (CAPI) refers to interviews conducted using handheld devices such as tablets or ultra-mobile personal computers. The questionnaire is configured using CAPI software and is loaded on to the device. The enumerator then reads the questions to the respondent, and inputs the responses directly into the device. This electronic data capture approach eliminates the need for post-interview data entry and thus allows for immediate data retrieval. Conversely, with pen-and-paper interviewing (PAPI), the responses need to be transferred to digital format through manual data entry after the interview.

The first Computer-Assisted Telephone Interviews took place in the USA in 1971, although it wasn't until 1987 that the first true nationwide CAPI survey was conducted in the Netherlands (Nichols and de Leeuw, 1996). Since then, there has been a proliferation in the use of CAPI techniques and a relative reduction in the use of PAPI. The main reasons given for this movement are (1) Improved data quality, (2) Increased speed and efficiency of data collection and analysis, and (3) Significant savings in time and cost of data collection.

Despite this trend, there has been little formal comparison of the two methods and the impact on data quality, survey length, survey costs, educational requirements of the enumerators and so on. Early research into the merits of CAPI focussed on concerns over operational issues and differences in data quality versus PAPI. More recently, the attention has turned to more detailed assessments of the technology and the effect it has on the data collection process. Many surveys around the world have made use of CAPI, and it is beyond the scope of this paper to discuss them all. However, until now, there has not been much discussion in the literature on how CAPI can be used to improve and evaluate the quality of socio-economic surveys.

This paper aims to highlight the use of CAPI in improving survey quality, thereby making an important contribution to the existing literature on data collection methods. EDI's extensive experience of using both PAPI and CAPI methods of data collection makes us well-placed to identify and compare the advantages in a development context.

In this paper, we add to the argument for more widespread use of CAPI by highlighting both the improvement in data quality that it brings, as well as its ability to better evaluate data quality. Our insights are drawn from our experience in conducting large-scale, complex surveys in East Africa for over a decade. All our surveys use our in-house CAPI software, surveybe. We have experience running data collection projects using surveybe for interviewing households, schools, health facilities, firms and government officials in both baseline and follow-up studies. As such, we have seen first-hand the

ability of CAPI approaches to both improve data quality, and to evaluate the quality of a dataset. Each of these arguments will be considered in greater detail in the following sections.

The rest of the paper is organised as follows: First, a detailed introduction of EDI and CAPI will be given, including our experience of using it. Secondly, aspects of CAPI that reduce errors during the actual survey itself will be discussed. Thirdly, features of CAPI that allow an overall evaluation of the quality of data, both during and after fieldwork, will be considered.

1.1 Introduction to EDI

EDI is a leading socio-economic data collection company and creator of the Computer Assisted Personal Interview (CAPI) Software, Surveybe. The EDI Group is made up of two sister companies based in East Africa and the UK respectively.

The first, EDI Limited is an East Africa-based, but internationally focussed research firm. Since our foundation in 2002, EDI has grown to become Tanzania's most respected research house. EDI was borne out of the desire to provide assistance towards economic and social development. EDI has a strong reputation both within East Africa and globally for high quality survey design and data collection services. We provide expert advice and support to a range of East African and International Organisations in the Public and Civil Society Sectors including (but not limited to): The World Bank, IFPRI, IRRI, ILRI, Centre for Global Development, Columbia University, Clinton Health Access Initiatives, Uganda Bureau of Statistics, DFID, DANIDA, Prime Minister's Office (PMO-RALG), Harvard University, University of Oxford, University of Copenhagen, ICAP – Columbia University, and the Amsterdam Institute for International Development.

Our sister company Economic Development Initiatives (EDI) Ltd, is a UK based developer of Surveybe, our Computer Assisted Personal Interview (CAPI) software which is used by data collection firms in over 50 countries around the world by leading academic and research institutions.

EDI has substantial experience in the design and implementation of surveys to provide data to understand poverty in all its dimensions and assist in formulating policies to fight it. At EDI, academic rigor and policy relevance complement each other. From survey design, over respondent sampling, interviewing techniques and quality control, up to data analysis and documentation, EDI has collaborated with the world's finest survey experts to established powerful best-practice techniques.

We understand that data collected locally, needs to be reported on globally. The uniqueness of EDI lies in its ability to bridge the gap between the local and the global. Our data abides to international

quality norms. We are committed to providing strong analysis through statistical figures and we enrich it with qualitative information and our excellent understanding of the local context. We understand that poverty is multidimensional concept. We have an excellent knowledge on debates regarding its measurement, while we can provide simple and relevant poverty indexes.

