

Exploring the Feasibility of Video Mail for Illiterate Users

Archana Prasad¹, Indrani Medhi¹, Kentaro Toyama¹, Ravin Balakrishnan²

¹Microsoft Research India
196/36, Scientia, 2nd Main Road, Sadashiv Nagar
Bangalore, India

²Department of Computer Science
University of Toronto
Toronto, Ontario, Canada

indranim@microsoft.com, kentoy@microsoft.com, ravin@dgp.toronto.edu

ABSTRACT

We present work that explores whether the asynchronous peer-to-peer communication capabilities of email can be made accessible to illiterate populations in the developing world. Building on metaphors from traditional communication systems such as postal mail, and relevant design principles established by previous research into text-free interfaces, we designed and evaluated a prototype asynchronous communication application built on standard email protocols. We considered different message formats – text, freeform ink, audio, and video + audio – and via iterative usage and design sessions, determined that video + audio was the most viable. Design alternatives for authentication processes were also explored. Our prototype was refined over three usability iterations, and the final version evaluated in a two-stage study with 20 illiterate users from an urban slum in Bangalore, India. Our results are mixed: On the one hand, the results show that users can understand the concept of video mail. They were able to successfully complete tasks ranging from account setup to login to viewing and creating mail, but required assistance from an online audio assistant. On the other hand, there were some surprising challenges such as a consistent difficulty understanding the notion of asynchronicity. The latter suggests that more work on the paradigm is required before the benefits of email can be brought to illiterate users.

Categories and Subject Descriptors

H.5.2 Information interfaces and presentation: User interfaces

General Terms

Design, Experimentation, Human Factors.

Keywords

ICT for development, video mail, illiterate users.

1. INTRODUCTION

Information and communication technology for development, focuses on computing applications for socio-economic development of underserved communities [12, 13, 14, 15, 16, 19, 20, 20, 22, 23, 24, 26]. One common characteristic of these underserved communities is illiteracy. Even conservative estimates of illiteracy suggest that there are over one billion illiterate people in the world [11]. As such, there is value in computing application intended to aid and be used by people in

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Conference AVI '08, Month 5, 2008, Napoli, Italy.
Copyright 2008 ISBN # 1-978-60558-141-5

exploring how computing can be made accessible to illiterate users. This poses a significant design challenge, as the sheer abundance of text in standard interfaces suggests that significant retooling of the interface or completely new interaction styles would be required to ensure usability by illiterate populations.

Previous work in designing user interfaces for illiterate and semi-literate populations focuses on broad principles, recommending features such as the use of graphical icons [6, 7, 12, 13, 14, 19, 20, 20], minimal use of text [2, 7, 12, 13], voice annotation [12, 13, 14, 20], easy navigability [2, 7, 12, 13, 17, 20, 20] and the use of numbers for people who may be illiterate but not innumerate [12, 13, 19, 20, 20]. These principles have been applied to applications in the areas of job search [12, 13, 15], healthcare [6, 14], map navigation [13] and microfinance [19, 20, 20], but has not yet been significantly applied to computer-mediated communication.

The majority of communication applications targeted towards illiterate users are in the area of agriculture and dedicated to query-based communications between an illiterate person and a literate agricultural expert [22, 24, 26]. These applications, and of course the increasingly ubiquitous mobile phone, currently only provide illiterate users with voice-based communication that is typically synchronous. To the best of our knowledge, there are no applications yet developed for asynchronous computer-mediated communication dedicated to illiterate users. In particular, the most common asynchronous communication tool, email, which has had a profound impact on the lives of the world's literate population, is essentially inaccessible to illiterate people.

In this paper, we explore the question of whether and how the benefits of an asynchronous communication tool like email might be made accessible to populations with little to no literacy. Our ultimate goal is to create a communication experience built on standard email protocols – thus enabling simple inter-operability with other systems used by the literate world – that do not require literacy to enable effective asynchronous communication. As a first step towards realizing this goal, we explore the use of video rather than text as the communication medium. We present the design and evaluation of a prototype video-mail application that uses a combination of graphics, animation and voice assistance to empower illiterate users to be completely self-reliant right from setting up accounts through communicating using it.

In the following sections, we further discuss the motivation and challenges of this problem space, previous research from which we draw design guidance, the user community we worked with, the iterative design and implementation of a fully functional prototype video mail system, and a two-stage user study that evaluated this functional prototype with twenty illiterate users from the urban slums of Bangalore, India. We conclude with a discussion of the insights gained from both our iterative design process and user study, and design recommendations for future instantiations of the concept of video mail for illiterate users.

2. MOTIVATION and CHALLENGES

There are over 1.2 billion email users in 2007, and this number is expected to rise to 1.6 billion by 2011 worldwide [25]. Thus, email is clearly a dominant asynchronous communication tool amongst literate people worldwide. However, users are expected to have several pre-requisite skills, including:

- Literacy or at least semi-literacy.
- The ability to recall login information that is a combination of text, numerals and symbols.
- The ability to set-up their account by understanding the standard settings of the service provider.
- Dealing with email client applications that are normally text-intensive.
- An understanding of and ability to use a navigational system which is heavily text-based.
- An understanding of hierarchical structuring of information (folders, etc).
- Constant decision-making from multiple existing choices to achieve a task.

