

Featherweight Multimedia for Information Dissemination

Gerry Chu, Sambit Satpathy, Kentaro Toyama, Rikin Gandhi, Ravin Balakrishnan, S. Raghu Menon

Abstract— Featherweight multimedia devices combine audio with non-electronic visual displays (e.g., paper). Because of their low cost, customizability, durability, storage capacity, and energy efficiency, they are well-suited for education and information dissemination among illiterate and semi-literate people. We present a taxonomy of featherweight multimedia devices and also derive design recommendations from our experiences deploying featherweight multimedia in the agriculture and health domains in India. We found that with some initial guidance, illiterate users can quickly learn to use and enjoy the device, especially if they are taught by peers.

Index Terms—Audio, illiteracy, information dissemination

I. INTRODUCTION

A major thrust towards bringing information and communication technologies to underserved populations is the development of low-cost computational devices such as the XO laptop from OLPC [1], Classmate PC [2], and even a potential \$12 laptop [3]. While these initiatives and other factors such as increased penetration of mobile phones will undoubtedly have tremendous impact in providing the global poor with better communication tools and access to information, the cost of these devices remain prohibitive to the 2 billion+ people worldwide who live on less than \$2 per day.

Beyond monetary cost, illiteracy remains a major hurdle. Conservative estimates of illiteracy suggest that there are over one billion illiterate people worldwide [4] for whom even a free, connected computer is useless given that some level of literacy and basic education is typically required to operate them; recent attempts at text-free interfaces [5] hold the promise of partial accessibility to computing by the illiterate, but still don't address the complexity of the device.

While it is certainly worthwhile trying to grapple with the cost and literacy constraints of “full-fledged” technology for many applications, there are other applications that may only require a little bit of technology in order to add significant value. One such broad application area is the information dissemination that government and non-profit organizations do to improve the education and health of people. In these populations, traditional information dissemination methods

include printed material, radio and television public service announcements (PSAs), and verbal instruction. These methods are certainly viable, but suffer from several shortcomings, such as requiring literacy to comprehend textual printed material, the relatively high cost of producing, distributing and consuming radio and television PSAs, and the arguably higher cost of verbal instruction by humans who might not even deliver a consistent message over time. Of these, only verbal instruction provides the possibility of interactive exchange with the consumer.

In this paper, we explore the notion of featherweight multimedia” – various combinations of electronic audio devices with non-electronic visual displays – for interactive multimedia information dissemination (Fig. 1). Featherweight multimedia devices require minimal power, are more rugged, and are dramatically cheaper than low-cost computers or feature-rich mobile phones. Yet, they integrate many of the best elements of traditional techniques with the enhancement of interactivity. Our contribution is not in the technology per se, but rather in 1) mapping out the design space of featherweight multimedia, 2) exploring factors affecting their use for information dissemination to illiterate or semi-literate populations using our experiences deploying them, and 3) reporting the reactions of the target population to initial usage of several featherweight multimedia variants. Our preliminary experiments suggest that non-literate users with little formal education can quickly understand how to use featherweight multimedia, but that social support is required to maximize effectiveness.



Fig. 1. Village residents using a talking book

Manuscript received September 22, 2008.

G. Chu, S. Satpathy, and R. Balakrishnan were with Microsoft Research India (email: gerrychu | ravin @dgp.toronto.edu, sambit407@gmail.com).

K. Toyama, R. Gandhi, and S. Raghu Menon are with Microsoft Research India, “Scientia”, 196/36 2nd Main, Sadashivnagar, Bangalore, 560 080, India (phone: +91 (80) 6658-6000; email: kentoy|rikingt-raghum@microsoft.com).

R. Gandhi is additionally with Digital Green

R. Balakrishnan is with the University of Toronto

II. RELATED WORK

Recently, there have been several initiatives adapting specific featherweight multimedia devices to the development

field. These evolved from devices that were originally created for the developed world. In addition, featherweight devices can trace their design histories to other multimedia initiatives that use audio and video in development.

A. Existing featherweight multimedia devices in development

Books of Hope is an organization that creates talking books in South Africa [6]. Each 16-page book is held within an electronic device that has 16 buttons. The user turns to a page, each of which has an icon. The user presses the button labeled with the same icon, triggering the appropriate caption to play aloud. Each book holds 5 minutes of audio, and with two AAA batteries, 200-500 plays can be achieved [7]. In an evaluation [8], a control group of grade 7 students attended a workshop about mental health, while an experimental group received talking books. Pre- and post-tests determined that both groups learned the same amount but a large percentage of students who received the talking books showed them to other students (80%), to their families (80%) and to church members (43%), pointing to possible viral dissemination benefits of cheap portable featherweight multimedia.

