

Article

# Investigating Multiple Household Water Sources and Uses with a Computer-Assisted Personal Interviewing (CAPI) Survey

Morgan C. MacDonald <sup>1,\*</sup>, Mark Elliott <sup>2</sup>, Terence Chan <sup>3</sup>, Annika Kearton <sup>4</sup>,  
Katherine F. Shields <sup>5</sup>, Jamie Bartram <sup>5</sup> and Wade L. Hadwen <sup>1</sup>

<sup>1</sup> Australian Rivers Institute, Griffith School of Environment, Griffith University, Nathan, QLD 4111, Australia; w.hadwen@griffith.edu.au

<sup>2</sup> Department of Civil Construction and Environmental Engineering, University of Alabama, Tuscaloosa, AL 35487, USA; melliott@eng.ua.edu

<sup>3</sup> Monash Sustainability Institute, Monash University, Melbourne, VIC 3800, Australia; terence.chan@monash.edu

<sup>4</sup> International Water Centre, Brisbane, QLD 4000, Australia; a.kearton@watercentre.org

<sup>5</sup> Water Institute, Environmental Sciences and Engineering, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; kshields@email.unc.edu (K.F.S.); jbartram@email.unc.edu (J.B.)

\* Correspondence: morgan.macdonald82@gmail.com; Tel.: +1-416-617-1226

Academic Editor: Ashok K. Chapagain

Received: 25 October 2016; Accepted: 29 November 2016; Published: 6 December 2016

**Abstract:** The investigation of multiple sources in household water management is considered overly complicated and time consuming using paper and pen interviewing (PAPI). We assess the advantages of computer-assisted personal interviewing (CAPI) in Pacific Island Countries (PICs). We adapted an existing PAPI survey on multiple water sources and expanded it to incorporate location of water use and the impacts of extreme weather events using SurveyCTO on Android tablets. We then compared the efficiency and accuracy of data collection using the PAPI version ( $n = 44$ ) with the CAPI version ( $n = 291$ ), including interview duration, error rate and trends in interview duration with enumerator experience. CAPI surveys facilitated high-quality data collection and were an average of 15.2 min faster than PAPI. CAPI survey duration decreased by 0.55% per survey delivered ( $p < 0.0001$ ), whilst embedded skip patterns and answer lists lowered data entry error rates, relative to PAPI ( $p < 0.0001$ ). Large-scale household surveys commonly used in global monitoring and evaluation do not differentiate multiple water sources and uses. CAPI equips water researchers with a quick and reliable tool to address these knowledge gaps and advance our understanding of development research priorities.

**Keywords:** multiple household water sources; computer-assisted personal interviewing (CAPI); Pacific Island Countries; development; surveys; climate change

---

## 1. Introduction

Most surveys conducted for global water, sanitation and hygiene (WaSH) research neglect aspects of household water management that are widespread and essential in many developing country settings. Notable among these is the exclusive focus of household surveys on the “primary” source of water for drinking and cooking [1]. This has led to underrepresentation of multiple water source use in household (HH) water management and its relevance for important and timely issues in global WaSH including hygiene, household water quantity needs, and adaptation of water practices by season and climate resilience. A recent review of on-plot drinking water supplies and health found only five

studies that investigated the use of multiple water sources [2]. Furthermore, none of the five articles cited present data on how multiple water sources are managed within the home.

The importance of multiple household water sources in many developing country settings and the associated gaps in knowledge have been acknowledged recently by WaSH researchers (e.g., [3,4]). However, research on and monitoring of multiple water source use is often perceived as too difficult and time-consuming and there have been few efforts to address these issues or to introduce tools that enable research on multiple sources [2]. This study reports on the potential of a computer-assisted personal interview (CAPI) to conduct fast and accurate research on multiple water sources; and on the application of the instrument in Pacific Island Countries (PICs).

It has been well-documented that some households depend on more than one water source for their domestic water needs [5,6]; however, the concept of multiple sources and uses remains poorly understood. One of the earliest explorations of alternative water supplies appeared in *Drawers of Water I* [7], and discussed a range of available water sources in three East African countries, as well as the utility of alternative water sources in households with unreliable piped water supplies. Thirty years later, in *Drawers of Water II*, the reported rate of multiple sources was higher, and it was found that water source selection was largely dependent on the intended use, with unimproved sources commonly reserved for non-consumptive purposes [8]. This association of water sources and uses has been referred to as the “rationality factor”, to describe the value- and preference-driven selection of a particular water supply for a given domestic function [9]. Multiple water sources are commonly employed in settings without affordable access to a single continuous source of high-quality water; however the complex behaviour of how this is managed has not received adequate attention [3]. The type and perceived quality of a given water source dictates how it is used by the household, with the highest quality source commonly reserved for drinking and cooking [9–11]. However, the academic literature on this topic is surprisingly limited. In rural Vietnam, rainwater was found to be the most common water source for a variety of uses, but during the dry season it was reserved for high-value consumptive needs [11]. In Cambodia, certain households appear to value the quality of rainwater more highly than water piped to their homes, and resort to purchasing untreated river water from tanker trucks when more preferred sources are unavailable [4]. The number and acceptability of available water sources drives a complex behavioural pattern for managing water within the home, made even more complicated by seasonal differences, extreme weather events, and increasingly pronounced climate variability. Unfortunately, most conventional WaSH surveys are ill-equipped to explore these multiple water sources and their uses.