These skill requirements of current email systems make them essentially inaccessible to the world's one billion illiterate people. Now, one could argue that this population might be better served with technologies such as mobile phones which have a lower usability barrier. However, the synchronous communication capabilities provided by mobile phones are clearly not suitable for all communication scenarios. More asynchronous forms of communication are also desirable, as the numerous "professional letter writers" who ply their services in the towns and villages of the developing world (particularly in India) vividly illustrate. While the age-old physical letter continues to serve its intended purpose, as these populations become increasingly mobile, traveling farther afield from their home villages in seek of work, the need for efficient communication with their now geographically distant relatives and friend becomes more acute. Thus, more efficient alternatives to the postal mail service merit serious investigation. Given email's established track-record as the asynchronous communication tool of choice amongst the technologically-literate peoples of the world, it is arguably reasonable to see if we can somehow morph this technology to also serve the communication needs of the illiterate.

Doing so requires significantly ameliorating the skill requirements of current email systems as enumerated above. The primary challenge is in providing both an interface and communication medium that does not rely on text. The difficulties inherent in text as the medium of communication could potentially be solved by using video and audio exclusively. The arguably more daunting remaining challenge is in designing a user interface that would allow asynchronous "email-like" video and audio communication in a facile manner for people who are not only illiterate but also completely novice computer users. In approaching this challenge, we draw from previous research interfaces used in other application domains.

3. RELATED WORK

There are three areas of related work which are particularly relevant to our research. First is the work on user-interface design for illiterate users. Second is the area of web-based asynchronous communication applications focused on novice but literate users. Third is the work on pictorial passwords for authentication.

3.1 Interfaces for Illiterate Users

Early research in this area placed emphasis on the need for contextual design methods to explore this problem, as illiterate users are very different from the target user imagined by most interface designers [3]. We follow this lead, and have spent literally hundreds of hours in the field working with non-literate people on a variety of projects. Most previous work with non-literate users focuses on the mechanics of the interface. In particular, researchers recognized the value of imagery in place of text, and extensive use of graphics is advocated by most of this work [6, 7, 12, 13, 14, 19, 20]. Some also explored the value of voice instructions and annotations. Much of the interesting work in this area focuses on the subtle interplay between graphics and audio to generate a compelling interface. Some authors note that it might be plausible to include numerals, as illiterate users are often numerate [12, 13, 19, 20, 20], while others have focused on ultra-simple navigation as a design goal [6].

These principles have been applied to create text-free user interfaces [12, 13, 14, 15] for similar user groups for other application domains such as a job-information system for illiterate domestic helpers [12, 13, 15], a health information dissemination system for illiterate patients [14], map-navigation [13] and microfinance [19, 20, 20]. One high-level goal of these systems was to create interfaces that an illiterate person can use, on first contact with a computer, to immediately perform useful tasks with minimal additional human assistance. Since our interface goals are clearly very similar, it is worth identifying some design principles that arise from this prior art. In particular, we draw upon earlier work by members of our research team [13], where design principles for text-free interfaces were elicited through an ethnographic study involving over 300 hours and 200 people from urban slums in Bangalore, India. They are as follows:

Liberal use of graphics and imagery; use of static hand-drawn representations with voice annotations: While the use of graphical imagery rather than text is an obvious feature, the exact nature of the graphics can make a huge difference. A comparative study [14] between different audio-visual media with 200 participants showed that static hand-drawn representations with voice annotations were best understood.

No use of text, but numbers might be acceptable: While less text makes sense for people who cannot read, it was discovered that illiterate people could often easily recognize numerals (0, 1, ... 9), hence it might be reasonable to use numerals in the interface, at least for some subset of the illiterate user population.

Voice feedback and help throughout: Pre-recorded human speech segments as a way for the interface to "converse" with the user was found to be extremely valuable. As such, it is recommended that voice feedback and help be provided throughout the interface.

Consistent "help": Easily accessible and always available auditory help allows an application to be more autonomously used, even for novice users. Additionally, an on-screen character could help users relate the voice to a visual representation.

3.2 Asynchronous Communication Applications for Novice Users

Researchers have looked at communication systems between novice users and domain experts through a query-based approach using voice [22, 24, 26] and video [26]. Both these examples are in the agriculture domain. While these do use voice and video like we intend to, they do not address the peer-to-peer personal communication domain that is the focus of our work.

In the personal communication domain, research on Chinese migrant workers and the interaction with their children has been studied [17, 18]. While these studies explore research questions similar to those we are interested in, they were essentially initial feasibility explorations that recommend voice-based communication augmented with video. No hi-fidelity working prototype was developed and as such no significant usability studies were conducted using a working system [17, 18].

3.3 Pictorial Passwords for Authentication

Researchers have explored the use of pictorial passwords for user authentication that provides better recall and usability [3, 5, 8]. However, most of this work aims at providing insights and principles for literate computer users. In a series of experiments, Katre [8] studied the use of pictorial passwords for illiterate populations and found that while users could easily recall previously seen pictures, they were not necessarily able to recall a set of pictures in a particular sequence as would be required for a typical password authentication scheme. Katre proposes various schemes that might mitigate the sequence recall problem, but these have yet to be tested in a real application.