Talking books such as LeapPad [9], WhizKid [10], and PowerTouch [11] are educational toys for children. They are operated by overlaying a piece of paper or workbook onto a touch or pen-sensitive screen. The device can identify the current page, so that when elements on the page are pressed, the device plays a recorded caption. There was an announced initiative in 2004 to deploy the LeapPad in Afghanistan [12], but nothing has been published about the deployment since. However, the PicTalk device, which uses LeapPad technology, is being used to teach Indian schoolchildren [13]. Students using PicTalk to learn English performed 40% better on post-tests than a control group learning using a traditional curriculum.

The Literacy Bridge project has created an interactive digital audio player/recorder costing roughly US\$5-\$12 and are deploying it in Ghana [14]. Although they call it a "Talking Book" the focus is on the audio and not on the paper book component.

B. Devices similar to featherweight multimedia

Audio guides (e.g., [15]) are widely used in museums and galleries. They play an audio clip when a numeric code is typed in. An interesting element of these sorts of audio interfaces is the spatial decoupling of the visual printed media from the audio device.

Similar to the aforementioned talking books, communication aides are used by people with developmental disabilities who cannot communicate verbally [16].

More advanced and expensive audio/paper devices such as the Audio Notebook [17] and the Pulse SmartPen [18] record audio as the user writes in a notebook. Tapping on a word plays the audio that was recorded when the word was written.

The Daisy Consortium is the publisher of a file format for talking books meant for the blind or those with learning disabilities [19]. The Dolphin company creates players and authoring tools for talking books in the Daisy format [20].

Public advertising on billboards, as well as research on more interactive public ambient displays have demonstrated designs intended to engage casual bystanders in information exchange [21][22]. A featherweight multimedia instantiation of this concept might involve billboards with phone numbers that passersby can call for additional prerecorded information, or paper posters with an audio player, a design that we explore later in this paper.

C. Related work in development

There have been two notable audio interfaces for developing countries. Tamil Market uses a spoken dialog system that allows illiterate people to query the system by voice to hear crop prices spoken aloud [23]. Similarly, Pakistani community health workers can call HealthLine, ask it verbally for information, and hear medical information over their mobile phone [24]. Both require backend computers to provide speech recognition, hence imposing a fairly significant computational and cost requirement.

A system using visual codes and mobile phone cameras, such as that explored by Parikh et al. [25] could also be used to trigger audio playback.

Work on text-free user interfaces [5] has focused on replacing text in computing interfaces with audio prompts that are played when the user mouses over a graphical interface element. The focus of this work has been on making conventional computers accessible to the illiterate population, and hence the usual cost and power requirements of standard computation remains. Featherweight multimedia can be characterized as being text-free user interfaces without an electronic visual display.

The concept of integrating paper and audio is supported by work by Medhi et al. [26]. They tested five different representations (video, photo, animation, static drawing, and text) of medical conditions with and without audio captions on Indian slum residents. They found that users best understood drawings that are combined with audio.

Video might be considered as the "heavyweight" version of featherweight multimedia. Like featherweight multimedia, it has both audio and visual electronic displays. Video has been used to teach Indian schoolchildren [27] and has also been shown to be effective in teaching illiterate users to use a job-search computer application [28]. The Digital Green project uses video to teach sustainable agricultural techniques to Indian farmers [29]. It has been shown to be six times more effective than conventional agricultural extension workers. Despite these successful examples of video for information dissemination in developing countries, video requires an expensive television and DVD player that depend on electrical power that can be unreliable in many development settings.

III. DESIGN SPACE

At first glance, our definition of featherweight multimedia – the combination of an electronic audio device with some non-electronic visual display – connotes a relatively simple range of possible instantiations. On closer investigation, however, we find a rich design space worthy of careful delineation.