A small number of academic papers have explored the phenomenon of multiple water sources and the way that they are managed at the household level. Howard et al. (2002), pioneered a novel study that characterized different water sources and uses within the homes of low-income communities in Uganda, as well as factors that influenced source selection [12]. In 2013 a report was published that highlighted the critical role of multiple sources in providing sufficient volumes of water to communities in South Africa, Ghana and Vietnam [3]. The report included the location of water use, a technique that more accurately measures the quantity of water consumed and characterizes off-site use, but the authors did not report on the effect of seasonal change on the availability of water and the impact this has on household water management. In light of the findings of our survey administered in the Solomon Islands and the Republic of the Marshall Islands, there is evidence to suggest that seasonal change plays a significant role in the selection and application of different household water sources [13]. Other authors have impressed how household management of multiple sources is strongly linked to patterns of water quality and availability that are governed by seasonal change [11,14]. A survey instrument developed by Whittington (2000), enabled data collection on multiple water sources and uses by providing an instrument for data collection on the interconnectedness of dynamic household water sourcing and seasonal variability [15]. While the data generated by this survey is considered the gold-standard in the differentiation of multiple water sources and uses, it has not been widely

adopted [2]. It is considered too time consuming and difficult to implement because of its intricate grid-pattern framework, numerous skip sequences and extensive length.

CAPI approaches using handheld tablets or rugged laptop computers to facilitate survey delivery by enumerators, have been widely applied in other fields [16–18] and are becoming more common in WaSH research, but the majority of studies are still performed using paper-based surveys, or pen and paper interviewing (PAPI). No studies were found that employed CAPI to investigate household water management. As survey instruments grow in complexity to include new parameters such as multiple water sources, they become cumbersome and difficult to use, especially for enumerators with limited training. Skip patterns and conditional logic statements, where new questions arise or are removed from the survey on the basis of respondent responses, create confusion when enumerators are required to navigate the survey using the kind of post-script instructions belonging to PAPI methods. In contrast logic statements are algorithms built into the CAPI framework that can be used to limit numerical data entry to valid ranges only, bypass questions not relevant to the respondent, and ensure responses to mandatory fields. In CAPI, skip patterns are automated—enumerators are not required to navigate the survey themselves—and this not only simplifies survey progress, but it also reduces data entry errors and missing responses [19]. This simplification and increased efficiency is believed to increase the speed at which surveys can be administered [20,21], without compromising data collection accuracy [22]. While some studies have reported that CAPI surveys take longer to complete [23] it has been posited that this is a function of survey design and study methodology [20], which have greatly improved over time through technology development [21]. These advancements of CAPI methods and digital survey platforms, such as SurveyCTO, have made it possible to develop and implement complex WaSH surveys that were previously believed to be too time consuming and too complex to be practical using PAPI methods.

This paper reports on a novel approach to the investigation of multiple water sources using a complex household survey administered using CAPI. Our research objectives were: (a) to determine whether the transition from paper-based to tablet-based surveys led to improved time-efficiency; (b) to determine whether the time per survey continued to decline as enumerators become more proficient with the tablet-based survey; and (c) to evaluate whether the tablet-based survey delivered better quality data. Critically, this study highlights the need for novel instrumentation to resolve the global deficiency of information on multiple water sources and the complex behavioural patterns associated with household water management.

## 2. Materials and Methods

The original PAPI survey used in this study evolved from a questionnaire designed to investigate multiple water sources and uses within households [15]. We expanded the survey from 44 questions and 11 pages to incorporate elements on location of use, and the impact of extreme events such as flood, drought and cyclones on household water management. These changes substantially increased the length and complexity of the PAPI version, adding 52 questions and 3 pages (Figure 1). Seven pages of the PAPI survey had 10 rows and between 12 and 18 columns, requiring the enumerator to enter information directly into the grid format, whilst also recognizing and adhering to the survey's 70 skip patterns and 7 nested loops of questions to be repeated for each viable household water source. The CAPI survey was developed with SurveyCTO technology based on open data kit platforms (<https://opendatakit.org>), and administered using a Samsung Galaxy Note Tab 3 Lite. It was designed to mimic the protocol of the paper-based survey, such as the question grid pattern, and provide the same kind of flexibility afforded by PAPI, including providing options for adding clarifying comments.

SN	SOURCES OF WATER IN THIS VILLAGE OR COMMUNITY.	9. In the rainy season, do you use water from ...SOURCE... for the following activities? (RECORD ALL THAT APPLY)  WATER USED: (1) AT SOURCE (2) IN THE HOME (3) BOTH 1 and 2 (4) NOT USED							10. Do any of these water uses from ...SOURCE... differ in the dry season from the rainy season?  (1) YES (2) NO (>>12)	11. In the dry season, do you use water from ...SOURCE... for the following activities? (RECORD ALL THAT APPLY)  WATER USED: (1) AT SOURCE (2) IN THE HOME (3) BOTH 1 and 2 (4) NOT USED							12. Have any of your water sources ever been unsafe to drink because of:  1) HUMAN FAECES CONTAMINATION 2) ANIMAL FAECES CONTAMINATION 3) HELMINTHS (WORMS) 4) SALTY TASTE 5) OTHER: _____
		1 DRINKING	2 COOKING	3 BATHING	4 WASHING	5 ANIMALS	6 FARMING	7 OTHER		1 DRINKING	2 COOKING	3 BATHING	4 WASHING	5 ANIMALS	6 FARMING	7 OTHER	
1	Lake																
2	River																
3	Public taps																
4	Public Well <input type="checkbox"/> Electric pump <input type="checkbox"/> Rope and bucket																
5	Private well <input type="checkbox"/> Electric pump <input type="checkbox"/> Rope and bucket																
6	Rainwater (shared) <input type="checkbox"/> Communal <input type="checkbox"/> Households (#___)																
7	Rainwater (private)																
8	Natural spring																
9	Bottled water																

