Numeric Paper Forms for NGOs

Gursharan Singh, Leah Findlater, Kentaro Toyama, Scott Helmer, Rikin Gandhi, Ravin Balakrishnan

Abstract—**Non-governmental organizations (NGOs) working in disadvantaged communities have a variety of data-collection and analysis needs, for example, for performing surveys or monitoring programs. Because much of this data collection occurs in environments with insufficient IT support and infrastructure, and among populations not always comfortable with technology, paper forms rather than electronic methods remain the predominant means for data collection.**

We consider the design of machine-readable paper forms for NGOs. We first examine the unique needs of NGOs that interact with underprivileged populations through interviews with eleven organizations and an in-depth investigation of one NGO's specific form-filling requirements. These explorations led to a focus on numeric forms – forms with questions requiring responses largely constrained to numbers.

We then present an experiment which evaluates how a variety of formats for numeric data would fare with users from backgrounds similar to those who might fill out such forms. Our goal was to balance the tradeoff between ease-of-use among our intended population and machine readability. Combining the results of the experiment with an analysis of machine-readability from a technical perspective, we propose the best numeric input methods for different NGO form filling requirements.

Index Terms— **machine-readable forms, paper forms, input methods, ICT for development**

I. INTRODUCTION

Monitoring, evaluation, measurement, and self-assessment are among several critical tasks for non-profit organizations working in global development. Knowing the nature of one's impact, ideally with accompanying quantitative information, allows for self-correction, reports to sponsors and potential donors, and external influence.

Most program assessment requires data collection in some form as a first step. Data collection can be tedious and expensive (in labor, time, and financial cost), and it is thus desirable to extract as much value from the effort as possible. In this regard, information technologies frequently play an important role in the storage, analysis, and display of painstakingly collected data. Microfinance institutions benefit from back-end databases that store client data [22]; healthcare institutions have need to maintain medical records; and NGOs in general can benefit from longitudinal data collected over the life of their programs.

Organizations have also considered the use of information technology for the task of collecting the data itself. Micro-finance accounting information systems [28] and healthcare information on PDAs [2] are two significant examples. The expectation is that these efforts minimize transcription of data from paper forms and allow for more rapid analysis.

There are, however, many who question the value of electronic means for data collection in poor environments. One cost-benefit analysis of mobile devices used by microfinance institutions to interact with their clients suggests that the benefits of electronic technology for data collection do not always outweigh the costs [9]. Others point out that electronic mechanisms are distrusted by populations who are used to physical evidence of transactions [23].

Paper, on the other hand, is a ubiquitous, low-cost, and well-understood medium. Even in the developed world, paper forms remain widely used for the purposes of gathering information, despite ongoing advances in digital technology. However, while it may be easy to collect data on paper, transferring that data into a format suitable for subsequent computer based storage and analysis remains a difficult problem that is currently typically resolved only by tedious manual data transcription.

In this paper, we first study how paper forms are used for data collection in the context of non-profit non-governmental organizations (NGOs) that seek to monitor their own programs. We performed interviews with eleven development-focused organizations involved in healthcare, microfinance, education, and agriculture. We then probed deeper with one of the organizations to better understand their data-collection pipeline. This investigation resulted in a proposal for data collection that uses paper for the "front-end" collection tasks, and subsequently uses a combination of scanner and PC to digitize the data. Rather than attempt to handle the myriad of possible paper forms right away, we instead focused on an important subset – forms used for collecting numeric input, which was found to be frequently used by NGOs. Although many additional factors besides input type will determine the value of such a system in a real environment, our study provides a first step towards building an effective system. We conducted a study of how

Manuscript received September 22, 2008. Revised February 20, 2009.
G. Singh was with Microsoft Research India, Bangalore, Karnataka, India (email: gursharan@alumni.upenn.edu).
L. Findlater is with University of British Columbia, Vancouver, BC, Canada (email: lkf@cs.ubc.ca).
K. Toyama is with Microsoft Research India, Bangalore, Karnataka, India (email: kentoy@microsoft.com).
S. Helmer is with University of British Columbia, Vancouver, BC, Canada (email: shelmer@cs.ubc.ca).
R. Gandhi is with Microsoft Research India, Bangalore, Karnataka, India (email: riking@microsoft.com).
R. Balakrishnan is with University of Toronto, Ontario, Canada (email: ravin@dgp.toronto.edu).

people with varying educational backgrounds responded to form-filling tasks using several different "input methods" for recording numbers. Finally we analyzed the literature on machine-readable forms, and gauged how each of the input methods would fare for subsequent automatic digitization of the data.

II. Related Work

We review four categories of research on which our work builds. First, a number of electronic devices and software tools have been designed to complement paper in information work. Second, some user studies have compared digital and paper-based methods of data entry. Third, some research has focused on using paper as an interface for systems that interact with underserved populations. And, finally, existing work on machine-readable forms will be reviewed in a separate section.

A. Complementing Paper

The importance of paper as a medium to capture data is recognized, even in technology-heavy developed-world environments: *"Rather than pursue the ideal of the paperless systems, we should work toward a future in which paper and electronic document tools work in concert and organizational processes make optimal use of both"* [25]. And, indeed, many tools have addressed this need.

Early research by Johnson et al. [13] showed how paper documents could be stored, indexed, distributed and processed by capturing them as images with scanners. Wellner's [29] classic Digital Desk system used a camera and a projector to record input and project documents.

Guimbretière's *PADD* [12] attempts to bridge the gap between the paper and the digital world. Built on top of PADD, Liao et al.'s [17] *Papiercraft* implemented a gesture interface for natural editing. In a lighter-weight approach, the Anoto pen [1] uses paper printed with unique marks and a small camera in the pen to synchronize changes on the paper to a digital version of the document. Arai et al. [3] show how a camera-assisted pen can capture words on paper by hyper-linking them to other content. Koike et al. [15] used projected Venn diagrams to record input.

Although these systems combine the strengths of paper and electronics, their benefits come with additional costs, which is a concern for NGOs.

B. Digital versus Paper

Some researchers have compared paper and digital systems for data entry, and overall tend slightly in favor of paper. Galliher et al. [8] found that people are more likely to complete paper forms than digital ones. They cite technical difficulties, as well as stolen or lost devices. On the other hand, errors of omission were more common on paper.