4. TARGET USER COMMUNITY

We conducted our research with people in three urban slum communities in Bangalore, India. To gain access into these communities, we worked with a non-governmental organization (NGO) called Stree Jagruti Samiti (SJS), which has had an established presence in these three communities for 15 years. SJS works primarily with the women and children in the slums. All of the people we worked with had three common background traits: (1) functional illiteracy or semi-literacy (typically somewhat numerate); (2) low levels of formal education (highest education attained being schooling up to the fourth grade); and (3) no experience whatsoever in using a computer. These traits make them an ideal user population with which to explore our ideas with regards to creating a voice and video based email surrogate suited for illiterate populations.

These communities we work with have their own unique characteristics, and these should be kept in mind when attempting to generalize the results we present later. For example, populations differ in terms of their attitudes toward illiteracy. The people we worked with were very frank with respect to illiteracy, attaching no shame to the inability to read; this is unlike illiterate individuals in developed countries who often hide this inability. Also, our users held strong positive associations of the English language (which they did not speak for the most part) with wealth and prestige – both a holdover from colonial British rule, as well as a modern-day fact due to the economic opportunities available to English speakers. These characteristics might have had an impact on our results, in a manner that could well have been different had our users been people from other locations and cultures. The state of illiteracy, poor education, and ignorance of computer technology are factors common across all our users and

are arguably those that impact our designs the most. Other demographic characteristics of this population likely had less influence on our particular problem space; however, we list them here for the sake of full disclosure: About half the people were female household workers who clean private homes, wash dishes, and so forth. The other half were males who are typically daily wage laborers like plumbers, carpenters, construction workers, mechanics, or fruit and vegetable vendors. Their primary language of communication is Kannada, but many speak additional languages such as Hindi, Tamil, or Telegu. The average household income was INR 800 - INR 3000 (approximately USD 18 – USD 67) per month, in line with general market statistics on wages in India [1]. A few had television sets, music players and gas burners, but these were not owned by all households. Some had seen computers in the houses of their employers, but due to class- and caste-based discrimination, were generally prohibited from touching the computer (even for the purposes of cleaning!).

5. PROTOTYPE VIDEO MAIL SYSTEM

While some aspects of our explorations into the feasibility of video mail as a synchronous communication technology for illiterate users could be accomplished using primarily low-fidelity methods, we chose to iteratively develop a higher-fidelity working prototype system instead. Since a working system will allow us to conduct true usability tests in the field, we believe it will allow us to obtain a more ecologically valid and deeper understanding into the issues surrounding the potential use of this technology. The prototype was designed and refined over three stages:

5.1 Initial Designs and Mental Models

In the first stage of our design process, we used principles suggested by prior research in text-free user interfaces [12, 13, 14, 15] to generate the initial user interface. In addition, we strove to promote a strong mental association with existing systems of communication as a method to ease novice users into this technology. Accordingly, we chose the postal system as a metaphor that would be easily understood by our target community. The user interface’s “look” had the essence of a familiar Indian postcard (Figure 1), and the functions and controls simulated the process of creating and posting it. A visually salient audio assistant was designed as a post-person who explained creating, sending and receiving video mail in terms of the postal system’s process that was familiar to this user population.

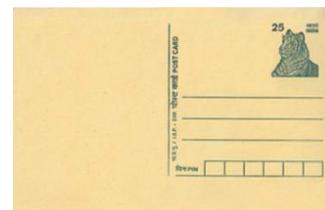


Figure 1. Indian postcard used as a familiar metaphor.

Our initial prototype design is illustrated in Figure 2. On the left of the screen is an image of a woman dressed as a typical Indian post-person that represents the audio assistant who provides context specific voice help when the user hovers the mouse cursor over it. This audio assistant remained on screen at all times as it was our intent to provide a consistent place where users could turn to for help at any time. Figure 2a shows the login screen, which consists of numbers for the user ID, and pictures for the password. Figure 2b illustrates the inbox of messages, consisting primarily of a photograph of the sender with color coding to indicate new

and old messages, and a time of day icon (moon and sun), and the area used for creating messages in four possible formats: freeform ink which we thought might be viable as a means of spontaneous expression, audio only messages, audio+video, and text – included for backward compatibility for the literate user. Clicking on the relevant message format icon immediately started the recording. We chose to support these three non-text formats rather than video alone as at this initial stage we did not want to presume that video was necessarily the best format for asynchronous communication by illiterate users.