1.2 Introduction to Surveybe

Surveybe is a new generation CAPI software that gives Survey Managers' direct control of survey design, accurate responses that have been validated and cleaned at the point of collection, time saving efficiencies in all phases of their projects and real-time, analysis ready, data. Surveybe has been designed by a team with a blend of backgrounds in survey management, economics and commerce along with expert software developers. Creating an easy to use, comprehensive and affordable software suite that can intuitively and simply handle all manner of complexities has been the driving ethos of our work.

Surveybe allows users to design and implement both, simple, linear face-to-face surveys as well as large scale complex longitudinal or panel surveys. The resulting product allows survey managers to fully embrace the undoubted benefits of CAPI over PAPI. EDI has supported many multi-topic household panel surveys in Sub-Saharan Africa and increasingly across the globe, with users in over 52 countries, in using electronic methods for collecting data.

1.3 Introduction to CAPI

A CAPI approach has several features that make it differ from PAPI. The first is its ability to produce an almost immediate electronic dataset that is ready for review and analysis. The elimination of the data-entry stage significantly reduces the turn-around time to produce a dataset after interviews have occurred. A second feature is the automated routing structure that can be built into the questionnaire. Questions can be disabled if deemed irrelevant by previous responses or not if not intended for certain types of respondent. Furthermore, CAPI can handle complex questionnaires where routing is dependent on external data or multiple variables. A third feature is the ability to configure consistency checks into the software. These checks can flag inconsistencies, missing data or unlikely values. A fourth feature is the ability of CAPI interviews to include attachments, such as images, photographs or audio recordings.

Such features have added to the consensus that CAPI is a cheaper and faster alternative to the more traditional PAPI methods of interviewing. In many data collection circles, the discussion has now moved on from the comparative merits of PAPA versus CAPI, and on to which CAPI hardware and

software is better. The increasing availability and affordability of high-quality tablets has shifted the CAPI paradigm once more, with further benefits deriving from their portability, battery life, storage capacity and additional features such as cameras and audio recordings. In Leeuw's (2008) cost benefit analysis of CAPI, it is argued that CAPI software allows for faster 'back-end processing'. The consistency checks shorten the data processing and cleaning stage that occurs after interviews, because much of the internal validity checks take place during the interview itself. Additionally, the elimination of a data entry stage allows for a significant reduction in the delivery time of data. However, a CAPI methodology does increase the cost and time of 'front-end processes'. At the launch of the survey, more time and resources need to be spent on configuring the questionnaire in the software and programming the consistency checks. Additionally, training may need to be given to both the programmer of the questionnaire, and to the enumerators on how to use the devices and the CAPI program. Furthermore, the higher hardware costs of CAPI interviews, which require the purchase or rental of hand-held devices, add to the overall front-end costs.

However, there is now a growing body of evidence supporting the idea that CAPI is a cheaper and faster form of data collection than PAPI. For example, Leisher (2013) found that "the cost per completed interview for the tablet-based survey was 74% less than the paper-based survey average." They attributed most of this saving to the reduced need for data cleaning and lower enumerator fees. Furthermore, they found that "the average time per interview question for the tablet-based survey was 46% less than the paper-based survey average." These results were in line with at least two other authors' findings: both King et al. (2012) and Zhang et al. (2012) found that the cost savings generated by using CAPI were greater than the initial outlay on hardware and software.

The use of CAPI for surveys began in Western countries from the 1980s, particularly for marketing research. However, the application of CAPI to social surveys in developing countries has only become more widespread in recent years, both due to the availability of cheaper ultra-mobile PCs and tablets, and the improvement of CAPI software. Additionally, several recent studies have promoted the use of CAPI through findings of improved data quality compared to PAPI surveys (see Banks and Laurie, 2000; Caeyers et al., 2012; Leeuw, 2008; Rosero-Bixby et al., 2005). Despite this progress, there is still a significant proportion of developing country studies that use a PAPI approach. CAPI sceptics are concerned with the high initial costs of the hardware and software, and the time spent configuring the questionnaire (Caeyers et al., 2012).

2. Reducing Errors During the Interview

2.1. Automated Routing

Automated routing and skip patterns are an essential element of CAPI software. Routing patterns are configured to update based upon previous responses, with changes to question enablements happening instantaneously. The respondent and enumerator are guided through the questionnaire, with only relevant questions being asked, and any missing responses being flagged during the interview. This built-in routing therefore accomplishes three purposes – it restricts answers for irrelevant questions, it gives error messages for questions that are enabled but not filled in, and it also provides errors if answers are changed and there are now values within fields that should be disabled.

An example is an education module which asks about subjects taught to the child in the household. If the respondent answered that the child currently attends school, then the education module should be enabled and an error message will be displayed if any of these questions are missed. Conversely, if the respondent's child does not currently attend school, then responses to the school questions should be restricted for this interview file to prevent irrelevant data. If the respondent initially said that the child did attend school, and answered the school questions, before changing his mind and claiming that the child did not attend school, then the software will show error messages for responses in the disabled questions. The enumerator confirms which scenario is correct while still with the respondent and clears any responses which should not be included in the final data.