TABLE 1. COMPARISON OF ELECTRONIC AUDIO PLAYERS

	LOW-END MOBILE PHONE	SMARTPHONE	CD PLAYER	PORTABLE AUDIO/MP3 PLAYER	TALKING BOOK (LEAPPAD)	CUSTOM EMBEDDED ELECTRONICS	GREETING-CARD AUDIO CHIPS
DEVICE COST	\$10-50	\$200-500	\$10-\$50	\$10-\$200	\$15-\$50	\$0.05-\$100+	\$0.05
CONTENT CUSTOMIZABILITY	High	High	Medium	High	High	Medium	Medium
CAPACITY	High	High	Medium	High	Medium	Medium	Low
AUDIO QUALITY	Medium	Medium	High	High	Medium	Medium	Low
DURABILITY	High	Medium	Medium	High	High	Medium	Medium
POWER EFFICIENCY	Medium	Low	Low	High	Medium	High	High
USABILITY	Medium	Medium	Medium	Medium	High	Medium	Medium
UBIQUITY	High	Low	Medium	Medium	Low	Low	Low
SENSORS	Low	High	Low	Low	Low	Low	Low

A. Electronic audio player

The audio component of featherweight multimedia can be sourced from various readily available technologies including CD players, portable MP3 players, mobile phones, museum audio guides, and embedded audio devices, including audio chips (as can be found in audio greeting cards [30]), embedded MP3 players, and other custom audio electronics. These vary along several interesting dimensions, summarized in Table 1:

Device cost. This can vary considerably, ranging from just a few dollars at wholesale prices to several hundred dollars for state-of-the-art equipment. An embedded audio chip of the sort found in greeting cards cost just several cents, and one could embed many of these chips into a single featherweight multimedia device. At the other end of the spectrum, museum audio guides can cost thousands of dollars.

Content customizability. CD/DVDs cost just a few cents to manufacture with content, while content on MP3 players can easily be uploaded from a computer. However, the distribution costs of getting those CD/DVDs to the user or providing a geographically distributed user base with access to a content-serving computer to update their MP3 players needs to be accounted for. Mobile phones offer the highly flexible solution of wireless access to content that may not even be stored on the phone itself, although costs for such access can vary considerably across regions. Many embedded devices only allow for a one-time write of the content; where they do allow updating of content, the costs are roughly similar to that of updating an MP3 player although updating an embedded system typically requires more specialized and complex tools than the mainstream ones available for MP3 players.

Capacity. Capacity ranges from 5 seconds for some embedded electronics to days with MP3 players.

Audio quality. This ranges from high-quality audio in CD and MP3 players to somewhat lower quality audio on mobile phones to tinny audio in greeting card style embedded devices. Volume also varies from personal levels in embedded devices to sharable levels in MP3 players with external speakers.

Durability. With the possible exception of CD/DVD media that is prone to scratching, the remaining options tend to rely on solid-state memory and electronics that are roughly equivalently robust. However, the housing for the electronics could make a difference, with greeting-card style embedded devices being relatively fragile compared to an MP3 player.

Power. Most of the devices surveyed have low power consumption requirements. However, they vary considerably

in the types of batteries used and whether or not they can be recharged and/or replaced easily. Some devices use readily available replaceable batteries, some are more specialized and can only be replaced by a technician, while others can only be recharged by plugging into mains power. Depending on the usage locale, the type of battery might be a crucial deciding factor in choice of a featherweight multimedia device.

Usability. Given the focus on information dissemination in low-education target audience, usability is arguably even more crucial than in more mainstream technology. For example, many cheap audio players have surprisingly complex interfaces that require the user to decipher instructions on a tiny LCD screen, thus negating the positive value of their inexpensive cost. We believe that for most usage scenarios, ultra-simple interfaces with just a few buttons (start, stop, rewind) or with one button per audio caption are generally best. A numeric keypad can added if random access to content is provided through numeric codes.

Ubiquity. Mobile phones have staggering penetration in many parts of the developing world. In contrast, more specialized embedded systems might be cheaper but might ultimately not be viable given their specialized nature.

Sensors. Smart phones and custom embedded devices might have cameras, RFID readers, and other sensors that can enable automatic access to the appropriate audio clip depending on the visual content, whereas simpler sensor-free devices would rely on the user to select audio clips for a given visual display.

B. Non-electronic visual display

This can be considered along several key dimensions:

Physical form factor and portability. The display can range from small and highly portable to larger fixed installations. Example form factors include sheets of paper, greeting cards, brochures, books, posters, and even billboards.

Cost. This can range from essentially free (e.g., hand-scribbled content on a sheet of paper), to a few cents (printed paper), to a few dollars (printed books), to thousands of dollars (billboards). The high cost of a billboard seemingly detracts from our goal of ultra-low cost information dissemination, but the display's cost can be amortized over the number of people that might use it. Indeed, a printed sheet of paper priced at just a few cents that is only used once by one person could cost more in aggregate than an expensive billboard that disseminates information to thousands of people over a longer period of time.