Figure 1. One page of the paper and pen interviewing (PAPI) survey adapted from Whittington (2000) [15], with 9 rows and 16 columns with a total of 144 potential responses.

The CAPI survey was designed to increase the quality of data collection, and facilitate its ease of use by local research staff. In order to reduce the number of data entry errors, the CAPI used embedded skip patterns that automatically triggered questions contingent on earlier responses. Dropdown lists and closed-ended questions were employed to avoid spelling ambiguities and reduce the frequency of unclear responses, and unique identifier numbers and temporal information on survey start and finish times were automated to increase efficiency. The initial investment required to construct the CAPI should not be underestimated; however, the open data kit platform employed by SurveyCTO uses a streamlined Microsoft Excel interface that reduces barriers and increases accessibility relative to other CAPI programming interfaces [20].

Six local enumerators, three in the Solomon Islands (SI) and three in the Republic of the Marshall Islands (RMI), attended three full days of detailed instruction and practical exercises. Information sessions on the importance of informed consent, operational definitions, question understanding, and survey structure formed the bulk of the training. Enumerators with less computer experience required more time to familiarize themselves with the CAPI survey. However, SurveyCTO runs on Android OS, the dominant operating system for smart phones and tablets in developing countries, and most enumerators had prior experience with the operating system. Critically, the CAPI survey reduced the need for explanation of rules around skip patterns and nested loop questions. The need to recruit enumerators fluent in multiple local dialects proved to be more important than previous computer or tablet experience. After training and prior to data collection, the PAPI survey was field tested in Nomoliki, a peri-urban community of Honiara (SI), and the CAPI survey was field tested in Jenrok, an urban community of Majuro (RMI). Adjustments were made to improve question clarity, facilitate delivery, and troubleshoot any technical issues with the CAPI.

The survey was conducted in five communities in SI and eight communities in RMI between August 2014 and November 2015. The CAPI version was implemented in three communities in

SI (households  $n = 56$ ) and eight communities in RMI (households  $n = 235$ ). Two communities in SI received the PAPI version (households  $n = 44$ ) before the study transitioned to CAPI. Only one of three enumerators from SI conducted household interviews with both PAPI and CAPI surveys. Enumerators from RMI used the CAPI survey exclusively because they were faster and easier to use and our opportunity for data collection in RMI was time-sensitive.

### 2.1. Analysis

Survey duration was automatically captured in the CAPI version, but start and finish times were not recorded as part of the PAPI protocol. Therefore, we use field notes and dated surveys to compare the mean of surveys performed per hour as a proxy measure of time-efficiency. Erroneous data points for survey duration had to be removed from the dataset for situations in which the two enumerators were working in different areas and only one global positioning system (GPS) was available. In these cases the GPS coordinates could not be collected until later in the day, generating a false result for survey duration. Data quality is assessed by the number and type of data entry errors made by enumerators, including: (a) missing responses, such as unanswered questions and incorrect navigation of skip patterns; (b) unclear responses, in which notes made by the enumerators are illegible or the language used is ambiguous; and (c) inappropriate responses, where the recorded response does not reflect the question for which it is intended.

Out of range responses, numerical entries that exceeded the range of the response field, were also classified as inappropriate responses. This is not to be confused with data outliers, which were not assessed or removed from the data set. The three error types discussed in this paper represent unusable data points that were identifiable as mistakes made by enumerators during survey administration. We had no way of verifying the legitimacy of values entered that fell within the range of the data entry field. Therefore, data entry errors that were not missing responses, unclear, or exceeding the numerical range of the response field went undetected and were not included in our analysis. Additionally, field notes entered by enumerators at the beginning and the end of the CAPI survey are used to identify points of confusion or difficulties with the survey.

Survey durations were strongly positively skewed and the Kolmogorov-Smirnov (KS) Test confirmed that they were not normally distributed. Therefore, survey duration in minutes was log-transformed and confirmed to be normally distributed by the KS Test. The “regression” function in the Microsoft Excel Data Analysis ToolPak was used to evaluate the log-transformed trend in survey duration versus the number of surveys delivered by each individual enumerator. The data for all enumerators were also pooled and the same analysis was performed. The difference in data entry error rates between CAPI and PAPI was assessed using an independent t-test in the Microsoft Excel Data Analysis ToolPak. All reported  $p$ -values are for two-tailed tests.