CAPI allows complex routing accuracy in many scenarios that would be near impossible for enumerators to complete manually. For example, some survey questions in a tracking study may only need one data point per respondent, and thus should not be enabled if the respondent has completed them in previous rounds. This is easily configured within CAPI software, by using reference data on whether the respondent has been interviewed before, and whether the module has been asked in previous rounds. For enumerators to do this using PAPI, paper reference tables would need to be compared against the respondent. Not only would this be highly time-consuming, but it would also likely be fraught with error. As such, CAPI allows for accuracy in routing scenarios, and is very effective in catching any missed questions or modules.

2.2. Consistency Checks

A further essential element of CAPI is the use of consistency checks, which alert enumerators to oddities or inconsistencies in the responses. These checks are run during the interview itself and can thus be corrected at a time where the enumerator is still with the respondent. Leeuw (2008) terms this an improvement in 'technological data quality'. It must be noted that within our Surveybe

software, the consistency checks are not run automatically, but must be started by the enumerators. Enumerators are normally trained to run the checks on each screen, and then to 'validate' the entire interview by running the checks on the entire interview file at the end. Enumerators learn that resolving issues screen by screen will result in fewer errors to resolve at the end so this incentivizes them to consistently run these checks during the interview even when they are not being observed. Supervisors or other data processing staff can then re-run these checks to identify any remaining issues.

Consistency checks can take two forms: warning messages and error messages. Warning messages can be used to signal unlikely responses, outliers, or potentially contradictory answers. Such messages ask the enumerator to confirm that the response is indeed correct, but allow for keeping the answer as is. Warning messages thus allow early identification for values not lying within reasonable boundaries or potentially inconsistent responses. Examples of unlikely responses include expenditure amounts that are too high for the study context, or asset ownership numbers that seem unlikely in a given region. Potentially inconsistent responses that could use a warning message may include a respondent who says that she completed primary school but does not know how to read. This form is the least strong consistency check because it implies a possibility that the response could be correct, but should be confirmed.

Error messages are a stronger form of consistency check, and indicate incorrect or impossible scenarios. Examples include relationship contradictions in the household, such as one household member being married to another household member who is single, or biological children with less than twelve years' age gap to their parents. These error validations can also incorporate calculations across different responses, with the calculation being used to identify the error. For example, in a time-use module, the software can add up hours in day to make sure they are less than 24, or the error can calculate total household expenditures and check this against household income.

Household consumption provides another example of a more complex calculation that can be incorporated into the checks (Caeyers et al., 2012). Items and units consumed could be matched with the corresponding calorific value, where calorie information is drawn from an imbedded reference table. This could then be totalled for the household, and the total household calories could be displayed to the enumerator. A warning or error could be programmed if the total does not lie within normal ranges, with the data thereby requiring correction or confirmation.

Panel or longitudinal checks would be extremely difficult to implement in paper-based interview approaches, but are easy to configure with CAPI. These checks can flag up responses that are too

different to previous responses from earlier rounds of data collection. For example, household member information from previous data collection rounds can be incorporated into the CAPI questionnaire, and an error can be generated if a household member's gender is reported as different, or if the reported age does not align with the previous age plus the number of years since the earlier round. If such an approach was to be used in PAPI, enumerators would be required to carry paper tables of previous household data. Comparing the answers for multiple household members across multiple variables would be highly time-consuming, subject to error, and possibly even neglected by enumerators in field.

Restriction of answer input is the strongest approach to prevent any inconsistent or impossible responses. At the simplest level, this ensures that only answers of the correct format are recorded. For example, if the questionnaire calls for the time of the interview, the software can restrict this so only time-stamps are accepted as answers. If the question asks for a number, the software restricts inputs to only numeric answers, rejecting any text or symbols. PAPI questionnaires cannot restrict answer types, and data might therefore need to be dropped if enumerators record incorrect answer types.

A more advanced version of restricting inputs is through the use of 'dynamic questions'. This refers to the updating of response options to align with previous responses, so that enumerators never have the option of selecting impossible or inconsistent responses. For example, a head teacher could select which subjects are taught at the school, and then the teachers at this school could only be allowed to select from this updated subject list which subjects they are currently teaching. Another example could be a communication module where the respondents select which sources of communication they use, and then the following questions could ask them to provide a ranking of the importance of each of their answered sources. The ranking could be a dynamic question which only displays the initial communication sources that were selected, and therefore prevents inconsistent new sources from appearing in the ranking.