Electronic integration. Normal visual displays for featherweight multimedia are non-electronic. However, more sophisticated versions might be inexpensively tagged to enable tighter integration with the associated electronic audio player. For example, RFID or optical tags might enable a talking book to determine which page is currently active, or optical codes on a poster could be used to index into auditory content on a camera-equipped audio player.

IV. USAGE EXPLORATIONS

We explored the viability of several forms of featherweight multimedia via usage explorations in different field settings. These were not intended to be formal experiments or extensive deployments, but rather initial forays. We experimented with two different domains (agriculture extension and healthcare), five device configurations, and five qualitatively different preliminary investigations. The investigations were intended to shed light on several questions we felt were critical to determine the overall promise of featherweight multimedia:

Q1. User literacy and education requirements: A major motivation for featherweight multimedia is that it relies on audio rather than text for communication. As such, we expect the technology to be usable by low-literacy populations. Is this in fact the case? Are there other cognitive barriers to use?

Q2. User engagement: In contrast to more full-fledged multimedia systems that provide a rich set of highly engaging interactive content, will the simpler and minimally interactive content provided by featherweight multimedia be sufficiently engaging that users will actually feel compelled to access information through such devices?

Q3. Social support requirements: Ideally, the device is usable by low-literacy populations on their own without any mediation from trained helpers or support from peers, but in reality, it is highly likely that some amount of social support will be required. To what extent is social support required for featherweight multimedia to be effective?

V. EXPLORATIONS IN AGRICULTURE EXTENSION

Our first exploration of featherweight multimedia was in the agriculture extension domain. We partnered with the Digital Green project [29], which seeks to teach targeted sustainable agricultural techniques to small and marginal farmers in India using video clips of their peers learning those same techniques. Their human-mediated video approach has shown higher adoption rates of the agricultural techniques as compared to traditional methods of dissemination such as paper posters or extension officers conducting 1-on-1 lessons.

The requirement of a TV and DVD player, however, poses challenges in terms of cost, portability, and electrical power. Interestingly, Digital Green's own assessments show that a version of their system where a poster is used in place of the TV/DVD can achieve much of the gains, though at a lower cost-effectiveness rate.

Instead of the TV/DVD, we experimented with several kinds of featherweight multimedia, and also with the devices in mediated and non-mediated contexts. In all cases, the audio was copied from video clips used by Digital Green, while the

video was replaced with a paper poster that illustrates the agricultural technique (Digital Green staff assured us that the audio was understandable without the original video). Since video is more visually engaging, we wanted to observe how featherweight multimedia fares in comparison.

Our explorations were conducted in a village in the state of Tamil Nadu, India. Attendance at the mediated meetings (Fig. 2), which take place in the evenings at a private home three times a week, stayed relatively constant at approximately 20 village residents throughout the meetings that we directly observed. The village comprises about 50 households or 160 individuals (50 men, 50 women and 60 children). Only 10% of the population could read or write.



Fig. 2. Mediator teaching agricultural information to village residents using featherweight multimedia – poster with audio player

A. 1st Prototype: Poster + Custom Audio, Mediated

Our featherweight multimedia device for this experiment was a custom audio device (Fig. 3, left) that was meant to be hung alongside a poster. When pressed, each of the seven buttons on the right play an audio file stored on a removable SD card. These audio clips are captions corresponding to mediator-drawn illustrations on the poster. On the left is a button that stops playback, below which is an on/off switch. For ruggedness and portability, the device is equipped with a handle and is covered with cardboard. A mp3 player [31], powered by three AA batteries, drives external speakers. The cost of this custom device was approximately US\$150, with nearly two-thirds going to the cost of the mp3 player.



Fig. 3. Custom audio player (left). MuVo V100 (right). Scales are not equal.

At the time of our research, eight Digital-Green-style meetings were conducted with this device (with no video). Usage of the device varied from taking up almost all of the hour-long meeting to almost none on several occasions. We attended four meetings and spoke with the mediator twice outside of the meeting time.

Observations and Feedback: Overall, both the mediator and the audience received the device well. We found that the audience did not get restless even when listening continuously to longer clips of 5 minutes or more, although this is surprising in light of previous research [29]. Half the audio clips on the SD card were played at each meeting. Although people did say they were tired of hearing the same content over and over again, this repetition has been shown to increase understanding. One village resident said that she could listen to the same audio two or three times before getting bored, and in fact needs to hear it that many times to fully understand the content. So, the featherweight device appeared to serve as a reasonable alternative to video.