The number of questions in the CAPI survey varies based on certain subject responses. Most notably, the length of the survey increases with an increased number of daily household water sources and with the types of extreme events (i.e., flood, drought, cyclone) that the subject reports having experienced. Therefore, it was necessary to validate the decline in interview duration with number of surveys an enumerator delivered (see Results section) to ensure that it was not an artifact of the variability in water sources or extreme events reported. To this end, we used the “regression” function in the Microsoft Excel Data Analysis ToolPak to evaluate the trend in both number of water sources reported and number of extreme event types reported with enumerator surveys delivered.

### 2.2. Ethics Statement

The study reported here was designed to investigate the use of multiple water sources in the Solomon Islands (SI) and the Republic of the Marshall Islands (RMI), for the impact of extreme weather events on household water management. It was approved by the human research ethics committee of Griffith University (ENV/47/13/HREC.), the University of Alabama (14-OR-425) and by the Historic

Preservation Office of the Republic of the Marshall Islands (2014-01) and the National Health Research and Ethics Committee of the Solomon Islands (HRC 14/29).

### 3. Results

#### 3.1. Primary Findings of the CAPI Data Collection System

The literature on multiple water sources and uses in less developed countries is limited. Preliminary findings of the CAPI survey indicate the importance of multiple water sources in household water management in Pacific Island Countries (PICs), and the capability of CAPI as an effective tool to address these knowledge gaps. The CAPI survey instrument was verified by household interviews in PICs, which identified the routine use of more than one water source in 92.1% of households surveyed. The average number of water sources reported for each household was 2.32 in RMI and 3.14 in SI. Of the 1026 water sources reported by households from both RMI and SI, 471 (45.9%) were found to have different uses between wet and dry seasons. A more in-depth analysis of the primary results on household management of multiple water sources and uses is being prepared for academic publication (Elliott et al. in preparation).

#### 3.2. Comparison of Survey Duration CAPI vs. PAPI

PAPI protocol did not require enumerators to record survey start and finish times, so a direct comparison of survey duration with the CAPI was not possible. However, field notes and dated surveys enabled the reconstruction of daily fieldwork activities with enough accuracy to estimate the number of surveys performed per hour, while accounting for five minutes of walking time between households.

The average duration of the first ten CAPI surveys performed by each of the five enumerators in the field was 46 min and 6 s, accounting for an estimated 5 min of walking time between households, this is equivalent to 1.3 surveys per hour. This is slightly slower but not significantly different from 1.35 PAPI surveys performed per hour, equivalent to 44 min and 26 s, calculated using the first block of eight households for each of the two enumerators who conducted the paper-based surveys. However, following this initial block of eight PAPI surveys, the number of surveys completed decreased to between 1.07 (56 min 4 s,  $n = 4$ ) and 0.84 (71 min 25 s,  $n = 7$ ), in the second and third blocks, respectively; indicating that survey duration was increasing. This is opposite to the effect seen with CAPI, where aggregate survey duration was reduced by almost seven minutes in the second block of ten surveys, increasing the mean number of CAPI surveys performed per hour to 1.47 (39 min 13 s). These declines in duration continued as the enumerators delivered more CAPI surveys.

#### 3.3. Learned Efficiency and Time Saving Using CAPI

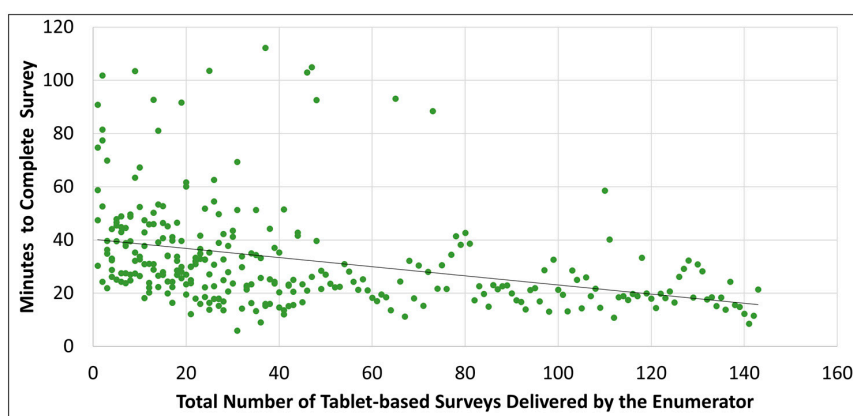
The average duration for all 291 CAPI surveys was 28.3 min (95% CI: 10.2–76.1 min). Survey duration declined with enumerator experience (i.e., the number of surveys delivered by an enumerator) Linear regressions performed on log-transformed survey durations yielded an intercept of 36.4 min and a 0.55% decline in survey duration per survey delivered. Therefore, the 100th survey would be projected to take 21 min (a 15.4 min decline). Regression results are provided in Table 1. The overall pattern of declines in survey duration with enumerator experience was also observed for each of the five enumerators and was statistically significant for all but one (Table 1). Increased enumerator experience resulted in a significant decline in survey duration ( $p < 0.0001$ ), as indicated by the pooled survey times in Figure 2. Individual data trends and regressions are plotted in the Supplementary Material (Figure S1).