In summary, CAPI allows the use of warnings, error messages and answer restrictions. These various forms of consistency checks allow immediate identification of problems in the responses while the enumerator is still with the respondent. Confirmation or correction with the respondent can thus happen immediately, thereby leading to both a reduction in errors during the interview itself, and a reduction in the length of the data-cleaning stage at the end of fieldwork.

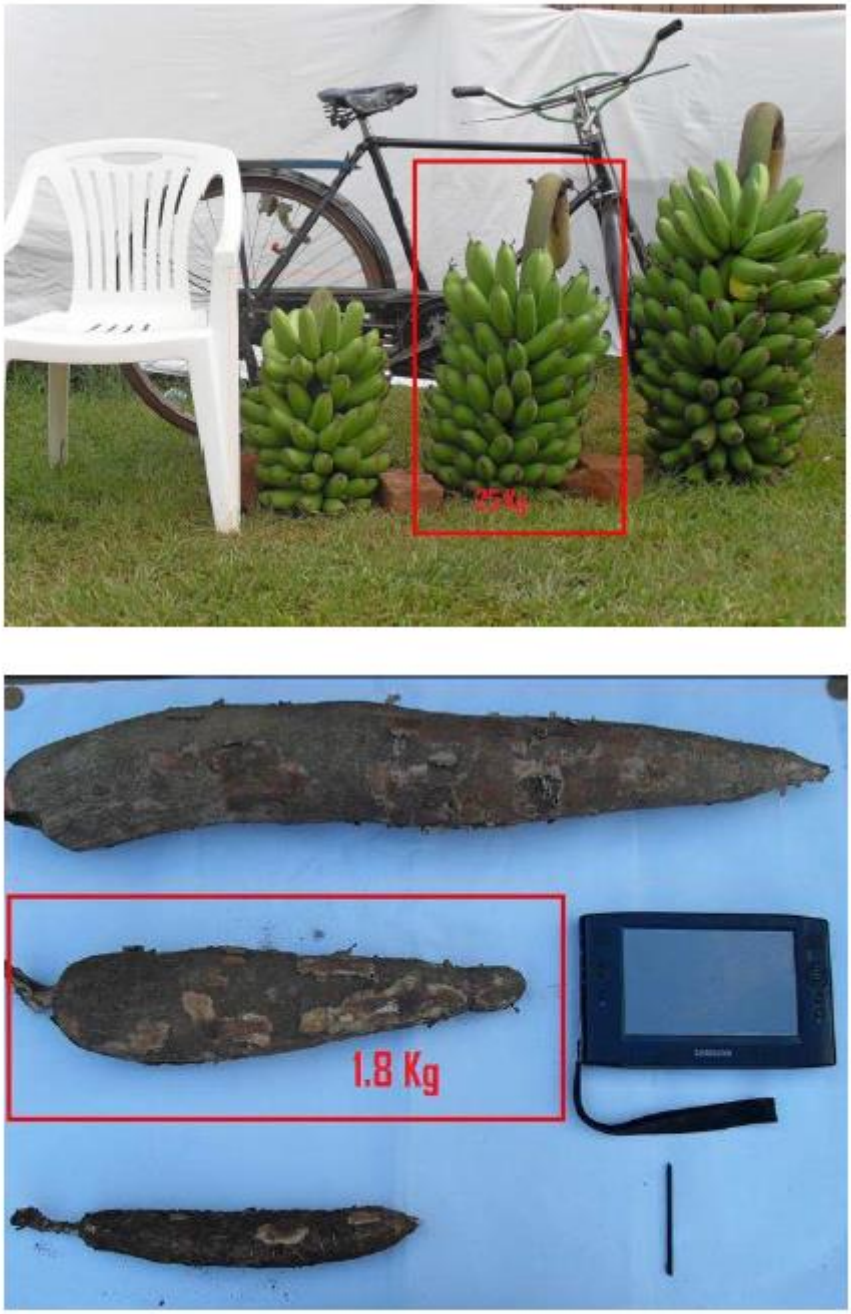
2.3. Images

Using images within a questionnaire can lead to several important improvements in the quality of data. Firstly, images can be used to standardize responses. This is particularly pertinent when referring to food-types where quantities tend to be expressed in non-standard and imprecise units, such as a 'small heap of cassava' or a 'medium bunch of bananas'. Such item-unit combinations are ambiguous and are thus subject to individual interpretation of what constitutes small, medium or large. If such terms are used, variables may be biased and could lead to under or over-reporting of consumption and may even result in incorrect conclusions about the poverty status of the household.

As such, images can be embedded in the questionnaire that specify the quantity being discussed or recorded. For example, a picture of a 'medium bunch of bananas' could be included, where the bunch is placed next to a constant-sized object for comparison purposes. Ambiguous quantities can then be more accurately mapped onto standardized metric units. Similarly, in health modules, pictures of medication may be used if the respondent does not remember the exact name of an antimalarial tablet. Images can thus play an important role in standardizing responses. See Figure 1 for examples of images that could be used.