Shelby-James et al. [26] disprove an often-made claim that electronic data capture is more accurate than paper-based methods. They found error rates with handheld computers were over sixty times that for paper-based data entry.

C. Data Collection for the Developing World

In spite of the trend to place digital technology directly in front of underserved populations, a few researchers have focused on using paper instead. Parikh et al. [24] systematize paper-based data collection by using paper user interfaces and automated forms processing for microfinance institutions. Mackay et al. [20, 21] propose using paper on top of tablet PC displays. Parikh et al.'s [23] CAM system demonstrates use of a camera-equipped mobile phone to read bar codes printed on paper forms. Bar codes cause the phone to issue audio annotations, which instruct the user to transfer data on the paper form into the mobile phone. DeRenzi et al. [6] use a PDA based system to significantly increase adherence to medical protocols for pediatric health care. These systems all incur a device cost for the person doing the data collection.

Ellison et al. [7] stress the importance of participatory monitoring and evaluation for a small scale NGO. Chand et al.'s [5] Jadoo, is a paper-only exploration of how structured content on paper can be made easy to deal with even for people with little formal education. Our early investigations agree with this work on the value of paper, and its likely continued use by NGOs.

Our research builds on the above work and focuses on the novel problem of designing paper forms which should, on the one hand, allow easy machine readability, and on the other hand, be easy to fill in for people with diverse educational backgrounds.

III. NGOs and Forms

Over the last several years, we have both formally and informally interviewed eleven organizations – five NGOs, four microfinance institutions, and two clinics – in India with respect to their data-collection needs and experiences with both digital and paper-based forms. These organizations varied significantly in size and in their use of forms (from requiring a handful of forms filled in per month, to hundreds of thousands per year), but all had in common that they had data-collection needs and continued to utilize paper forms for this purpose, despite isolated trials with electronic data collection.

Below, we outline what we learned from these organizations, and go into some detail about the data-collection system for one organization that typifies the NGO experience with forms.

A. Overall NGO Experiences with Forms

The need for data collection is widespread among NGOs, and means for doing it efficiently and subsequently being able to analyze the results easily are, in the words of one NGO, "the need of the hour" [14]. Despite the prevalence of paper forms, however, NGOs do not seem completely satisfied with their systems (or lack thereof) for incorporating collected data in their routines.

The goals of most data collection fall into one of several categories. For surveys, the intent is simply to understand the state of a population as it exists. For monitoring and evaluation of projects, the objective is accountability, both in terms of whether NGO staff and beneficiaries are performing tasks they have committed to, or whether the intended outcomes of a project are being accomplished. For baseline surveys, the collected data helps identify various parameters and tune efforts before work on a particular project is started by an NGO, for example, the majority occupation of the population in the area.

Forms can be filled in by different people with different backgrounds, but for the most part, there seem to be two classes of people filling in forms: those who are paid by the NGO to perform data collection and those who are direct beneficiaries of the NGO's programs. The former category consists of people who are almost always literate, and typically have completed secondary school; some have advanced degrees, as well. The latter range wider in their backgrounds, but frequently have not had much formal education. Anyone asked to fill in a form will likely be literate, but there are cases where the form-fillers can just barely sign their name and write numerical digits. Furthermore, their handwriting may not be very good.

Forms involve recording of numerical data (e.g., age, price, income, dates, durations), textual data (e.g., name, free-text comments), and binary or multiple-choice questions. It is worth noting that a significant fraction of the questions asked on forms appear to be reducible to numeric information, with multiple-choice answers. The key exceptions are names and free-form comments. Also, in some cases, photographs were part of the data collected.

Overall, the repeated concerns expressed by NGOs regarding data collection were: cost, time, training, data accuracy and consistency, storage, and means of data analysis. Among those NGOs who had experimented with electronic systems, these issues were again highlighted, in addition to difficulties with infrastructure and maintenance. We discuss each of these concerns briefly below.

Cost: NGOs are run on tight budgets, and many expressed concerns even about the cost of paper and printing. (Some of the NGOs we spoke with explicitly requested a ream of "white paper" as a meaningful gift in exchange for their time.) Digital equipment, of course, can be prohibitively expensive, and even those that can afford the capital costs have difficulty with costs of maintaining technology. One interesting point here is that centralized technology in an NGO office is easier to justify than devices that must scale in proportion to the number of respondents (or equivalently, data-collection staff).

The cost of staff hours for data collection also adds up, although these costs match labor costs of the local area, and are typically very low by developed-country standards for hourly wages. Among the NGOs we consulted, data collecting staff were paid as little as US$4 for a 10-hour day of data collection. In another case, a bicycle courier was paid 15 cents per village to collect sheaves of paper records and bring them back to the NGO office.

Infrastructure and maintenance: Among the organizations that had either experience with or expressed interest in information technology (IT) systems for data collection and processing, all cited challenges with power and maintenance issues. Fig. 1 shows a person filling out a form with light from a flashlight, a frequently encountered situation.

Fig. 1. A participant fills out a form during a power outage.

Training, data accuracy, and consistency: Data-collection tasks often require some training, and systems involving digital devices only add to this need. Training is a time- and management-intensive activity that NGOs invest heavily in.

Ensuring data accuracy and consistency of routine surveys is another headache for many NGOs. Staff, who themselves are not paid particularly well, may have little motivation to perform data-collection tasks with care. In many cases, the data-collection staff itself needs to be monitored, to ensure that data is being correctly collected according to specification.

Even dedicated staff, however, often embed form-filling tasks into what can otherwise be busy daily schedules. Thus, data accuracy and consistency can suffer.

Data storage and analysis: Almost all of the NGOs we spoke with understood the value of digitizing data once it was collected. Many employed data-entry staff whose sole job was to take paper forms and convert them into digital format; others outsourced these tasks to transcription services; in some organizations volunteers handled this task. These methods are all costly – for some organizations, the costliest part of the data pipeline – or irregular, and depressingly often, we were shown stacks of paper forms that had yet to be processed; in some cases, they had been gathering dust for years.

Analysis of data most frequently involved the use of spreadsheet software as well as tables in word-processing software filled in by hand. Two microfinance institutions used custom software, which produced fixed reports once the data was entered.