Interestingly, no one else besides the mediator operated the audio player, or expressed interest in doing so. It was seen as something only the mediator should operate.

Meeting participants pointed out that since the device was custom-built by the authors, when it broke (for one meeting), a normal repairman could not fix it. They also commented that the device looked ugly. Perhaps the most significant design flaw was the device's lack of a pause button. Pressing stop "rewound" the audio to the beginning of the particular audio clip. For this reason, the stop button was rarely used. On the positive side, people from neighboring villages heard about the meetings and asked the mediator about them with interest.

B. 2nd Prototype: Poster + Off-the-shelf Audio, Mediated

Based on the feedback from the first iteration, we abandoned the custom audio player and switched to a commercially available one (Creative Technologies MuVo V100: Fig. 3, right). The commercial device is more durable, is capable of playing more than 7 audio files (10+ hours), and has pause/rewind/fast-forward capability. The total cost of the system was US\$38. Being a low-end device, the audio player's screen cannot show the name of more than one audio file at a time. In addition, it does not have playlists and does not support titles in the local language. The audio player allows grouping of audio files by folder but the user interface for switching between folders is difficult to use.

Therefore, we created a numbering scheme where the title of each audio clip consisted of two numbers separated by a hyphen. The first number is a code for the topic, for example (1=azolla, 2=vermicompost). The second number is the clip number within the topic. A printed guide was given to the mediator explaining the code. Navigation is accomplished by skipping forwards and backwards through the list of tracks by flicking a bi-directional self-centering spring-loaded switch. Since this is cumbersome with a large number of audio clips, we only loaded a few topics at a time onto the audio player.

Posters were designed similar to the first experiment, and

the sessions were again mediated.

Observations and Feedback: Despite some usability issues, this combination of device, poster, and mediation worked very well. In fact, Digital Green has been actively continuing the use of this device for five months, with over 50 meetings conducted to date. They report that adoption rates have been steadily increasing using our featherweight-multimedia solution, and that the rates are coming close to adoption rates with TV and video content. This, together with other benefits of the device such as portability, cost, and less reliance on power grid (Fig. 4), makes featherweight multimedia a powerful contender for this scenario.

The experiment has not been without problems: even after several meetings, the mediator had not mastered the bi-directional switch. One cause of the difficulty is the fact that the switch is functionally overloaded: when flicked, it switches between tracks, but when held, it fast-forwards or rewinds. This suggests that controls should be modeless, especially to accommodate users new to electronics.



Fig. 4. Featherweight multimedia working during a power outage!

C. 3rd Prototype: Talking Book, Non-Mediated

In this third exploration, we look at a more integrated featherweight multimedia solution: talking books. These devices are relatively cheap, have had some traction in developed countries for childhood education, are built robustly because they are designed for children, and provide access to audio and visuals via a single integrated package.

We repurposed a VTech WhizKid [10] device for agriculture information dissemination. The WhizKid can operate in a standalone fashion, or it can be hooked up to a laptop via USB port, in which case the WhizKid tablet and pen are treated as a mouse by the operating system. In the latter mode, it is possible to customize the content and interaction completely, so that an inserted page of our own design can be mapped to audio files we choose. (Should this solution work, the intent would be to work with the manufacturer to produce agriculture content for the device in its standalone mode.) The retail cost of the device without the laptop is approximately US\$25. The WhizKid has another benefit in that its robust, colorful design is less intimidating than a regular PC. During our studies, the laptop was hidden so as to not intimidate users. Screen shots from the video were used for visual content (Fig. 5), and relevant portions of audio were copied from the videos.



Fig. 5. An audio-augmented page for a talking book

We walked around a village on two different days, for a total of six hours, to various households and requested residents to try out the device in their homes. Besides the second author, a translator and a mediator were also present. Each user was given 20 minutes of free use to work with the device, and was allowed to continue if he/she wished. We provided no up-front instruction, but if the user failed to engage with the device at all after five minutes, the mediator or the translator provided gentle and low-grade assistance.

A feedback questionnaire was administered verbally that asked for the user's background information and opinions about the device. For the second day ($N=4$), we also administered a verbal test and feedback session after each use, to evaluate how much of the presented information the user had learned and to hear comments about the device.

Observations and Feedback: In spite of some initial hesitation, most of the villagers who tried the device were comfortable with it and could use it with little or no assistance. Those participants who took the pre- and post-tests showed increased knowledge of the content presented.