Images also have an important role in tracking. Minimizing attrition in longitudinal studies is a key focus of researchers to ensure the validity of their conclusions, and thus tracking the original respondent is an essential aspect of follow-up studies. At baseline, the CAPI software can be used to take a photograph of the main respondent. The photographs could then be embedded into the questionnaire in follow-up surveys, thereby allowing enumerators to verify that they are interviewing the right respondent. Moreover, images can aid the tracking efforts overall with enumerators showing other community members the photograph (given this approach has been approved by the relevant ethical review board and the respondent has given consent) in order to locate the household. Photographs can also be used to reduce errors in interventions, and subsequent data collection, by ensuring that the intervention was administered to the correct respondent.

Figure 1. Examples of images that can standardize imprecise quantities



Source: Authors

2.4. Evidence on Improved Quality of CAPI Data

Many studies have asserted the improvement in data quality due to CAPI but have not empirically set out to prove this. For example, Leeuw (2008) presents a summary of the mechanisms which allow CAPI data to reflect a reduction in errors. Similarly, Banks and Laurie (2000) describe the expected data quality benefits that triggered a move from PAPI to CAPI for the British Household Panel Survey, although no rigorous study on the actual achieved benefits of this change was conducted. Rosero-Bixby et al. (2005) set out the rationale of CAPI and provide a list of twelve features that are desirable in CAPI software. This list includes skips, consistency checks, precise time and space coordinates, and external information attachments.

Caeyers et al. (2012) represent one of two studies which compare the data obtained from CAPI and PAPI surveys in a systematic way. The authors conducted a randomized experiment on Pemba Island of Tanzania, among 1840 households, with three study arms. The first arm comprised of households who were administered a consumption module by pen-and-paper. The second arm contained households who were administered it by CAPI software with automated routing, but with no consistency checks. The final study arm was the remaining households who were administered the consumption module using CAPI software with both routing and 366 in-built consistency checks.

Their results show that CAPI routing significantly reduced the number of errors per survey compared to PAPI interviews, and then the use of consistency checks further reduced the number of inconsistencies. They find that the share of questionnaires with more than one missing or impossible value was 2% for CAPI questionnaires with consistency checks, 40% for CAPI questionnaires without consistency checks, and 83% for PAPI questionnaires. Furthermore, their results show that the pattern of incorrect entries is dependent on household characteristics. Dropping these data points can thus lead to biased inferences – for example, their analysis shows that 21% of households in the PAPI arm were not eating sufficient calories per day, whilst this figure is only 8% of CAPI households. As such, the high proportion of inconsistencies in PAPI surveys compared to CAPI is an important source of mismeasurement.

A second paper which empirically considers the use of CAPI on data quality is that of (Fafchamps et al., 2012). The authors aim to measure sales and profits for microenterprises in Ghana over four rounds of data collection in 2008 and 2009. Firm profit and sales data tend to be noisy and seemingly subject to mismeasurement, with both high standard-deviations of profits compared to the mean sector profit, and low autocorrelation of sales and profits for a particular firm over time. As such, the authors implemented two types of CAPI consistency checks when collecting firm data to minimize some of the suspected mismeasurement. The first check was a cross-sectional comparison of sales

against profits for a particular firm, whilst the second check was a panel check that flagged up any large differences in sales and profits from the previous round. These checks were progressively added across rounds, with only the third and fourth round comprising of the total consistency checking system.

Their results show that these consistency checks do minimize the variation of responses, and increase the autocorrelation of sales and profits for a firm across rounds of data collection. However, they find that only 3-13% of errors end up being corrected, and thus the overall effects on the sample are modest. This suggests that variation and volatility is inherent to the microenterprise sector, and thus seasonality and shocks, rather than mismeasurement, play a larger part in this noisy data. As such, consistency checks are perhaps more limited in their ability to improve data quality for firm profit and sales, as they are in other types of surveys such as household questionnaires.

Caeyers et al. (2012) also consider the effect of other features of CAPI in improving data quality. They find that automated routing plays a significant role in the reduction of errors. Their results show that 94% of the errors in PAPI that were due to routing errors were avoided through the automated routing system of the CAPI software. Similarly Leeuw (2008), in summarizing the evidence on CAPI, argues that the research to date has all agreed that automated routing significantly improves the quality of the data, mainly by preventing missing data. Caeyers et al. (2012) also consider the impact of using images to more accurately record ambiguous quantities which can then be better mapped onto standardized metric units. They find that 42% of item-unit combinations in the PAPI questionnaires were ambiguous, and thus possible bias-inducing assumptions would be needed before any analysis could be done. The use of images to add more precision to these responses is thus an important feature of CAPI.