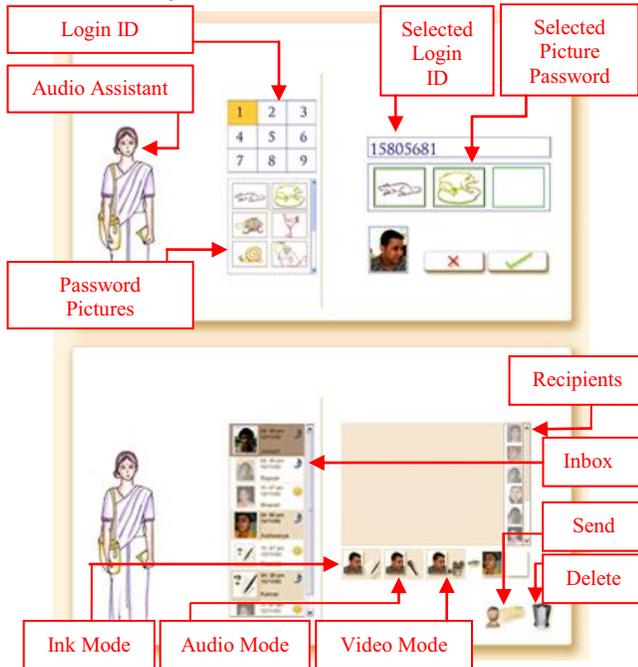


Figure 2. Initial prototype. (top) Login. (bottom) Inbox & mail creation. Red annotations are for illustration only.

Using this initial prototype, we sought feedback via informal user testing and interviews with three people representing illiterate, semi-literate and seasoned computer user groups. Note that we deliberately included a semi-literate and seasoned computer user because we wanted to get a sense of how the perceptions and usability of our designs might differ depending on the user population. These sessions led to several key observations:

- Login has to be simplified. User name as a combination of numerals and the password as a combination of images required recall on two separate parameters, and this proved difficult for the users.
- Vertical scrollbars were not understood by the illiterate and semi-literate users, as the arrows simply went unnoticed.
- Color coding for old and new messages was not understood.
- Actions should occur only when the user conducts a specific function and not automatically. For example, the prototype started recording immediately upon clicking the relevant message format button, rather than using additional “start/stop recording” buttons, resulting in some confusion.
- Animation is needed when transitioning from one screen to the next. Abrupt switching as is typically done in regular applications resulted in confusion due to a loss of continuity.

5.2 Free-Form Study of Revised Prototype

In response to the key observations in the previous stage, we made several modifications to the prototype in turn:

- Changed the login mechanism to use photos of users as the login ID instead of the numeric ID. The pictorial password mechanism used in the first stage was retained. (Figure 3a)
- Replaced the vertical scrollbars with much larger up/down arrow icons at the top/bottom of lists as required (Figure 3b).
- Removed the color coding of old and new messages. Unfortunately, we were not able to devise an alternate mechanism for distinguishing between old and new messages that we felt would work for this population of users.
- Added explicit start, stop, and play buttons to control the playback and creation of mail messages (Figure 3b).
- Added smooth animated transitions between screens.
- Added an account creation phase for first time users.

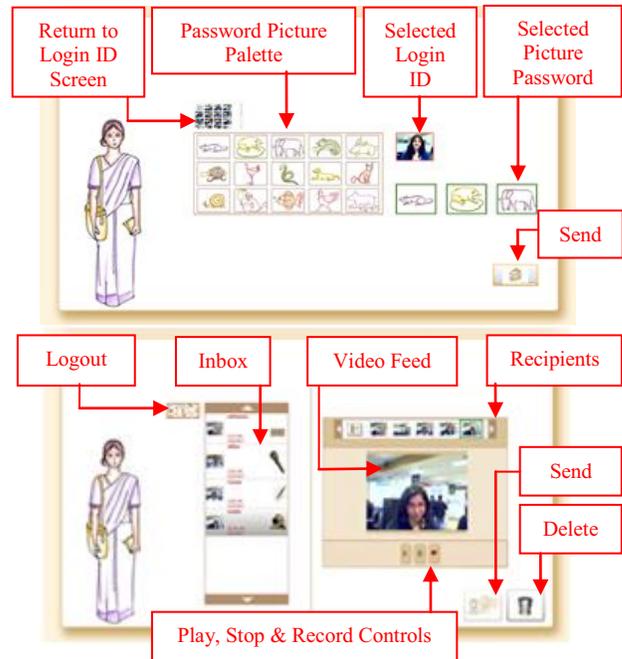


Figure 3. Revised prototype. (top) Login. (bottom) Inbox & mail creation.

In order to further evaluate this revised prototype, we installed the software on an unattended computer accessed by a group of thirty people whose background was comparable to our ultimate target user community. These users were employed as cleaning and facility maintenance staff at our corporate office in Bangalore. The software ran on this computer for a six-week test period. We chose to do this phase of testing in this setting rather than with the ultimate target users in the slum communities because we wanted to refine our prototype as much as possible within our controlled facilities before taking it out into the field.

The task required the participants to set-up their own mail account on the system, login and send a mail. We observed whether the application could be used without human assistance, if in fact it would be used at all, and the hurdles faced while using it.