**Table 1.** Linear regression results for survey duration by enumerator.

Enumerator	Tablet Surveys Delivered	Slope ( $\Delta\%$ Time/Survey Delivered)	Average Decline in Time (min) per Survey over 100 Surveys	Intercept (Time)		R <sup>2</sup> Value	p-Value
		% Decline		min	log <sub>10</sub> min		
1	49	1.48%	0.265	39.3	1.5941	0.163	0.004 *
2	143	0.73%	0.232	45.3	1.6558	0.338	<0.00001 *
3	43	1.48%	0.296	43.8	1.641	0.17	0.0059 *
4	28	0.85%	0.201	39	1.591	0.02	0.46
5	28	2.75%	0.327	42.9	1.633	0.287	0.0033 *
All	291	0.55%	0.154	36.4	1.5613	0.19	<0.00001 *

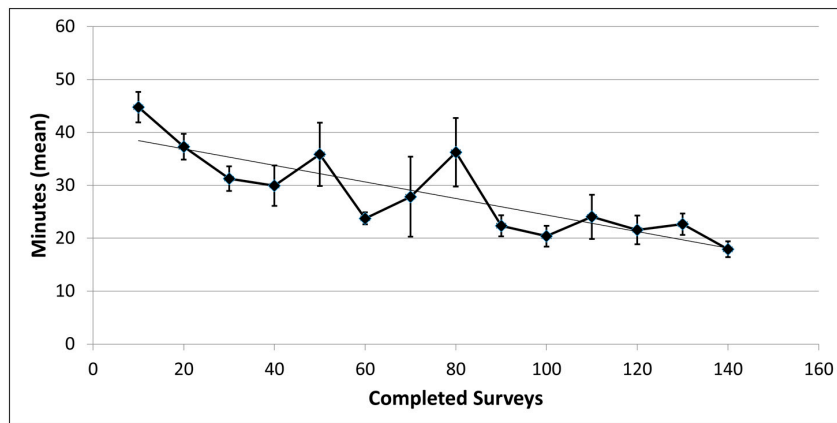
Note: \* Indicates significance at  $\alpha = 0.05$  level.



**Figure 2.** Relationship between computer-assisted personal interviewing (CAPI) survey duration and enumerator experience. A linear regression reveals diminishing survey durations associated with increasing enumerator experience achieved through practice with CAPI.

Figure 3 represents the mean duration of surveys in bins of ten, aggregated across all enumerators. Values plotted beyond 49 completed surveys belong to a single enumerator, who worked on the project longer than all of the others. Therefore, beyond 49 surveys completed the plotted values represent the mean duration of surveys performed by enumerator number two in bins of ten. Further analysis of these data reveal that 70% of the change in mean duration is predicted by the number of surveys performed ( $R^2 = 0.70$ ,  $F = 28.10$ ,  $p = 0.0002$ ). This method using binned data yielded a decline in survey duration of 0.48% per survey delivered. Observations of our data suggest that learned efficiency of the CAPI may have reached a possible asymptote at approximately 20 min per survey after approximately 90 surveys performed; less than half of the average duration for initial CAPI surveys. However, further research is needed to confirm this phenomenon and determine whether it is true across all survey types and a wider pool of enumerators.

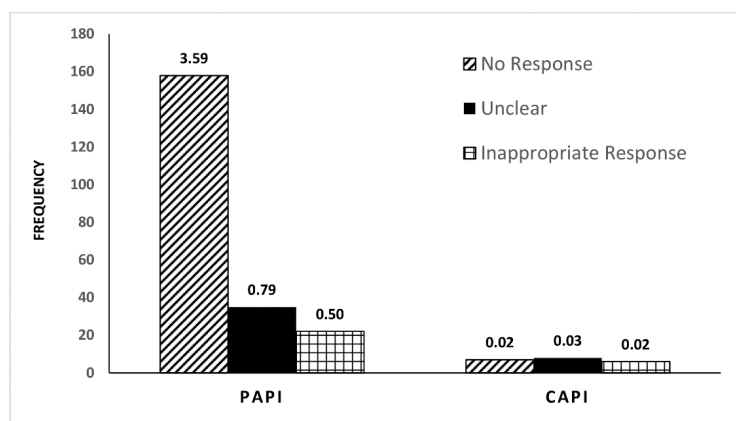
The number of questions in the CAPI survey increases as subjects report increased number of daily household water sources and extreme events (i.e., flood, drought, cyclone). It is possible that with increasing familiarity, enumerators recorded answers that they knew would not trigger additional questions, thereby biasing survey results and durations. Therefore, it was necessary to validate the decline in interview duration with the number of surveys delivered to ensure that it was not an artifact of the variability in water sources or extreme events reported. Regression analysis revealed that neither the number of water sources reported nor the number of extreme events reported declined with number of surveys delivered, indicating that the decline in survey duration was not an artifact of the latter communities having less water sources or experiencing less extreme events. Regression plots for number of water sources reported and number of extreme event types reported can be found in supplemental material in Figures S2 and S3, respectively. These results also suggest that the enumerators were not ‘gaming’ the CAPI instrument in order to complete surveys more quickly.



**Figure 3.** Average CAPI survey duration diminishes as a function of the number of surveys performed. Each point on the graph represents the average CAPI survey duration for each bin of ten surveys performed, aggregated across all five enumerators.