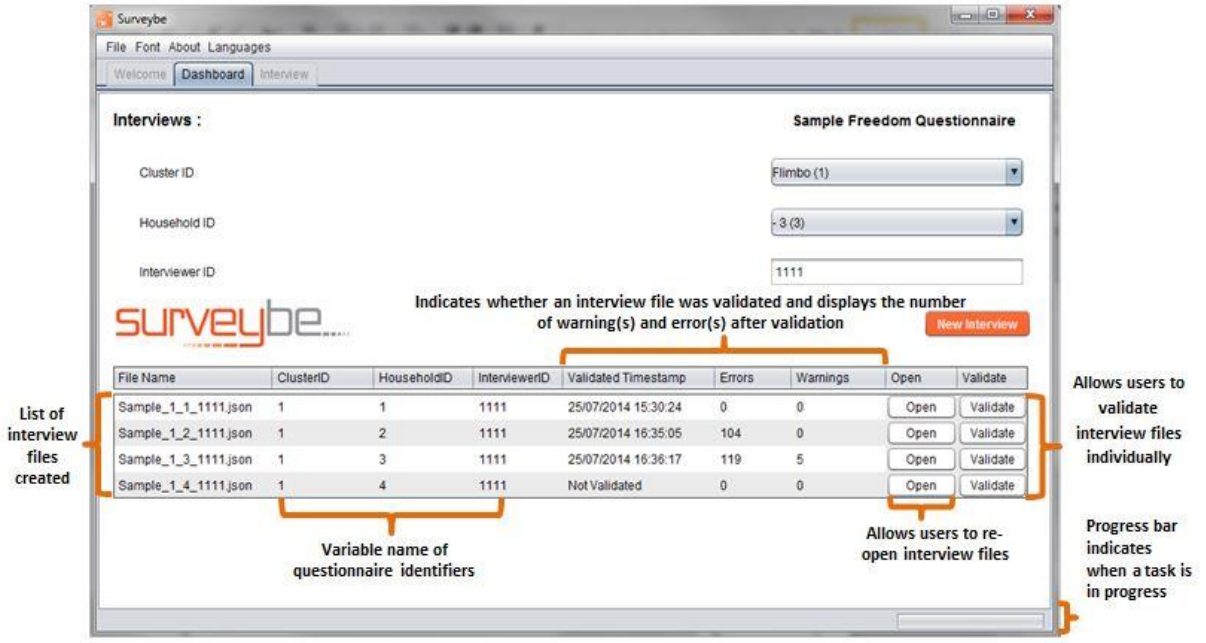
3. Evaluating Survey Measurement Quality

The previous section discussed features of CAPI that allow a reduction in errors in the data. Empirical evidence on the improvement in the quality of data using CAPI was also presented. In this section, features of CAPI that allow the evaluation and measurement of the quality of a dataset will be introduced. It will be argued that CAPI software contains several features that enable the evaluation of the quality of the data in an efficient and accurate way.

3.1. Proportion of Errors and Warnings

CAPI software allows survey managers to easily calculate the number of inconsistencies within an interview file. With surveybe, the 'dashboard' includes metadata on problems within the data, which can be easily copied into excel and other statistical analysis programs. The 'validated' heading shows if the entire interview file has had the consistency checks run. If a date for a validation is missing, it suggests that the enumerator did not check the data for inconsistencies, and is thus not following correct protocol. There are also columns showing the number of errors and warnings in the interview file. Errors should have been dealt with in an interview, and thus should be zero. There is room for some warnings in an interview file although it suggests an unusual scenario which should be confirmed. See Figure 2 to see the visual layout of the surveybe dashboard metadata.

Figure 2. Dashboard Metadata in Surveybe



Source: Authors

This metadata can be used to detect the number of problematic interview files in the raw data, such as the total number of files that did not have consistency checks run, or the total number of interview files that have non-zero errors. This could be used to evaluate the overall quality of data coming out of the field. This metadata can also be used to evaluate enumerators' quality of work, by merging in the enumerator code associated with each interview file. Comparison of the number of errors by enumerator could then be conducted. Similarly, the number of interview files per enumerator that are not validated could be assessed. Such an approach could be used to detect enumerators who are consistently submitting poor data. Further training could then be given or disciplinary action against them could then be taken.

In summary, CAPI software allows an immediate retrieval of metadata from the interview files. This metadata contains important information on whether enumerators are following correct field protocols, and whether the consistency checks are being implemented. Copying this metadata into statistical programs allows for an efficient assessment of the quality of raw data, by looking at proportion of interview files with errors, as well as the proportion of enumerators consistently submitting data with errors.