Attention was paid to understanding which features were most used and why. This stage resulted in the following observations:

- a) Users struggled with setting-up their account as we had not completely eliminated the requirement for text input in this phase (*i.e.*, names had to be typed in as text). This was not unexpected, and we were concurrently working on a text-free account setup design to be implemented in the final prototype.
- b) The application sometimes offered more than one way for users to accomplish their task. In particular, when creating a new mail, users had to choose between video, audio, ink, and text. This choice was found to be confusing by some users.
- c) Video mail was the most used. Users occasionally used ink mail (drawn images as mail), rarely used voice mail, and predictably never used text mail.
- d) Interface elements that were not to be used and were grayed-out perplexed them.
- e) Users found innovative uses for video mail such as creating a song chain or reading out the new headlines (if they were semi-literate) for their illiterate co-workers.
- f) Often they would help each other with setting-up and using the application. Superfluous visual elements that were used to

enhance the post-card metaphor proved to cause confusion.

- g) Some users faced difficulty in remembering the order of the graphical passwords.

5.3 Final Prototype

Based on the feedback from the previous stage the application was further revised in several ways.

First, we redesigned the account setup interaction to avoid any text input, relying instead on taking a photo of the user and using it as the login ID with a unique computer generated identifier assigned automatically to that photo. Further, given the difficulties faced by users in remembering the order of the graphical passwords – a finding similar to that found by Katre [8], we replaced the pictorial passwords with numeric ones instead. While we do not know *a priori* whether or not this would work better than pictorial passwords, we felt it was worth exploring as pictorial passwords were clearly not feasible. Figure 4 illustrates.

Second, we significantly simplified the rest of the interface, in particular retaining only the video mail feature and removing the ink, audio, and text mail formats as choosing between multiple formats was found to be confusing by users in the previous stage. Audio assistance was enhanced, and retained the postal worker metaphor. Figure 5 illustrates.

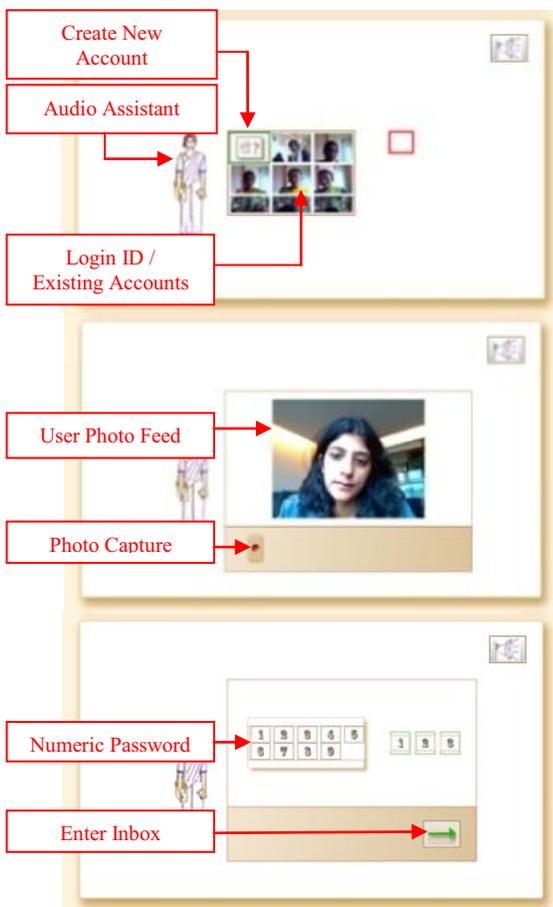


Figure 4. Final prototype: account creation. (top) Opening page showing existing accounts and a “create new account” icon. (middle) Capturing a photo for a new user ID. (bottom) Selecting numeric password.

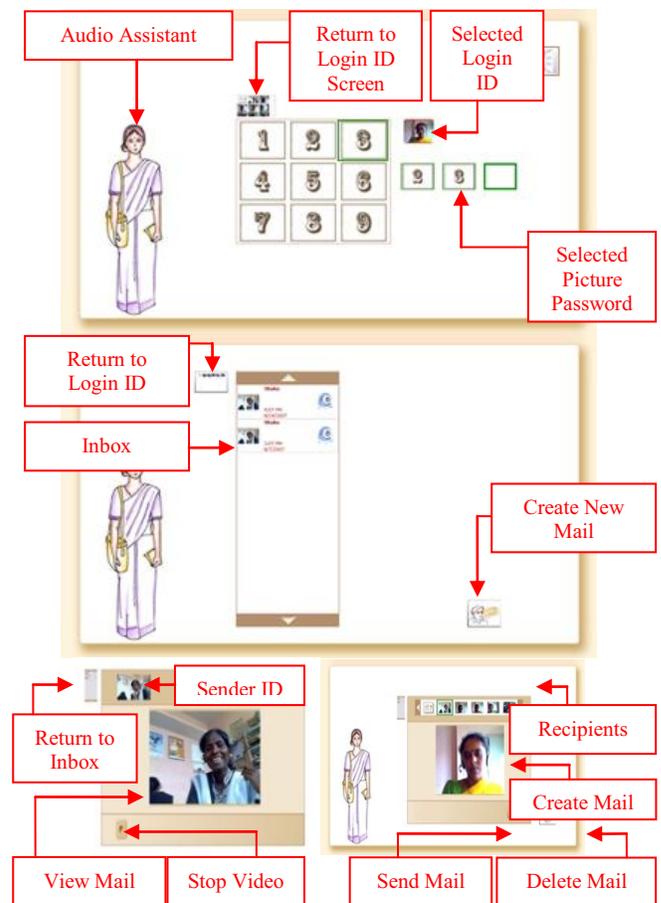


Figure 5. Final prototype. (top) Login screen. (middle) Inbox. (bottom left) View video mail. (bottom right) Create video mail.