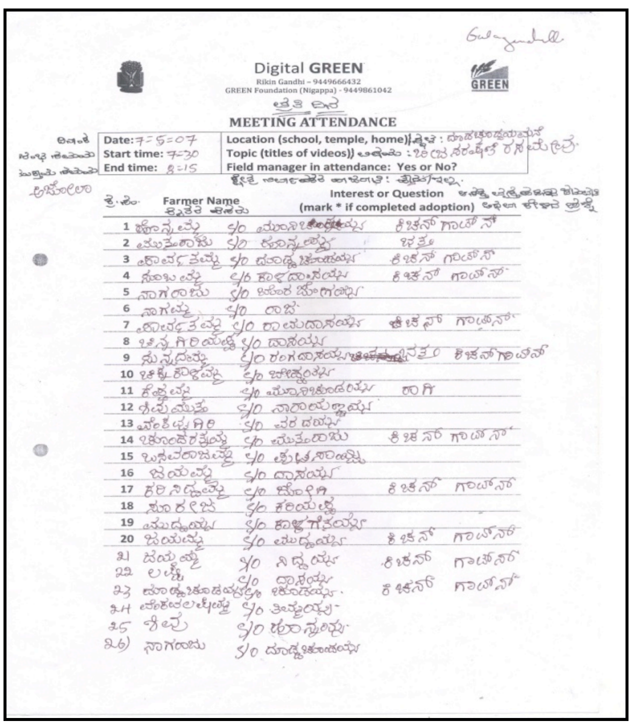
Other issues: One thoughtful NGO head mentioned that there would be some value in systems that helped organizations develop good questionnaires for data collection. He felt that a lot of data collected by NGOs was done without a clear understanding of what information was most useful for evaluation purposes [14].

B. One Particular NGO's Experience

GREEN Foundation is a small NGO based in Bangalore, India, whose mission is to spread sustainable agricultural practices among rural farmers. With an annual budget of about US$100,000, they have 15 field staff tasked with covering an operational area of approximately 100 villages.

Recently, they began a project that involves group sessions with farmers, where facilitators mediate discussion based on video content [11]. As part of the monitoring and evaluation of this project, both facilitators and farmers are requested to fill in paper forms. The program has been running for over a year, in each of 12 pilot villages, with typically three sessions per week. The first author visited these sites several times over a period of two months to better understand how the forms were used.

Their case is particularly interesting for us, because it provides information about the two broad classes of form-fillers: (1) two groups of paid staff of the NGO, all of whom are literate (senior staff with ten or more years of formal education, and junior staff, with at least eight years of education), and (2) smallholder farmers, who earned no more than $2 a day, and rarely had more than six years of education; many were all but illiterate.



Digital GREEN
Rikin Gandhi – 9449666432
GREEN Foundation (Nigappa) - 9449861042

MEETING ATTENDANCE

Date: 7-5-07
Start time: 7-30
End time: 8-15
Location (school, temple, home)
Topic (titles of videos)
Field manager in attendance: Yes or No?

Farmer Name | Interest or Question (mark * if completed adoption)

Fig. 2. A GREEN form filled by the junior staff.

After every video session, the farmers attending the session are asked questions related to the practices shown in videos, their interests, suggestions, and so on. The junior staff then summarizes feedback from every session onto a form, shown in Fig. 2. These per-session forms are analyzed by the senior staff weekly, and finally aggregated onto a monthly form. These forms are currently filled in plain text in Kannada (the local language in the region). The filled forms are then sent to the regional office. The data is entered into a database by a human transcriber and is then analyzed to spot trends and results in the villages of operation.

We observed that the form-filling sessions in villages tend to become interactive "classroom sessions" (Fig. 3) leading to vital exchange of information between the NGO staff and the participants. Forms filled by the farmers convey what they want to be taught, and forms filled by on-field staff help document their staff progress.

Overall, our key findings were that: (1) data collection and form-filling are important activities for many NGOs; (2) cost and ease-of-use are major concerns, often preventing technology-heavy systems; and (3) digitized data is desired, but digitizing data is *the* bottleneck for data-collection efforts. These findings confirm findings from earlier work and additionally identify an important problem faced by many NGOs.

Fig. 3. The verandah of a house used for meetings and data collection.

IV. Proposed Data Collection Process

A. Proposed System

Based on our findings from NGO interactions, we believe that the following combination of paper forms and hardware can solve many of the data-collection challenges outlined above, while making a good tradeoff among the constraints.

For hardware, one PC, a scanner, and a printer is required. We then envision that specially designed paper forms will be printed and used in the field to collect data. Once completed, the forms are scanned (a scanner with an auto-feeding mechanism would be ideal), and special software digitizes the filled-in content. To do this, some software innovation is required to allow those creating the forms to easily design customized, machine-readable forms. In addition, there must be software that can robustly digitize any content that is scanned.

Due to the limitations of technology for digitizing free-form handwriting, the expectation is that any given form will be implemented using multiple-choice or numeric responses as much as possible, because these responses are comparably easy to digitize. Free-form text will be given space within blank rectangles on the form, which can be scanned and stored as images (and not converted to electronic text).

The proposed system keeps costs low. No per-staff device is required, and all of the equipment is available as mass-market off-the-shelf hardware, which helps to keep costs low and alleviates technical maintenance needs. Since NGOs that collect data would like to have it in digital form eventually, the willingness to invest and maintain at least one PC is assumed (many will already have a PC). Printers and scanners add costs that are small compared with the PC itself, and paper is the only additional ongoing cost. The system would maintain all of the advantages of paper, namely its low cost, a well-understood “interface”, the lack of need for power or maintenance, and robustness in the field. Finally, the proposed system addresses the data-entry bottleneck by providing an automated means of digitizing much, if not all, of the data.

This proposal still leaves us with the following challenge: on the one hand, existing techniques for machine reading of hand-marked forms is reliable only when the forms are designed and filled out in a particular way (e.g., “bubbles” on standardized test forms) or when digits are neatly written (e.g., post-office automated digit-reading systems for sorting mail); on the other hand, the groups that we anticipate will fill in these forms are less familiar with standardized forms, and many have borderline penmanship.