3.2. Time-Stamp Analysis

Time-stamps refer to question types which record the time at the point when this question is selected. In CAPI software, time-stamps should automatically become 'read-only' once collected and cannot be re-entered or changed by enumerators, thereby eliminating tampering with this variable. Furthermore, the questionnaire can be configured to ensure that modules are not enabled until after the time-stamp has been recorded. This approach guarantees that time-stamps cannot be added in post-interview, and must have been recorded at the correct moment. It must be noted that the time-stamp draws its information from the date and time settings of the actual hardware, and so it is important that these are set correctly prior to field launch and not altered during fieldwork.

Time-stamps provide useful metadata which can be used to check enumerator behaviour. Comparing individual interviews against the average time for interviews can be used to detect short interviews that have either been rushed, or even interviews with made-up data. This average time analysis can also be conducted by enumerator. Enumerators with an average survey duration beyond one standard deviation from the mean duration across the study may imply systematic poor performance from these field staff. Long interview times may imply that the enumerator is struggling to smoothly read questions, is not keeping respondents on track, and may unnecessarily be using up project resources if they are not able to complete the interviews in a timely manner. Conversely, short interview times may imply that the enumerator is rushing, not reading all instructions and transition statements, not

reading all response options when prompted to do so, not allowing the respondent time to think, or not probing sufficiently for responses. In either case, further monitoring, investigation, training and potential disciplinary action would then need to be considered.

Time-stamps could also be included at the beginning and end of certain sections. This would allow survey managers to check the length of important modules and detect any enumerators that could be cutting corners. For example, the length that the household roster module takes could be plotted against the number of household members. Outliers that do not fall within the generally linear pattern of this relationship may signal interviews where the enumerator did not obtain accurate answers from the respondent, and would thus need further investigation. It must be noted that for exceptionally long interviews, these section time stamps can also help researchers identify the particularly time-consuming sections when trying to reduce survey length during piloting. If a significantly long section contains less essential variables, this can be identified as a possible section to eliminate to allow the focus to be on the most essential sections of the survey relevant to the research questions.

Analysing time-stamps would be possible in PAPI questionnaires, but the accuracy of these time-stamps would be a lot less certain. Caeyers et al. (2012) found that the most common errors in PAPI interviews are impossible survey durations, which occurred in 24% of PAPI questionnaires in their experiment, but in almost no CAPI questionnaires. This is aggravated by the fact that in some African countries time is told in a different format, where the day (1 o'clock) begins at sunrise. PAPI does not include an option to add in consistency checks to flag up answers that may be recorded in the African time format. Furthermore, Caeyers et al. (2012) found that time stamp errors were positively related to the household size, which suggests that enumerators have reduced concentration when recording time with bigger households, and are thus more likely to make mistakes.

In summary, single click time-stamps in CAPI questionnaires are read-only and tamper-proof responses to simply record the date and time. These answers can then be accurately used in the evaluation of data quality by looking at the mean time of interviews across the study and by each enumerator. The time-responses in PAPI questionnaires are much more error-prone with enumerators often making mistakes in long questionnaires, particularly when Western time is told in a different format.

3.3. Back-Checks

An essential benefit of CAPI software is the ability to obtain real-time data, where the data-entry stage is avoided altogether. This benefit is made use of in the implementation of back-checks. Back-checks refer to the re-interview of a sample of respondents by senior staff members with an abridged version

of the questionnaire, containing particularly important and unchanging questions or questions that trigger new sections. Back-checks are an important form of quality control for surveys where high number of discrepancies between the original data and the re-interview data suggest that the original data may have been poorly collected, or the question was not well-understood by the respondent.

The benefits of real-time data in conducting these back-checks are several-fold. Firstly, real-time data allow immediate access to the list of respondents interviewed up to that point. This could be used to create reference data which is built into the re-interview questionnaire. The reference data could be used to randomly select the respondents who should be re-interviewed. This approach of CAPI random selection thus allows a diverse, random sample of checks across different enumerators, and cannot be manipulated by those conducting the checks. Such an approach may have been possible with PAPI, but re-interviews would be significantly delayed until after data has been entered, and the list of re-interview respondents then sampled. Delays in re-interviews are undesirable because variables tend to change over time, thereby leading to more unexplained discrepancies between the original and re-interview data.

Secondly, the real-time data also allow immediate access to the re-interview dataset. This dataset can be compared against the original dataset as soon as any back-checks have been conducted. A high proportion of discrepancies from an interview file might suggest the need for the original interview to be recollected. The number of discrepancies can also be compared by enumerator, with enumerators with high number of discrepancies across their interview files needing additional training or monitoring. Furthermore, consistent discrepancies for a particular question often suggest that the question has been misinterpreted or misunderstood, either by the enumerators or the respondents. The ability of CAPI to obtain original and back-check datasets very rapidly would allow such problems to be detected early into fieldwork, and thus ensure that questions can be re-phrased or additional trainings can be conducted before too much time has passed.