### 3.4. Data Collection Quality

Approximately 46,800 data elements from 291 questionnaires were entered using CAPI methods. Only 21 errors were found (error rate: 4.49 per 10,000), acceptable by the established standard of 10 per 10,000 fields for paper-based parallel data entry [24,25]. Error rates for PAPI methods were significantly higher ( $t = 11.58, p < 0.0001$ ), with 215 errors in 44 questionnaires and over 7000 entry fields (error rate: 307 per 10,000), suggesting that the transition from PAPI to CAPI led to a reduction in error rate by a factor of nearly 70. Error types and frequencies are given in Figure 4. Common mistakes found in CAPI surveys were in fields that required either a numeric or written response, as opposed to closed-ended questions. Seven surveys failed to secure a GPS coordinate, eight contained unclear written entries in a field reserved for enumerator comments, and six contained *inappropriate responses*, where the recorded response did not reflect the question posed, for water conductivity readings or GPS coordinates. The number and frequency of mistakes was controlled in CAPI surveys with logic statements that blocked certain questions and required responses to others, while data collection using PAPI surveys was more sensitive to enumerator spelling and penmanship, and more vulnerable to incorrect survey navigation.



**Figure 4.** Comparison of data entry error type and frequency between paper-based (PAPI) and computer-based (CAPI) methods. Numbers above bars indicate error rate per survey for each type of error. Diagonal lines: *No Response*—skipped question; Solid black: *Unclear*—response is illegible or ambiguous; Hatch: *Inappropriate Response*—recorded response does not reflect the question for which it is intended.



#### 4. Discussion

This paper provides evidence of the utility of CAPI to advance the understanding of multiple water sources and uses in developing country settings. For our study of rural and remote communities in two PICs, CAPI was both faster and easier to use, and led to fewer data entry errors and skipped questions. Lightweight and transportable in the field, CAPI surveys were easier to use than PAPI surveys, and had the added benefit of creating interest amongst interviewees with the use of a tablet. Time required to complete each CAPI survey decreased as a function of the number of surveys performed, with mean durations consistently less than 70% of initial survey times. CAPI, also eliminated the need for parallel data entry, a costly and time-consuming process, and avoided paper-based surveys that can be heavy and difficult to transport, as well as make field research dependent on the availability of a high-quality printer. With CAPI, completed surveys were stored in a digital file format (CSV is common) and uploaded to a secure server, ensuring data safety and making it almost instantaneously available with internet access [26].

Computer size and weight have been cited as limiting factors for survey delivery in remote fieldwork [20]; however, a 17.8 cm high-definition screen made visualization possible even in direct sunlight, and the 317.5 g tablet was easy to carry, less than the weight of ten paper surveys. Aside from being easier to carry, the CAPI offered greater data security by storing completed surveys on a web-based server, when Internet was available. Alternatively, data was stored on the tablets and backed-up daily with a laptop maintained by the field supervisor. Access to a reliable power source was difficult. The enclosed lithium-ion battery was capable of up to ten hours of fieldwork, but required periodic recharging. In remote and isolated communities throughout the Pacific, our battery charging arrangement varied according to local circumstances, and included solar panels, diesel generators, and vehicle power supplies.

CAPI can also be designed with quality control measures, such as real-time data monitoring and visualization, and logical checks. In this study, logic statements and embedded answer constraints were used in order to ensure logical data entry. Also, household interviews were monitored by an in-field supervisor to guide and facilitate remote fieldwork activities and to ensure high-quality data collection. Despite this level of supervision, there was some initial concern that enumerators could have entered responses to manipulate skip patterns, thereby reducing the length of the delivered survey. However, our results indicate that CAPI durations were not an artifact of shorter surveys, with non-significant differences between completion times of surveys that incorporated additional questions for 'extreme event' modules and those that didn't ( $p = 0.561$ ). These tests validate the reported effect of diminishing CAPI duration with increasing enumerator experience, and further reinforce the notion that CAPI is an accurate and effective tool for conducting complicated WaSH surveys.

We hope that the findings of this study will help to evolve WaSH research and monitoring by stimulating more research on multiple water source management within the home. While appropriate for many high income countries, the concept of a single water source is inadequate and unrealistic in the majority of countries, including many PICs [27]. Although knowledge gaps have been recently acknowledged by WaSH researchers [3,4], there has been little effort to address them. Complicated and time-consuming surveys have discouraged the kind of large-scale investigation of water source selection, preference, budgeting and seasonal cycling that is widespread in lower and middle income countries. The CAPI survey employed in this study provides WaSH researchers with a quick and effective tool to address this knowledge gap; for which the SurveyCTO code is available in the Supplementary Materials (Table S1: MacDonald\_SurveyCTO\_program).

We are aware of limitations to this study and offer them for discussion. First, the sample size is somewhat small and unbalanced between CAPI and PAPI datasets. The original study design was not intended to assess the value of one survey method over the other, but to advance the understanding of multiple water sources and uses in PICs, and the changes caused by seasonality and extreme weather events. For this reason, the number of surveys performed with each method is unbalanced, but the data still enabled an evidence-based discussion of the improvements in convenience and quality

of CAPI over conventional PAPI methods. Our findings concerning diminishing survey durations and the appropriateness of CAPI for multiple water source research are unaffected by the imbalance between groups. Future studies will be better equipped to confirm some of our findings by employing more enumerators and increasing the sample size. Increasing the number of surveys performed by each enumerator will also assist in the identification of an efficiency plateau, beyond which greater experience with the CAPI survey would not result in decreased survey duration. Second, with the exception of two surveys for which we had exact times, PAPI duration was reconstructed using research notes and dated questionnaires. This prohibited a direct comparison of the two survey methods, requiring the use of the mean number of surveys performed per hour in order to draw limited conclusions. However, this simple comparison revealed that PAPI duration did not decrease as a function of the number of surveys performed, as it did with CAPI. Third, because record keeping of survey duration was automated with CAPI, we cannot account for time spent on unrelated events, such as interruptions by the interviewee's children. Still, it can be assumed that the same or similar issues affected the delivery of PAPI surveys.