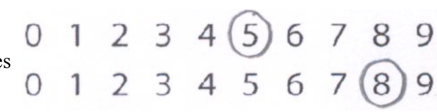
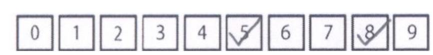
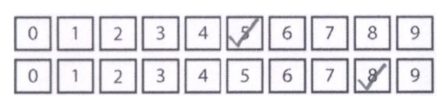
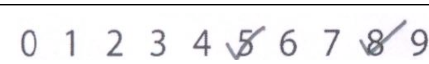
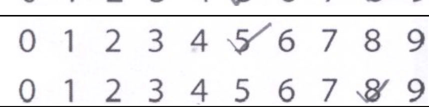
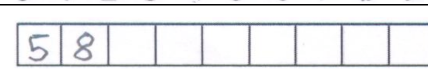
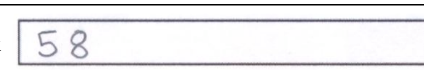
This is a non-trivial challenge that requires both engineering and interface design. In the remainder of this paper, we consider how best to handle numeric input on paper forms with the goal of achieving a reasonable tradeoff between ease of user comprehension, user accuracy, and machine readability.

V. Numeric Input Methods

We considered 10 different methods for entering numbers on paper forms, shown in Table 1. These were various versions of numeric entry, as commonly found in forms, and ranging from those that were specifically designed for machine readability (e.g., bubbles), to those that were handwritten. Eight were marking-based input methods, where the user marked the desired number either by filling in a bubble, circling the number, drawing a checkmark inside a box, or ticking the number (Table 1a-h). The marking-based methods were tested in both coded and un-coded versions, as described below. We also tested 2 handwritten methods (Table 1i-j): digit per box, where each digit was written in a separate box, and digits in one box, where all digits were written in a single box.

The un-coded marking methods provided a row for each digit, so, for example, a 2-digit number would require two rows of input (Table 1b shows the number “58” entered in un-coded Bubbles). Coded marking methods, in contrast, only provided a single row for all digits (see Table 1a). The advantage of the un-coded methods is that they allow for a complete set of numbers: for example, with 4 rows users can enter any number from 1 to 1000. The coded marking reduces the set of numbers that can be input: for example, “58” and “85” would appear the same when entered as coded input and numbers with repeating digits, such as “22”, cannot be input at all. However, the advantage of coded marking is that it greatly reduces the visual complexity of the form and saves physical space, which in turn saves paper. Coded marking could be useful for situations where the numeric data is nominal rather than scalar. If the form is used to record attendance at village meetings, for example, each individual may have an identification code assigned to them. Since these are nominal values they could be assigned with the goal of entering them as coded input.

Table I
Numeric input methods considered.

a.	Coded Bubbles	
b.	Un-coded Bubbles	
c.	Coded Circles	
d.	Un-coded Circles	
e.	Coded Checkbox	
f.	Un-coded Checkbox	
g.	Coded Ticks	
h.	Un-coded Ticks	
i.	Digit per box	
j.	Digits in one box	

VI. Technical Issues with Machine Readability of Numeric Input

Designing paper forms for automatic processing is a mature field in the developed world, with successful applications in mail sorting using postal codes (e.g., [27]), academic marking systems, and systems for processing applications (e.g., [10]). Moreover, there are numerous commercial software packages available that require little more than a PC and a document scanner. The process itself involves two stages. The first is form registration, which aligns the document so that values can be extracted. The second, and more challenging task is extracting the values from particular locations in the document. These values are extracted using either optical mark recognition (OMR) or intelligent character recognition (ICR), depending on the input method.

Processing forms in the developing world, however, adds numerous challenges. The cost and ease of deployment may increase since many commercial applications may lack support for local languages. Another challenge, based on our NGO interviews, is that since forms may be filled out in less than ideal conditions, they may be wrinkled or dirty, which can affect both the form registration and value extraction. Commercial software packages also generally assume that the end users (the individuals filling out the forms) are literate and have a certain level of education. These users are expected to have some level of competency in filling out, which is not always the case in the developing world. All of these challenges need to be weighed against the ease of use for different input methods.

The un-coded and coded checkbox and bubble input methods could be processed using OMR. In general, this method of input is nearly 100% accurate, assuming the user has followed the instructions [4]. This rate is reported for systems requiring specialized scanners and forms, which is not applicable for most NGO use; systems likely to be used by NGOs will have less accuracy. In addition, in the developing world, users will likely be less familiar with such forms, thus resulting in more user errors for OMR processing (for more detail on errors see Results section, below). Moreover, if the forms are degraded because they have not been filled out or stored in antiseptic conditions, less specialized systems may be more likely to mistake smudges, wrinkles and dirt for marks. The circling and ticking input methods we tested would also fall under OMR, but these types of marks are not generally supported in commercial software. The reason for this is possibly that the recognition accuracy is poor given the variability in the location and form of the marks, especially when contrasted to checkboxes or bubbles.

The digits in one box and digit per box input methods, on the other hand, require ICR processing, which is less accurate. ICR technology leans upon advances in machine learning that have allowed systems to learn how to discriminate characters from samples of handwriting, and is still an active area of research [18]. Currently, even the most accurate ICR systems are only capable of 98% accuracy [19], and this is the case for digits written by a more literate population than our target users. In the case of forms requiring recognition of characters of partly non-alphabetic systems, the accuracy is much worse. For example, for Kannada, an Indian language of interest to the NGOs we surveyed, the recognition rate can be less than 90% [16] and as mentioned earlier may not be supported by commercial software. It should also be stressed that accuracy rates in ICR are stated for sanitized data sampled from a completely different distribution of people. It is certain that illiteracy, poor handwriting, and form degradation will reduce the accuracy and this adds uncertainty as to whether the technology is applicable in this setting.

VII. Experiment

Our goal in designing paper forms is to achieve a reasonable balance between ease of user input and machine readability. We thus also conducted a controlled study to test user performance, accuracy and preference with respect to the 10 different input methods.

A. Participants

We recruited 40 participants who ranged in age from 17-50 years (M = 26.2). Their formal education varied from four years of schooling to undergraduate university level. Most of them spoke at least three of the following languages: Hindi, Telugu, Kannada, Sinhala, Tamil, Bengali, Konkani and English. Every participant could write in at least one language, although many had not held a writing implement since finishing school. A very basic vision test (reading a series of numbers aloud) was used to screen participants before starting the experiment. The participants came from disparate professional backgrounds: auto-rickshaw (three-wheeled mini-cab) drivers, farmers, restaurant workers, security guards, housekeeping personnel, cab drivers, army soldiers, and machine operators. Participants' literacy levels, ages, genders, and occupations were recorded. Participants were compensated with a small gift for their time.