Over six hours spread over two days, we were able to find ten people (5 male, 5 female) willing to try the device to learn agricultural content (Fig. 6).

Individuals were very hesitant to press anything at first. The elapsed time between when an audio clip had finished playing and the next selection was chosen was initially as long as 10 seconds. With more use, this reduced to about one second. Many users repeated aloud certain pieces of audio information after they heard it spoken from the device.

Older users were slower to learn the device and slower to learn the content – they listened to all the clips at least twice. One user kept the pen pressed down while the audio was playing and never figured out that tapping would suffice.

On the first day, we had black-and-white printouts of the video. Color turned out to be essential for some of the audio descriptions, and so on the second day, we returned with color-printed sheets.

Users did not experiment very much. They never pressed the Pause button unless it was pointed out to them. They only played the audio in sequence, even while listening to the audio captions for the second time to review the material. When questioned later, users said that they feared that they might spoil the device.

During the post-experiment feedback session most of the users were able to answer the factual questions on the agricultural processes (mean of 75% answers correct for $N=4$). This confirmed that most users had been able to retain the key pieces of information that they had heard.

Crowds immediately gathered when a participant used the device. Many of the bystanders often prompted the primary user. Also, there were a number of instances when people who listened to the audio information enthusiastically participated in the post-experiment discussion. One user said that she would be willing to share the device if the device was left at a common place in the village and remarked that it will be quite useful on days she missed the Digital Green video-screening sessions. This suggests that featherweight multimedia can be complementary to other forms of information dissemination.



Fig. 6. A user navigating through the talking book for information.

D. Discussion

These explorations revealed some preliminary insights relative to our study questions.

In terms of literacy and necessary user background, featherweight multimedia appears able to bridge challenges of illiteracy. Audio playback seems sufficient for this purpose, and the addition of even static visual images helps anchor discussion in mediated sessions. This is not a surprise, but it is assuring that there were no unexpected surprises with respect to adults with little formal education being able to make sense of these devices. Many of the participants in this phase of the study had also been previously exposed to agricultural content on video from Digital Green, yet none explicitly expressed a negative comparison with video.

The simple interface of featherweight multimedia also seems to reduce barriers for the first-time user. Users related well to the use of local content and illustrations depicting local surroundings and people, enabling viable user engagement. In several cases, both in mediated and non-mediated use, it was seen that users repeated pieces of information to re-affirm their knowledge. This process of repeating the narration may enhance retention and assimilation of information. For us, it confirms that the users were paying attention in the first place.

Featherweight multimedia also seems sufficiently engaging so as to enable effective information dissemination, both under mediated and non-mediated circumstances.

In mediated sessions, audiences were not only willing to sit through uninterrupted audio playback of five minutes or more, in many cases, they requested repeated playback of the audio.

In non-mediated situations, users expressed initial trepidation handling the device. One remarked that he was afraid he might spoil the device, and others visibly hesitated in early interaction, until their confidence grew. In spite of the initial reluctance and limited direct guidance, most users were able to engage meaningfully with the device after just a few minutes of exposure to the device.

During post-experiment feedback sessions, several people expressed a desire for shorter audio clips that summarized longer content. On the other hand, longer audio clips engaged a greater number of passive users, attracted more passers-by, and generated more discussion because of the clips' longer playing time. Whereas longer clips sustain the attention of a larger group of people, shorter clips cater to an individual's short attention span.

With respect to the need for social support, a human facilitator is essential, at least in the initial introduction of the device, regardless of the device's perceived ease-of-use. In all of the mediated sessions, the audience fully took in the content without concerns about the device. In contrast, users in the non-mediated sessions needed prompting and encouragement to use the device at all.

The need for such mediation, however, appears to decrease with time and also with group interaction. When crowds gathered in non-mediated sessions, the actions of the primary user were influenced by people loitering around. They either voiced common requests or prompted him when the user was stuck, serving as encouragement. Such voluntary peer guidance further alludes to the effectiveness of social support.

VI. EXPLORATIONS WITH HEALTHCARE INFORMATION

In this second exploration, we sought a more institutional setting in which featherweight multimedia devices might be used. We partnered with Sankara Eye Hospital, located in Coimbatore, Tamil Nadu, India. It provides free eye-care surgeries to 700-1000 patients per week, funded by the much smaller segment of patients who can afford to pay. Sankara identifies non-paying patients by conducting 3-5 weekly outreach screening camps in villages. The patients requiring eye surgery are bused to the hospital, where they receive free treatment, room and board over several days, followed by a bus trip home. Two of the authors spent a week visiting Sankara Eye Hospital to find opportunities for deploying featherweight multimedia. We observed screening camps in rural areas and shadowed patients around the hospital.