## 5. Conclusions

Tablet-based CAPI surveys enable the delivery of complex questionnaires for the investigation of multiple water sources and uses. Results revealed superior speed and accuracy of CAPI surveys over the more traditional PAPI method, with significantly lower error rates ( $p < 0.0001$ ) and diminishing survey durations by 0.55% decline in survey duration per survey delivered. This approach has the potential to enable large-scale regional surveying capable of characterizing multiple water sources, and generating more advanced datasets that are better equipped to inform water, climate change and development policy. Information on multiple sources, multiple uses and location of use would provide a more accurate depiction of the actual household water budget, and permit greater insight into the amount of water used for less understood domestic applications, such as the many aspects of hygiene.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2073-4441/8/12/574/s1](http://www.mdpi.com/2073-4441/8/12/574/s1), Figure S1: Relationship between CAPI survey duration and enumerator experience for individual enumerators, Figure S2: No significant relationship between number of water sources reported and enumerator experience, Figure S3: No significant relationship between number of extreme event types reported and enumerator experience, Table S1: MacDonald\_SurveyCTO\_program.

**Acknowledgments:** The Australian government is acknowledged for its support of this research through the Australian Development Research Awards (ADRA) scheme within the Department of Foreign Affairs and Trade. The authors would also like to thank all of the study participants without whom this research would not have been possible, as well as enumerators Dustin Langidrik, Hilda Tango, Trevor Palusi, Malynne Joseph, Patricia Kennedy and Adelma Louis for their exceptional work in the field and keen insights into community life.

**Author Contributions:** Morgan C. MacDonald, Wade L. Hadwen, Annika Kearton, Mark Elliott, Terence Chan and Jamie Bartram conceived and designed the experiments; Morgan C. MacDonald, Mark Elliott and Terence Chan performed the experiments; Mark Elliott and Morgan C. MacDonald analyzed the data; Morgan C. MacDonald and Mark Elliott modified and expanded the Whittington (2000) [15] survey tool; Katherine F. Shields and Jamie Bartram contributed materials/analysis tools; Morgan C. MacDonald, Mark Elliott, Jamie Bartram and Wade L. Hadwen wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Bartram, J.; Brocklehurst, C.; Fisher, M.; Luyendijk, R.; Hossain, R. Global monitoring of water supply and sanitation: A critical review of history, methods, and future challenges. *Int. J. Environ. Res. Public Health* **2014**, *11*, 8137–8165. [[CrossRef](#)] [[PubMed](#)]
2. Overbo, A.; Williams, A.R.; Evans, B.; Hunter, P.R.; Bartram, J. On-plot drinking water supplies and health: A systematic review. *Int. J. Hyg. Environ. Health* **2016**, *219*, 317–330. [[CrossRef](#)] [[PubMed](#)]