B. Task

The task was to enter 20 numbers on a paper form, with five each of the following: 1-digit numbers, 2-digit numbers, 4-digit numbers and 8-digit numbers. Numbers to be entered were displayed on the forms themselves. A sample form for the digit per box input method is shown in Fig. 4. The numbers included in the tasks were randomly generated with no digit being duplicated in a number (for the sake of the coded input methods) and every participant saw the same set of numbers (though not in the same order). No participant entered the same set of numbers for more than one type of input method.

C. Experimental Design

A single factor within-participant design was used: each participant completed the experimental task with 10 forms, one for each input method. Presentation order was counterbalanced using a Latin square design and participants were randomly assigned to an order.

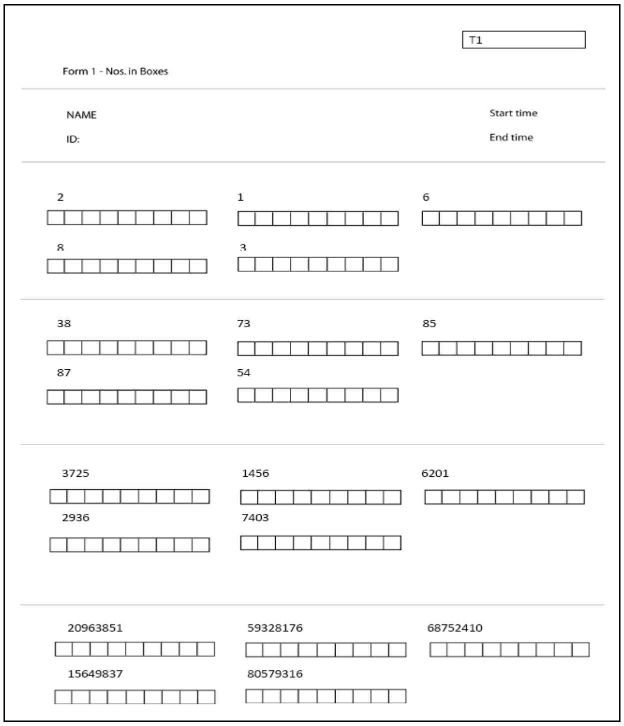


Fig. 4. Sample digit per box form. The numbers above each set of boxes are the numbers to be entered for the experimental task.

D. Measures

We measured speed and accuracy, and asked participants to provide preference feedback. Speed was measured as the time to complete the entire task for each input method and was measured manually with a stopwatch. Errors were counted after the forms were complete. A mark on a number that logically should not be there, or would obviously make a machine register the input incorrectly was counted as an error. Subjective feedback was collected for each input method using 5-point Likert scales on difficulty, confusion, and perceived speed.

E. Procedure

Experimental sessions took 45–80 minutes per participant, as they were all given as much time needed to complete or quit the task. Most of the participants were from villages in north and south Karnataka and we chose a quiet, isolated environment in which to conduct the sessions in each of these villages (Fig. 5). The remaining sessions were conducted in an office environment, where participants were called in. We ensured that there was sufficient lighting in both locations.

Participants were first given a background questionnaire to collect demographic information. Then, for each of the 10 input methods, participants were given time to examine the form and ask questions about the task. Once they were comfortable with the nature of the task, they were asked to begin. Times were recorded for each of the 4 subtasks, namely, 1-digit numbers, 2-digit numbers, 4-digit numbers and 8-digit numbers. After each input method, a questionnaire was used to collect subjective feedback. Finally, at the end of the study a short interview was conducted to compare the experience of using different input methods.

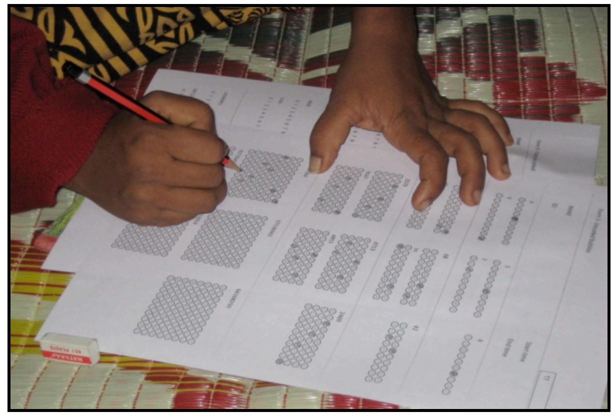


Fig. 5. A participant in a village filling a form on floor – locally known as the 'free desk'.

VIII. Results

Participants were divided into two educational groups: those with up to middle-school education (7 years or less, M = 6.1), and those with more (M = 11.9). A 2x10x10 (educational group x input method x presentation order) repeated measures (RM) ANOVA on the main dependent variable of speed showed no significant main or interaction effects of presentation order, so we simplify our analysis by only examining the effects of education group and input method (2x10 RM ANOVAs). All pair-wise comparisons were protected against Type I error using a Bonferroni adjustment. Where *df* is not an integer, a Greenhouse-Geisser adjustment for non-spherical data has been applied.

One outlier in each of the low and high education groups were excluded because they were more than two standard deviations away from the mean on the dependent variable of speed. Thus, we report on data from 38 participants (10 in the low education block and 28 in the high education block).

A. Speed

Average time to complete the forms with each input method is shown in Fig. 6. Participants in the lower education group took 31.8 minutes on average to complete all forms, while the higher education group took only 21.1 minutes on average. Note that we ran the RM ANOVA on a log transform of the speed data since the original speed data violated the homogeneity of variance assumption (significant Levene's test).

Both the education level of participants and the input method significantly affected the time it took to enter numbers into the forms (main effect of education block: $F_{1,36} = 23.3$, $p < .001$, $\eta^2 = .393$; main effect of input method: $F_{4.99,180} = 58.8$, $p < .001$, $\eta^2 = .620$).

More interestingly, some input methods were relatively better for the higher education group than for the lower education group (interaction effect between input method and education: $F_{4.99,180}$ = 2.37, p = .041, η^2 = .062). To understand which input methods were better within each group of participants, we performed pair-wise comparisons and summarize the significant results ($p < .05$) as follows.