Overall, CAPI software implies real-time data, which allows immediate analysis into the quality of the data against re-interview data. Comparing the original data against the re-interview data generates a range of statistics that can be used to evaluate the survey quality, such as the proportion of discrepancies by interview file or by enumerator. The early detection of these discrepancies means that questions can be re-phrased, or additional trainings on the meaning of questions can be given while still early into fieldwork.

One final note on conducting back-checks is that the approach should not involve detailed scrutiny of individual data points as there is no “correct” answer upon which to base the assessment of quality.

Instead, the focus should be on *trends* across many interviews, either of all interviews collected up to a certain date or across enumerators. Therefore, in using the comparison of original interview data to re-interview data collected during a quality control exercise, usual statistical principles should be followed where trends across observations from a sufficient sample size rather than individual responses tell the story of data quality.

4. Conclusion

This paper has considered the use of CAPI in socio-economic surveys and sought to shed light on some of the advantages of using this method, particularly for complicated surveys. We have examined a range of factors which help to explain why CAPI has become the preferred method of choice for most data collection services, particularly in cases where the survey requires more advanced approaches, such as the use of reference data.

In this paper we have also argued that CAPI effectively and effortlessly prevents error in the data collected through its in-built automated routing, consistency checks and use of images. This significantly reduces the need for back-end processing and data cleaning as the data is more consistent and reliable. This in turn leads to cost savings for the organisations in involved and reduces the time between fieldwork and publication of final results. We expect that these benefits will continue to increase in the future as more sophisticated technology becomes available, enabling CAPI to perform an even wider range of survey tasks.

This paper then purported that CAPI better allows survey managers to assess the quality of data. An analysis on both the number of consistency errors in the raw field data and the average duration of interviews allow survey-managers to detect any problems. Additionally, an immediate analysis of back-check data against original data will also capture discrepancies. Overall, these analysis approaches allow survey managers to infer if the data is of a good quality or if further action needs to be taken, such as the monitoring of certain enumerators, or the re-phrasing of misunderstood questions.

Overall we have shown how the correct use of CAPI can results in significant improvements to the quality of the data collected, whilst at the same time reducing the time, effort, and cost associated with quality control activities. Furthermore, CAPI can also be a useful tool in the evaluation of survey quality, enabling managers to perform a range of tests of the overall data and certain enumerators if necessary. We would therefore like to encourage others to consider using CAPI where possible, and to contribute to the literature on the benefits of CAPI and potential future developments.

References

Banks, R., Laurie, H., 2000. From Papi to Capi The Case of the British Household Panel Survey. Social Science Computer Review 18, 397–406.

Caeyers, B., Chalmers, N., De Weerdt, J., 2012. Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment. Journal of Development Economics 98, 19–33. doi:10.1016/j.jdeveco.2011.12.001

Fafchamps, M., McKenzie, D., Quinn, S., Woodruff, C., 2012. Using PDA consistency checks to increase the precision of profits and sales measurement in panels. Journal of Development Economics 98, 51–57.

King, J. D., Buolamwini, J., Cromwell, E. A., Panfel, A., Teferi, T., Zerihun, M., . . . Vienneau, D. (2012). A Novel Electronic Data Collection System for Large-Scale Surveys of Neglected Tropical Diseases. *PloS One, 14*.

Leeuw, E.D. de, 2008. The effect of computer-assisted interviewing on data quality: a review of the evidence [WWW Document]. URL http://dspace.library.uu.nl/handle/1874/44502 (accessed 5.20.16).

Leisher, C. (2013). A Comparison of Tablet-Based and Paper-Based Survey Data Collection in Conservation Projects. *Social Sciences, 3*, 264-271.

Nicholls II, W. and de Leeuw, E. (1996). "Factors in Acceptance of Computer-Assisted Interviewing Methods: A Conceptual and Historical Review.", Proceedings of the Survey Research Methods Section, American Statistical Association

Rosero-Bixby, L., Hidalgo-Céspedes, J., Antich-Montero, D., Seligson, M.A., 2005. Improving the quality and lowering costs of household survey data using Personal Digital Assistants (PDAs). An application for Costa Rica, in: Meeting of the Population Association of America Philadelphia, PA, March.

Zhang, S., Wu, Q., Velthoven, M. H., Chen, L., Car, J., Rudan, I., Scherpbier, R. W. (2012). Smartphone Versus Pen-and-Paper Data Collection of Infant Feeding Practices in Rural China. *Journal of Medical Internet Research, 14*.