3. Evans, B.; Bartram, J.; Hunter, P.; Williams, A.R.; Geere, J.-A.; Majuro, B.; Bates, L.; Fisher, M.; Overbo, A.; Schmidt, W.P. *Public Health and Social Benefits of At-Home Water Supplies*; University of Leeds: Leeds, UK, 2013; pp. 1–61.
4. Shaheed, A.; Orgill, J.; Montgomery, M.A.; Jeuland, M.A.; Brown, J. Why “improved” water sources are not always safe. *Bull. World Health Organ.* **2014**, *92*, 283–289. [[CrossRef](#)] [[PubMed](#)]
5. Itama, E.; Olaseha, I.O.; Sridhar, M.K.C. Springs as supplementary potable water supplies for inner city populations: A study from Ibadan, Nigeria. *Urban Water J.* **2006**, *3*, 215–223. [[CrossRef](#)]
6. Bakker, K.; Kooy, M.; Shofiani, N.E.; Martijn, E.J. Governance Failure: Rethinking the Institutional Dimensions of Urban Water Supply to Poor Households. *World Dev.* **2008**, *36*, 1891–1915. [[CrossRef](#)]
7. White, G.F.; Bradley, D.J.; White, A.U. *Drawers of Water: Domestic Water Use in East Africa*; University of Chicago Press: Chicago, IL, USA, 1972.
8. Thompson, J.; Porras, I.T.; Tumwine, J.K.; Mujwahuzi, M.R.; Katui-Katua, M.; Johnstone, N.; Wood, L. *Drawers of Water II—30 Years of Change in Domestic Water Use and Environmental Health in East Africa*; International Institute for Environment and Development: London, UK, 2001; pp. 1–122.
9. Almedom, A.; Odhiambo, C. The rationality factor: Choosing water sources according to uses. *Waterlines* **1994**, *13*, 28–31. [[CrossRef](#)]
10. Madanat, S.; Humplick, F. A model of household choice of water supply systems in developing countries. *Water Resour. Res.* **1993**, *29*, 1353–1358. [[CrossRef](#)]
11. Özdemir, S.; Elliott, M.; Brown, J.; Nam, P.K.; Thi Hien, V.; Sobsey, M.D. Rainwater harvesting practices and attitudes in the Mekong Delta of Vietnam. *J. Water Sanit. Hyg. Dev.* **2011**, *1*, 171. [[CrossRef](#)]
12. Howard, G.; Teuton, J.; Luyima, P.; Odongo, R. Water usage patterns in low-income urban communities in Uganda: Implications for water supply surveillance. *Int. J. Environ. Health Res.* **2002**, *12*, 63–73. [[CrossRef](#)] [[PubMed](#)]
13. Climate Resilient WaSH in the Pacific: Multiple Household Water Sources—A Traditional Strategy for Addressing Rainfall Variability. Available online: <http://www.watercentre.org/resources/attachments/programming-brief-2> (accessed on 5 December 2016).
14. Tucker, J.; MacDonald, A.; Coulter, L.; Calow, R.C. Household water use, poverty and seasonality: Wealth effects, labour constraints, and minimal consumption in Ethiopia. *Water Resour. Rural Dev.* **2014**, *3*, 27–47. [[CrossRef](#)]
15. Whittington, D. Chapter 14: Environmental Issues. In *Designing Household Survey Questionnaires for Developing Countries—Lessons Learned from 15 Years of the Living Standards Measurement Study—Volume 2*; Grosh, M., Glewwe, P., Eds.; The World Bank: Washington, DC, USA, 2000; pp. 5–30.
16. Metzger, D.S.; Koblin, B.; Turner, C.; Navaline, H.; Valenti, F.; Holte, S.; Gross, M.; Sheon, M.; Miller, H.; Cooley, P.; et al. Randomized controlled trial of audio computer-assisted self-interviewing: Utility and acceptability in longitudinal studies. *Am. J. Epidemiol.* **2000**, *152*, 99–106. [[CrossRef](#)] [[PubMed](#)]
17. Gfroerer, J.C.; Tan, L.L. Substance use among foreign-born youths in the United States: Does the length of residence matter? *Am. J. Public Health* **2003**, *93*, 1892–1895. [[CrossRef](#)] [[PubMed](#)]
18. Barrow, W.; Hannah, E.F. Using computer-assisted interviewing to consult with children with autism spectrum disorders: An exploratory study. *School Psychol. Int.* **2012**, *33*, 450–464.
19. Bernabe-Ortiz, A.; Curioso, W.H.; Gonzales, M.A.; Evangelista, W.; Castagnetto, J.M.; Carcamo, C.P.; Hughes, J.P.; Garcia, P.J.; Garnett, G.P.; Holmes, K.K. Handheld computers for self-administered sensitive data collection: A comparative study in Peru. *BMC Med. Inform. Decis. Mak.* **2008**, *8*, 11. [[CrossRef](#)] [[PubMed](#)]
20. Caviglia-Harris, J.; Hall, S.; Mulllan, K.; Macintyre, C.; Bauch, S.C.; Harris, D.; Sills, E.; Roberts, D.; Toomey, M.; Cha, H. Improving Household Surveys Through Computer-Assisted Data Collection: Use of Touch-Screen Laptops in Challenging Environments. *Field Methods* **2011**, *24*, 74–94. [[CrossRef](#)]
21. Caeyers, B.; Chalmers, N.; De Weerd, J. Improving consumption measurement and other survey data through CAPI: Evidence from a randomized experiment. *J. Dev. Econ.* **2012**, *98*, 19–33. [[CrossRef](#)]
22. Byass, P.; Hounton, S.; Ouedraogo, M.; Some, H.; Diallo, I.; Fottrell, E.; Emmelin, A.; Meda, N. Direct data capture using hand-held computers in rural Burkina Faso: Experiences, benefits and lessons learnt. *Trop. Med. Int. Health* **2008**, *13*, 25–30. [[CrossRef](#)] [[PubMed](#)]
23. Baker, R.P.; Johnson, R.A. Computer-assisted personal interviewing: An experimental evaluation of data quality and cost. *J. Off. Stat.* **1995**, *11*, 413.

24. Neaton, J.D.; Duchene, A.G.; Svendsen, K.H.; Wentworth, D. An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. *Stat. Med.* **1990**, *9*, 115–124. [[CrossRef](#)] [[PubMed](#)]
25. Day, S.; Fayers, P.; Harvey, D. Double data entry: What value, what price? *Control. Clin. Trials* **1998**, *19*, 15–24. [[CrossRef](#)]
26. King, J.D.; Buolamwini, J.; Cromwell, E.A.; Panfel, A.; Teferi, T.; Zerihun, M.; Melak, B.; Watson, J.; Tadesse, Z.; Vienneau, D.; et al. A novel electronic data collection system for large-scale surveys of neglected tropical diseases. *PLoS ONE* **2013**, *8*, e74570. [[CrossRef](#)] [[PubMed](#)]
27. Hadwen, W.L.; Powell, B.; MacDonald, M.C.; Elliott, M.; Chan, T.; Gernjak, W.; Aalbersberg, W.G.L. Putting WASH in the water cycle: Climate change, water resources and the future of water, sanitation and hygiene challenges in Pacific Island Countries. *J. Water Sanit. Hyg. Dev.* **2015**, *5*, 183–191. [[CrossRef](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).