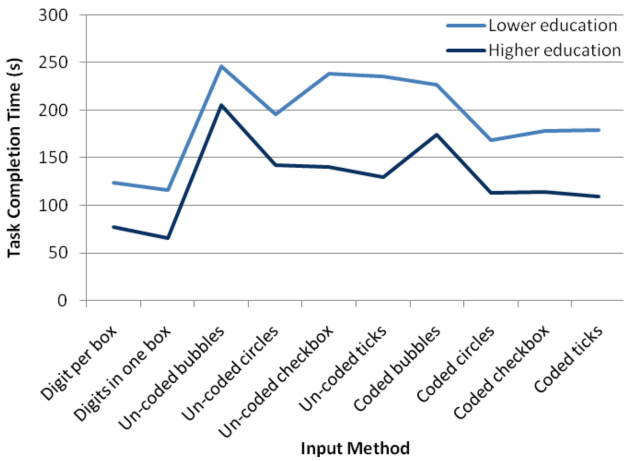


Fig. 6. Task completion time by education and input method. (N = 38)

1) The handwritten input methods were generally fast, and this result was clearest with the higher education group. These two methods were not found to be any different from each other in terms of speed. For the higher education group both handwritten methods were faster than all other methods. The lower education group, on the other hand, was not as relatively fast with digit per box: it was not found to be different from several of the coded input methods (circles, checkbox and ticks).
2) The bubble input methods were the slowest for the higher education group, but did not have as large a negative impact on the lower education group. For the higher education group, both un-coded and coded bubbles were slower than all other input methods with only one exception (coded bubbles and coded checkbox were no different from each other). For the lower education group, however, the bubble methods were not found to be significantly slower than any of the other 6 types of marking-based input methods.
3) Coded input methods were also generally faster than their un-coded counterparts, suggesting that the added complexity of having several lines of input instead of one increased the difficulty of the task. For all coded versus un-coded comparisons of ticking, checkboxes and circling in both education groups, the coded versions were faster with one exception (coded ticking for the higher education group was not faster than un-coded ticking).

B. Accuracy

Error rates were uniformly low; on average, less than 1 error was made per form (Fig. 7). Fig. 8 shows examples of entries that were counted as errors and a variety of additional anomalies (not counted as errors) that could be problematic for machine readability, as discussed in Section VI. We counted three types of errors: (1) when the number of digits in an entered number was large (four or eight) participants often found it hard to match the positions of digits and rows in un-coded versions (Fig. 8a) and marked two numbers in the same row; (2) ill-formed numbers and numbers flowing out of boxes were also problematic (Fig. 8b-c); (3) sometimes it was difficult to distinguish which number had been marked on the control (Fig. 8d).

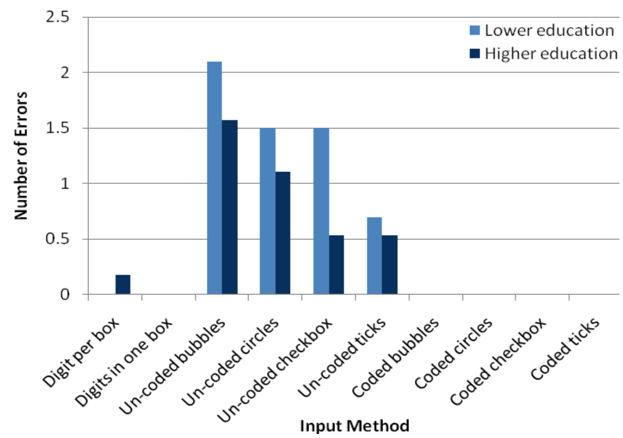


Fig. 7. Average error rate by education and input method. (N = 38)

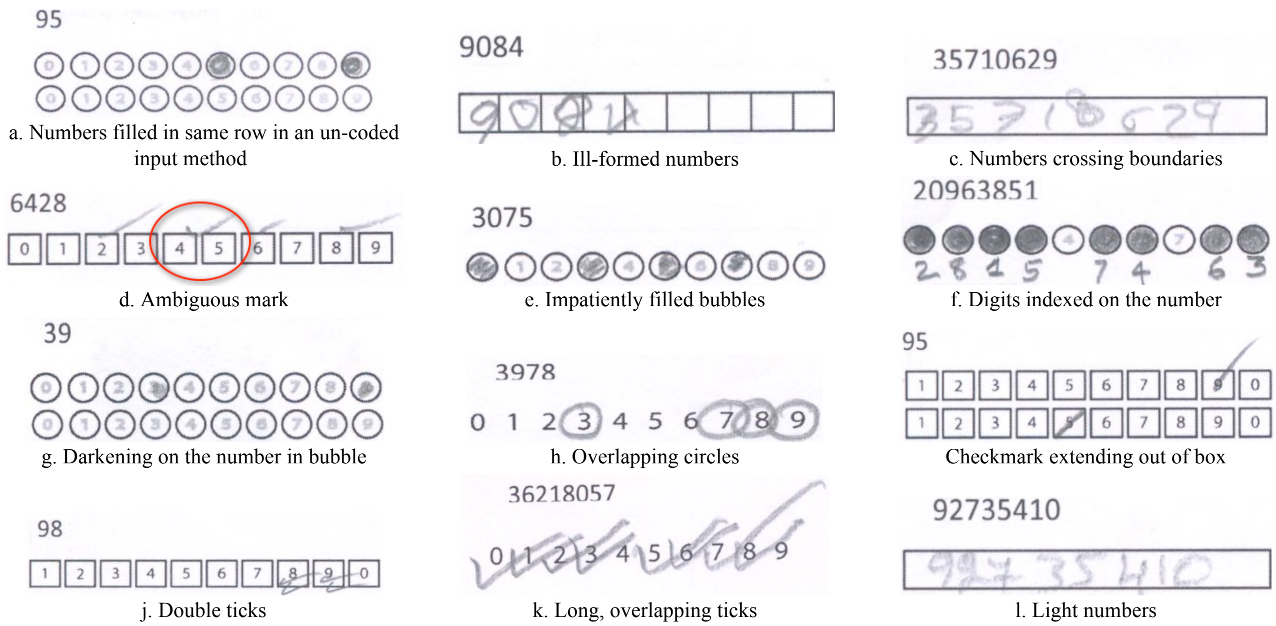
There was a significant main effect of input method on the number of errors ($F_{2.38,85.5}$ = 3.02, p = .046, η^2 = .077), but using a Bonferroni adjustment indicated that no pair-wise comparisons were significant. No main or interaction effects of education group were observed.