6. USER STUDY

6.1 Goals and Design

We evaluated the final prototype with 20 participants (10 female, 10 male) from the target user community described in section 4. Participants ranged in age from 25 to 45 years.

The study was intended to determine if our target users would be able to understand the overall concept of video mail and perform the actions required to setup an account, login, and send/receive mail. The study was conducted in the homes of the participants, and we used a tablet PC and a pen as we felt it more closely resembled the paper-and-pen letter-writing metaphors that participants were used to. A representative of the NGO we work with acted as a primary contact person with whom the participants would communicate via the video mail system. This primary contact person was well known to all participants, and pre-recorded a welcome email message that was shown in the inbox of all new accounts. We conducted the study in two stages:

Stage 1:

Using a pre-authored script, participants were walked through the system by the experimenter (the first author). The overall concept of asynchronous email communication was explained to them using analogies to the postal mail system that participants were already familiar with. Participants were shown how to use the tablet PC and pen and told that they can seek assistance from the audio assistant at any time.

They were then asked to perform a set of tasks as follows:

- Set up their own video-mail account. This involved taking a user photo of themselves and selecting a 3 digit numeral-based password as shown in Figure 4.
- Login to their new account (Figure 5a).
- Retrieve a welcome message in their inbox created by the primary contact person
- Compose and send a response to the welcome message
- Logout

Stage 2 (10 days later):

The primary contact person looked at and replied to all the mails sent to her from the previous stage before the same set of participants began this stage. We deliberately conducted this second stage 10 days after the first stage was completed as we wanted to see if after some time away from the system, participants could recall their passwords and how to use the system without the help of the experimenter. Unlike in the first stage, in this second stage the experimenter did not do an initial walkthrough of the system, but was available to assist if participants got completely stuck and were unable to proceed otherwise.

Participants were asked to perform a set of tasks as follows:

- Log into their video mail account
- Retrieve new mail from the main contact
- Compose and send a response
- Log out

6.2 Results and Discussion

Overall, each participant took about 5 to 20 minutes to complete the task in stage 1, and 5 to 10 minutes in stage 2. One female user did not show up for the second stage of the study, but the remaining users all eventually completed the task in both stages. Overall, we found that the male users completed the task faster and were more at ease with the technology than the female users.

We used four techniques for data collection: detailed notes taken by the experimenter in-situ while the participants were performing the tasks, continuous screen captures using a software tool (Community Clips) to record all on-screen activity, a video camera that recorded participants actions from an “over the shoulder” vantage point, and a software logger that recorded all mouse and keyboard inputs within the application.

The following are key observations gleaned from careful analysis of the detailed notes taken by the experimenter, and manual coding of the many hours of screen captures and video camera data. We ended up not using the data from the software logger that recorded mouse and keyboard inputs, as we found that data at such low level of detail was not necessary to determine the essential usability issues.

A key aspect of our overall design was the availability of the audio assistant at all times, and we were very interested in how this style of assistance fared with our target users. We found that:

- Users were unable to follow multiple linear audio instructions, and most often just followed the first or last in the series. For example if the audio assistant says “Please click on the Play button to play the video mail, the Stop button to stop the video mail and the Record button to record a video mail”, users will disregard the fact that they have a choice and simply follow through with either Play or Record.
- Unless prompted by the experimenter, users did not use the audio assistant in stage one of the study. However, during the second stage of the study it was observed that the users were significantly more confident and needed less prompting from the experimenter to use the audio assistant. As Figure 6 illustrates, the audio assistant was used extensively during both stages of the study, despite the relatively simple nature of the tasks performed by the users. This indicates that having continuously available help is crucial for these users.

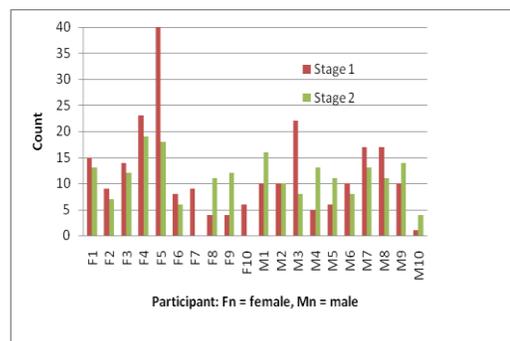


Figure 6. Number of invocations of audio assistant by each user during both stages of the study.

- Audio instructions that relied on color may be misleading even when used in combination with another parameter. For example, if the audio assistant instructs the user to click on a green arrow, the user is likely to try and click on anything

green, including green parts in the background of a video-still. This suggests that color coding should either be avoided completely or used only when the rest of the screen does not contain the same color.