After these visits, we decided to focus on improving Sankara's informational poster on post-operative care, which few patients were paying attention to despite the crucial nature of the information. This information is also read aloud by a nurse to the patients during discharge, indicating compatibility with featherweight multimedia.

The poster is 0.9m x 0.6m in size. It explains ten things patients should do on the left-hand panel, while on the right, it lists ten things they should not do (Fig. 7). The images on the poster are static cartoon drawings along the lines recommended by Medhi et al. [26], although the poster additionally includes text captions in Tamil. That paper's

findings that audio annotations enhanced comprehension of drawings provided further justification to our intent to add audio to these posters.



Fig. 7. Post-operative do's and don'ts poster. We added numeric codes to turn it into a "talking poster" when coupled with an audio player.

A. Technology

Based on our agriculture participants' difficulties with the MP3 player interface, we decided to explore a different device for this setting. We wrote a program that displays a large numeric keypad on a HTC Touch smartphone's touchscreen (Fig. 8). As numbers are typed in, they appear in the textbox at the top of the screen.



Fig. 8. HTC Touch running featherweight multimedia application. Unnecessary buttons below the touchscreen are covered.

On the poster, numeric codes are printed in each box (Fig. 7). When one of these codes is entered via the numeric keypad, the appropriate audio caption plays in the manner of museum audio tour guides. No "enter" or "play" button is needed, since the numeric codes are all the same length (in our case, two digits). There is also no "clear" button. If an unrecognized code is entered, both the textbox and the program's numeric buffer are cleared. These design decisions were made to simplify the user interface as much as possible.



Fig. 9. Ward at Sankara Eye Hospital where we conducted our observations.

B. Iterative Explorations

We conducted four sessions of testing with the device, each with slightly different usage scenarios and with a different group of patients, in a ward with hundreds of beds (Fig. 9). Around 200-230 patients were registered in each ward, but occupancy ranged from 70-230 at any time. Tests were conducted in the early evening when patients were most likely to be present. A sampling of 75 patients over the four sessions showed that the majority were illiterate; the few who were literate often had difficulty reading text due to poor eyesight.

All tests were conducted by one of the authors and a translator. We tried to interview anyone who interacted with the device or poster after they had finished.

1) First Session: Regular Poster, Non-Mediated

To establish a baseline, we set up a non-augmented poster, and sat where we could observe people viewing the poster. Over the course of a 90-minute observation, we estimate that 150-200 people (non-unique) passed by the poster. Of those, only 7 people looked at the poster (Fig. 10); the rest passed by without a glance. We questioned 5 of the 7, and among them, only one was able to answer simple questions about the content of the poster. Among the rest, one person even asked us whether it was permitted to sleep on the operated side, despite the fact that the poster expressly prohibits this.



Fig. 10. The regular poster, with no audio annotation: This is a rare patient who is actually taking the time to look at the poster. Our translator is the gentleman in white; on the left, two women are passing by without showing any interest.

2) Second Session: Audio Poster, Occasionally Mediated

We set up the poster with numeric codes in the same location as in session one, but this time we put the audio device in front of the poster on the table (Fig. 11), with speakers positioned behind. The backlight on the device was left on throughout the session.

We ran this session for 3 hours, during which time a total of 230 non-unique patients passed by the poster (we estimate 60-70 unique patients). Because it was immediately apparent (as with the first session) that very few patients would interact with the poster or the device unprompted, we tried a number of different things throughout the session to encourage them.



Fig. 11. Augmented poster set up with audio device and speakers

First, the translator used the device in an animated, demonstrative manner for three minutes, gesturing to the numbers on the poster and theatrically pressing the onscreen keyboard. Three patients gathered around, but after the translator finished demonstrating, the patients started to leave. We spoke to one of the three, and he was able to explain the purpose of the device and the correspondence between the poster, the printed numbers, and the resulting sounds. When prompted to try the device for himself, however, he demurred, saying that he was afraid to break the device.

During the second two hours, we played a pre-recorded Tamil-language audio prompt, “Learn about eye care by pressing the numbers,” every time a person passed or glanced at the poster, simulating a motion-sensitive trigger. The audio prompt never once achieved its intended goal, and the patients always left without trying the device.