C. Subjective Measures

Difficulty: Participants generally found all methods easy to use: on a 5-point scale from easy to difficult the average overall rating was 2.0 (*SD* = 0.8). Some methods were felt to be more difficult (a 2-way RM ANOVA showed a main effect of input method: $F_{5.12,184}$ = 3.94, p = .002, η^2 = .099). Pair-wise comparisons show that the digits in one box method was less difficult than un-coded bubbles and un-coded ticking. No other significant main or interaction effects were found.

Confusion: Participants generally claimed not to find the input methods to be confusing: the average rating on a 5-point scale was 4.2 (*SD* = 1.8), where 5 represented “not at all confusing”. A 2-way RM ANOVA revealed that some input methods were more confusing than others, but pair-wise comparisons showed that the only statistically significant comparison was that numbers in a box was less confusing than un-coded ticking (p < .013). Though not statistically significant, digits in one box received the best rating from participants (*M* = 4.8, *SD* = 0.6). No other significant main or interaction effects were found.

Perceived speed: In general, perceived speed matched actual speed. There was a main effect of input method on perceived speed ($F_{6.22,224}$ = 5.17, p < .001, η^2 = .126). Pair-



a. Numbers filled in same row in an un-coded input method; b. Ill-formed numbers; c. Numbers crossing boundaries; d. Ambiguous mark; e. Impatiently filled bubbles; f. Digits indexed on the number; g. Darkening on the number in bubble; h. Overlapping circles; Checkmark extending out of box; j. Double ticks; k. Long, overlapping ticks; l. Light numbers

Fig. 8. Examples of errors (a-d) and other anomalies.

wise comparisons showed that both the handwritten types (digit per box and digits in one box) were perceived to be faster than coded bubbles ($p < .05$). Digits in one box was perceived to be the fastest, since participants also felt it was faster than two of the un-coded input methods (bubbles and checkbox). Participants in the lower education group rated the tasks as slower than those in the higher education group, a result that matched actual speed (main effect of education block on perceived speed: $F_{1,36} = 7.29$, $p = .010$, $\eta^2 = .168$). No significant interaction effect was observed between input method and education group.

D. *Qualitative Observations*

We made a number of observations that, though not directly the focus of the study, illustrate practical issues that may be encountered with paper-based forms in this context. Several of the less educated participants expressed great enthusiasm with the study. They savored holding a pencil since they had some formal education but their jobs never required them to write or read anything. Other participants were more apprehensive about holding a pencil because they did not know how to write, but they agreed after being encouraged by earlier participants.

A few participants without computer backgrounds felt a natural inclination to add leading zeros where there were unused "slots". For example, in a preliminary investigation prior to the formal tests, they marked the month of February in the date field as "02", filling both the input columns, instead of only "2" and leaving a column blank.

Some circumstances of rural life affect the form-filling task in unexpected ways. In one case, we experienced a power outage during a preliminary field experiment (again prior to our formal trials). One of the participants, who was over 50 years of age, found it particularly difficult to continue, although his vision was fine in bright light.

One participant commented that she depended on her husband for filling out textual information since she did not know how to write text. However, she was comfortable with numbers, and felt that if forms only required numeric entry, she could do it herself.

E. *Summary of Results*

In terms of speed, degree of formal education impacted the effectiveness of individual input methods. In particular, both un-coded and coded bubbles were the slowest input methods for the higher education group but were no slower than any other input method for the lower education group. We had anticipated that the handwritten methods (digits in one box and digit per box) would be relatively faster for the higher education participants than the lower education ones, but both groups of participants, it turned out, were faster with these than with most other input methods. Coded input methods were also generally faster than un-coded methods and resulted in fewer errors. As expected, higher education participants completed the task more quickly than lower education participants.

Note that since the lower education group had one third as many participants as the higher education group, there was less statistical power to detect differences among the input methods; this could explain why some pair-wise comparisons were significant for the higher education block but not for the lower education block.

IX. Discussion

Combining our understanding of the state of the art for machine-readability of numeric forms with the user study, we find that we can make the following recommendations for typical NGO conditions.

At the highest level of granularity, if all of speed, machine accuracy and user accuracy are desirable characteristics, coded checkboxes seem to be the best fit irrespective of education level of users. If the users are literate, digits in one box might be a better choice as it was the fastest method for the more highly educated group, with low error rate (zero for our participants), and if carefully written these digits have reasonable machine readability characteristics.

The relative performance differences between the higher education and lower education groups suggest that there may be a tradeoff between the benefits of structure and simplicity for different levels of education. The non-handwritten methods may be useful for low literacy participants since they do not require the same writing proficiency as the techniques that require handwritten numerals (digits in one box and digit per box). Our lower education participants had on average 6 years of schooling. In our experience from interviewing NGOs, we found that some users of their forms may have even less education, which could magnify the relative performance differences we found based on education. This highlights the need to consider the educational background of the target user population.

Although coded methods were in general quicker to fill than their un-coded counterparts, and also resulted in fewer errors by participants, they are not necessarily suitable for all types of forms. In particular, coded methods cannot handle numbers with repeating digits. As such, if a form has some fields that require entry of such numbers, it might be best for the sake of consistency to stick to an un-coded method throughout the entire form.

Interestingly, participants rated handwritten numbers best overall. Both digits in one box and digit per box methods were reported to be less confusing and easy to fill. In contrast, the bubble methods were not particularly favored, and were the slowest.

Taken as a whole, our results indicate that the proposed system of using structured paper forms with subsequent automated scanning is likely viable. However, the results also caution that there are nuanced tradeoffs that need to be made when choosing the type of input method to use. The fastest and most preferred techniques (handwritten numerals) are generally the least accurately recognized by a machine, although the digit per box method is arguably a reasonable compromise in this regard since the box provides some structure that eases the task of the recognition algorithm.