- d) Users tended to hold their mouse over the audio assistant through a whole message, continually hovering the cursor over the icon without clicking.
- e) Congratulatory audio messages seem to produce excitement and encouragement. For example, after going through the login process of selecting the login photo and correct password, an audio congratulatory message informing the user in a congratulatory tone that they had successfully entered their inbox and that they could now retrieve mail or create a new one produced a lot of positive excitement.

With regards to the concept of video mail and receiving/creating new mail, the data shows that receiving personalized video mail was clearly seen as an exciting event. It was very interesting to observe that many participants did not understand that the welcome video mail was pre-recorded. They attempted to have conversations with the video mail, even when it was re-played several times over! A possible hypothesis is that their mental model of the synchronous telecommunication system overrides our intended asynchronous postal-system mental model. The moment users are faced with a video of a person talking, they immediately respond as they assume that this is similar to a telephone conversation – that this is a video phone. This finding was a bit disappointing, as we had put in significant effort in designing the interface to project an asynchronous model. Clearly further work is required to get this aspect of the system right.

With regards to the authentication scheme, we found that the combination of photograph as the login ID and numeric passwords worked reasonably well. However, we observed that users sometimes had difficult deciphering all numerals and confused the numbers 2, 3 and 5. From interviewing users on this issue, we discerned that many tended to remember the numbers by their placement within the on-screen number pad rather than by actual recognition of the numerals per se. This indicates that one should keep the on-screen visuals of the numerals consistent across versions of software, and potentially use a similar login screen for multiple applications for this population in order to reduce any recognition confusion. Regardless of the mechanisms by which they recalled the numeric passwords, in stage one, all the male users were able to create their passwords in the account setup stage and immediately thereafter reenter their password on the login screen with no errors. Six of the ten female users made mistakes when reentering their password on the login screen, even though they had just created the password moments ago in the account setup stage. We attribute this to differences in numeracy between the males and females in our user population. Interestingly, however, in the second stage of the study, male and female users were equally adept at recalling the passwords they had created ten days earlier, with just three men and three women making mistakes on the first attempt at password entry.

In the rare case that two users share the same password and one of them logs in with the other user's photo, it was observed that they were unable to decipher that they were in the wrong inbox. This leads us to believe that a personalized welcome message when a user enters their inbox is required.

The data also revealed several other key issues with regards to various interface elements. While these were observed within our video mail application, many are general issues that would apply to any application for illiterate populations and hence potentially have implications beyond the present work:

- a) Linear progressions are not conceptually understood. Users did not understand that they were being taken from one screen to the next. Thus icons that show the previous screen as a means to get to the previous screen may be moot.
- b) Users had some difficulty identifying 2D thumbnail photos of themselves. This could be due to poor quality of images taken by webcam and the small thumbnail size and poor eyesight that some users might have. A potential design solution might be to increase the image size on mouse-over.
- c) Users did not seem to realize that they need to click on the 'Stop' button to stop the action of recording a message. This was in spite of clear audio instructions to do so. We suspect this might be due to users thinking of the recording as a synchronous open communication channel where there is no explicit end. We continue to seek better ways of reinforcing the notion of asynchronicity.
- d) Similar to the 'stop' issue above, users did not seem to see the need to "exit" the application on completing their tasks. The notion of an application that had to be started and stopped is clearly foreign to this population. Perhaps a kiosk-style "always on" appliance might be more appropriate.
- e) Several users clicked on the lower (older) mail in the inbox when attempting to access their new mail. As noted in the earlier phases of our work, standard grouping techniques such as color coding of new and old messages were ineffective. We intend to explore other techniques such as putting old messages in a completely separate space on screen.

7. CONCLUSIONS

Our work suggests that providing a personal asynchronous communication system for illiterate users could be viable. Our user study showed that users were able to grasp the basics of the application and complete the given tasks. Most importantly, they were able to do so even after a ten-day break from the initial demonstration by the experimenter. While users clearly required help throughout, they were able to get this help mostly from the onscreen audio assistant, which indicates that such systems will likely not require a human expert attendant beyond the initial demonstration. To the best of our knowledge, this is the first articulation of the viability of video mail for illiterate users. Further, the design insights gained in our work also contributes to the growing literature on designing interfaces for this population.

While the use of text-free graphical interfaces for applications focused on illiterate users is not new, our work expands upon the literature by applying these principles into the previously largely unexplored domain of asynchronous personal peer-to-peer communications. Our experience in designing and evaluating this prototype video-mail system clearly showed that there remains much to be learnt in the area of designing interfaces for this population of users, as simple application of previous design principles did not immediately result in a usable system. In particular, it is important to note that although the tasks in both stages of our study using the fully functional final prototype are almost trivially simple from the perspective of seasoned literate

computer users, they are anything but trivial for illiterate users who have never previously used a computer. Despite our best efforts in earlier phases of the work to reduce interface complexity, our study revealed various highly nuanced issues that remain to be solved. Many of these issues would not have manifested themselves in a more literate population, indicating that significant challenges need to be surmounted in order to make even the simplest applications accessible to illiterate users.

Next steps in this work include a limited deployment of a working system to determine if it will actually be used by, and be useful to, the community over an extended period in the field.