X. Conclusion and Future Work

We have explored the problem space of data collection by NGOs in underserved communities where resources and technological literacy are typically scarce. In particular, we focused on how a particular class of paper forms – those that require only numeric data entry –might be best structured to support subsequent automated data transfer to enable computerized storage and processing. The main contributions of our work are: (1) an interview-based analysis of challenges faced by NGOs around data collection and form filling, (2) a proposal for a semi-automated system that uses paper as the interface for data collection, and a scanner and PC for digitization, (3) an empirical study with 40 participants of how 10 different number-based input methods fare among people with both lower and higher educational backgrounds, and (4) an analysis of how easily the forms filled using each of these input methods can be automatically digitized.

Our main findings are that users tend to prefer and perform best with techniques that require handwritten numbers. However, these are not the most easily recognized by a machine. If multiple factors such as speed and accuracy of entry, and machine readability are taken into account, coded checkboxes are likely the best option, although choosing a coded method limits the type of numbers that can be captured (e.g., numbers with repeating digits are not easily handled by the coded techniques we considered). Another key finding is that the educational background of the user can significantly impact performance with different input methods, and as such the demographics of the target user population should be carefully considered when making decisions as to the type of input method to use.

In our study, participants used the same input method across each form. Given the tradeoffs we found with respect to coded and un-coded methods, it would be interesting to probe further into how people might perform with forms that have a mixture of input methods, each optimized for the types of questions being asked on the form. It is unclear if having more than one input method on a form would be overly confusing to users, particularly those with limited levels of literacy.

Finally, there are many other challenges that will need to be addressed in building an effective system to collect and process data using paper forms as a front-end, and automated input recognition as a back-end. Different types of data can pose different problems. Our work with only numeric data is a first step in this direction. It would be interesting to actually build a system that handles the processing of forms in real use, and to study its performance in the field. We intend to explore this in the near future with a partner NGO.

References

1. Anoto, http://www.anoto.com
2. AED Satellife, http://www.healthnet.org
3. Arai, T., Aust D., and Hudson S.E. (1997). PaperLink: a technique for hyperlinking from real paper to electronic content. *Proc. CHI '97*, p. 327-334.
4. Bergeron, Bryan P. (1998, Aug.). Optical mark recognition. *Postgraduate Medicine*. 104(2).
5. Chand, A., Dey, A.K. (2006). Jadoo: a paper user interface for users unfamiliar with computers. *Proc. CHI 2006.* p. 1625-1630.
6. DeRenzi, B., Lesh, N., Parikh, T., Sims, C., Mitchell, M., Maokola, W., Chemba, M., Hamisi, Y.,

Schellenberg, D., Borriello, G. (2008). e-IMCI: Improving Pediatric Health Care in Low-Income communities. *Proc. CHI 2008.* p. 753-762.
7. Ellison, M. Firm foundations: Improving NGO data collection, http://www.id21.org/id21ext/s7bns1g1.html
8. Galliher, J.M., Stewart, T.V., Pathak, P.K., Werner, J.J., Dickinson, L.M., Hickner, J.M. (2008). Data Collection Outcomes Comparing Paper Forms With PDA Forms in an Office-Based Patient Survey. *Annals of family medicine*. *6*. p. 154-160.
9. Gogineni, M., Ratan, A.L. Evaluating the viability of a mobile phone-based, SMS/GPRS-enabled, client data collection channel for urban microfinance. *Microsoft Research India Technical Report.*
10. Gravic, http://www.gravic.com/remark/remarkcustomers.html
11. Green Foundation, http://www.greenconserve.c*om*
12. Guimbretière, F. (2003). Paper Augmented Digital Documents. *Proc. UIST 2003*. p. 51-60.
13. Johnson, W., Jellinek, H., Klotz, L., Rao, R., Card, S.K. (1993). Bridging the paper and electronic worlds: the paper user interface. *Proc. CHI '93*. p. 507-512.
14. Khan, Aqueel. ASKIndia, private communication, 2008.
15. Koike, H., Sato Y., Kobayashi Y., Tobita H., and Kobayashi M. (2000). Interactive textbook and interactive Venn diagram: natural and intuitive interfaces on augmented desk system. *Proc. CHI 2000.* p. 121-128.
16. Kunte, S.R., Samual, R.D.S. (2007). An OCR System for Printed Kannada Text Using Two - Stage Multi-network Classification Approach Employing Wavelet Features. *Proc. ICCIMA 2007.* p. 349-353.
17. Liao, C., Guimbretière, F., Hinckley, Hollan, J. (2008). Papiercraft: A gesture-based command system for interactive paper. *ACM Transactions on Computer-Human Interaction*. 14(4). p. 1-27.
18. Liu, C., Fujisawa, H. (2008). Classification and Learning Methods for Character Recognition: Advances and Remaining Problems. *Studies in Computational Intelligence, SpringerLink.*
19. LeCun, Y., Bengio, Y., Bottou, L.and Haffner, P. (1998). Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE.* 86(12). p. 2279-2324.
20. Mackay, W.E., Pagani D.S., Faber L., Inwood B., Launiainen P., Brenta L., and Pouzol V. (1995). Ariel: augmenting paper engineering drawings. *Proc. CHI '95*. p. 421-422.
21. Mackay, W.E., Pothier G., Letondal C., Bøegh K., and Sorensen H.E. (2002). The missing link: augmenting biology laboratory notebooks. *Proc. UIST 2002*. p. 41-50.
22. Mifos, http://www.mifos.org.
23. Parikh, T.S, Javid, P., Sasikumar K., Dhosh, K., Toyama, K. (2006). Mobile phones and paper documents: evaluating a new approach for capturing microfinance data in rural India. *Proc. CHI 2006.* p. 551-560.
24. Parikh, T., Aditya, V.P.S, Vellayutham, M. (2003). Automated Forms Processing and Paper User Interfaces for Data Collection from Village Microfinance Groups. Unpublished draft, http://people.ischool.berkeley.edu/~parikh/papers/itira2003.pdf
25. Sellen, A.J., Harper, R.H.R. (2001). The Myth of paperless office. *MIT Press*.
26. Shelby-James, T.M., Abernethy, A.P., McAlindon, A., Currow, D.C. (2007). Handheld computers for data entry: High tech has its problems too. *Trials*. 8(5). http://trialsjournal.com/content/8/1/5
27. Siemens, http://www.industry.siemens.com/postal-automation/
28. SKS Microfinance http://www.sksindia.com/techno-logy.htm
29. Wellner, P. (1993). Interacting with paper on the DigitalDesk. *Communications of the ACM*. 36(7). p. 87-96.