8. VIDEO

A video demonstrating the system and aspects of the user study can be found at www.youtube.com/videomailapp

9. ACKNOWLEDGMENTS

We thank all who participated in our studies.

10. REFERENCES

1. Bery, S. (2006). National Council of Applied Economic Research. **Error! Hyperlink reference not valid.**
2. Chand, A. (2002). Designing for the Indian rural population: Interaction design challenges. *Development by Design Conference*.
3. Cooper, A. and Reimann, R. (2003). *About Face 2.0, The Essentials of Interaction Design*. Wiley Publishing Inc. USA.
4. Davis, D., Monrose, F. and Reiter., M. (2004). On user choice in graphical password schemes. *13th USENIX Security Symposium*. p. 151–164.
5. De Angeli, A., Coventry, L., Johnson, G. and Renaud, K. (2005). Is a picture really worth a thousand words? Exploring the feasibility of graphical authentication systems. *International Journal of Human-Computer Studies*, 63(1-2), p. 128-152.
6. Grisedale, S., Graves, M. and Grunsteidl, A. (1997). Designing a graphical user interface for healthcare workers in rural India. *ACM CHI Conference*. p. 471-478.
7. Huenerfauth, M. (2002). Developing design recommendations for computer interfaces accessible to non-literate users. *Master's thesis, University College Dublin*.
8. Jermyn, I., Mayer, A., Monrose, F., Reiter, M. and Rubin, A. (1999). The design and analysis of graphical passwords. *USENIX Security Symposium*. p. 1-14.
9. Katre, D. (2004). Using mnemonic techniques as part of pictorial interface for self identification of illiterate villagers. *International Conference on Human Computer Interaction, Bangalore, India*.
10. Kumar, R. (2004). eChoupals: A study on the financial sustainability of village internet centers in rural Madhya Pradesh. *Information Technology and International Development*. 2:1. p. 45-73.
11. Lourie, S. (1990). *World literacy: where we stand today - One billion non-literates*. Editorial, UNESCO Courier. July 1990.
12. Medhi, I. Pitti B. and Toyama K. (2005). Text-free UI for employment search. Proc. of *Asian Applied Computing Conference, Nepal*.
13. Medhi, I., Sagar, A. and Toyama K. (2006). Text-free user interfaces for illiterate and semi-literate users. *International Conference on Information and Communication Technologies and Development, Berkeley, USA*. p. 72-82.
14. Medhi, I., Prasad, A. and Toyama K. (2007). Optimal audio-visual representations for illiterate users. *International World Wide Web Conference*. p. 873-882.
15. Medhi, I. and Kuriyan R. (2007). Text-free UI: Prospects for social inclusion. Proc. of *International Conference on Social Implications of Computers in Developing countries, Brazil*.
16. Mitra, S. (2005). Self organizing systems for mass computer literacy: Findings from the hole in the wall experiments. *International Journal of Development Issues*, 4(1). p. 71-81.
17. Moraveji, N., Ho, R., Huynh, D., and Zhang, L. (2005). An exploration in interface design for the Chinese migrant worker population. *Designing for the User Experience, San Francisco, CA*.
18. Moraveji, N., Ho, R., Huynh, D., and Zhang, L. (2005). Communication gap: Designing an interface for Chinese migrant workers. *Usability Professionals' Association, User Experience Magazine*, 5(2).
19. Parikh, T., Ghosh, K. and Chavan, A. (2003). Design considerations for a financial management system for rural, semi-literate users. *ACM CHI Conference*. p. 824-825.
20. Parikh, T., Ghosh, K. and Chavan, A. (2003). Design Studies for a financial management system for micro-credit groups in rural India. *ACM Conference on Universal Usability*. p. 15-22.
21. Parikh, T. (2002). HISAAB: An experiment in numerical interfaces, Media Lab Asia Panel Discussion, *Baramati Initiative on ICT and Development*.
22. Plache M. and Prabaker M. (2005). Tamil market: a spoken dialog system for rural Indi. *Extended Abstracts of the ACM CHI Conference*. p. 1619-1624.
23. Ratnam, B. V., Reddy P.K., and Reddy, G. S. (2005). eSagu: An IT based personalized agricultural extension system prototype – analysis of 51 Farmers' case studies. *International Journal of Education and Development using Information and Communication Technology*, 2(1). p. 79-94.
24. Ramamritham, K., Bahuman, A., Duttagupta, S., Bahuman, C., and Balasundram, S. (2006). Innovative ICT tools for information provision in agricultural extension. *International Conference on Information and Communication Technologies and Development, Berkeley, USA*.
25. Radicati Group Q3 2007 Market Update: www.radicati.com/uploaded_files/news/Q32007_PR.pdf
26. Ramamritham K., Bahuman A. and Duttagupta S. (2006). aAqua: a database-backed multilingual, multimedia community forum. *ACM SIGMOD Conference*. p. 784-786.
27. Wiedenbeck, S., Waters, J., Birget, J.-C., Brodskiy, A., and Memon, N. (2005). Authentication using graphical passwords: Basic results. *Human-Computer Interaction International Conference*. p. 1-12.