Occasionally during the two hours, the translator again either demonstrated the device himself or actively encouraged patients to try it, and managed to get a total eleven people to try the device. Every time the poster was used, a crowd of 2-12 people gathered (Fig. 12), but would then disperse when the person operating the device stopped.



Fig. 12. A brave patient interacts with the poster (seen from the back) while a crowd looks on.

Nine of those who tried the device were interviewed. All but two had either glanced at the poster or had been in a crowd of people watching the device being used, then had been taught and encouraged to use the poster by the translator.

The other two were instances of patients teaching other patients how to use the device. In one instance, a young man in his mid-20s who was accompanying one of the patients used the device through encouragement by the translator. He then taught another man of the same age, who was also there to accompany a patient. He used the device for several minutes. One of them then came back an hour and fifteen minutes later to continue interacting with the device.

In the other instance, a man and a woman stood in front of the poster. The woman was about to use the device, but the man tried it first. After each audio caption is played, they nodded approval together. The woman then took over, and a crowd of six gathered. She pressed and listened to a few captions. Another woman joined her, and the first woman systematically pressed all the numbers, taking 6 minutes. After each caption was played, the first woman made sure that the second woman understood the content.

There were at least two people who were not able to understand the use of the device even after the translator's explanation.

For this group of patients, active human mediation seems necessary even to get to a point where patients will try the device, and even then only a small minority will do so. However, more people expressed curiosity and stayed to listen if someone else was operating the device. The audio prompt to invite usage (without mediation) was a failure.

3) *Third Session: Audio Poster, Mediated*

In this session, we asked a nurse to explain the usage of the audio poster to all patients in the ward over the PA system. Immediately afterwards, a crowd of about 12 patients gathered around the poster, and the nurse demonstrated to that group. Individual patients in the crowd then tried using the device assisted by the nurse. Within eight minutes, the crowd dispersed. A second PA announcement an hour later did not result in significant sustained interest.

4) *Fourth Session: Audio Brochure, Peer-Mediated*

For the fourth session, we tried a more intimate form factor and a different kind of mediation. Since patients spend a lot of time sitting on their beds waiting for treatment or recovering, we decided to take the content to them instead of having them come to the poster. To this end, we created an audio brochure, which is simply the poster shrunk onto A4-sized paper that could be passed around. All words and numbers remained legible, and the audio-guide device remained the same.

On the first day of this experiment, we taught one of the male patients how to use the device. The patient was asked to use the device for himself, teach others to use the device, and then to pass the device on. The patient took on his responsibility eagerly, perhaps overly so. Over the course of 10 minutes, he demonstrated the audio brochure to 11 patients, letting them hear a few captions before moving on to another patient, without relinquishing the device.

The next day, the translator taught a group of several patients to use the device and pass it along (Fig. 13). After a few minutes, the group taught a group of women sitting across from them how to use the device. Two women were then observed sitting at the feet of the men and having a discussion

about the device and the health information. One of the men in the group was a retired 70-year-old teacher. Five minutes later, the teacher was gesturing animatedly to the assembled people around him, demonstrating the device, but keeping control of it. The translator then told the teacher to encourage others to use the device rather than to use the device himself as a tool to teach the content. The teacher then moved to different parts of the ward, teaching other groups and having discussions with them (Fig. 14). Most of the time he did not relinquish control. Often, a member of the group pointed to a number on the brochure, which the teacher then typed in. A few times, other people were observed pressing the numbers themselves. Thirty-five minutes after the teacher got control, he returned the device and the brochure to us, and reported that he had completed teaching all 30 patients who were awake!



Fig. 13. Translator (standing) teaching a group of patients how to use smartphone with brochure



Fig. 14. Patient (a retired teacher) teaching other patients

We interviewed 10 people (besides the teacher) who had either tried the device or learned from the teacher. For the most part, they reported that they felt that the information was very important and were able to recall some key points. There were no negative responses, and no one said that they had difficulty in using the device. One woman even expressed a desire to buy the device, but said that unfortunately she did not have enough money on her!

C. Discussion

The social framework that we found to work in institutional healthcare settings involves (1) approaching groups, (2) using a portable form factor (in this case the brochure/smartphone), (3) providing small-group tutorials with instructions to the group to teach others. In some cases, one member of the group will take the initiative and take charge of teaching others. We make several further observations:

First, technical manipulation of the interface was not a problem for most of the patients (except for two) who experienced the device. In fact, even patients who did not directly interact with the device, but instead observed others operating it, were able to understand what it did and to associate the audio with the visual information.

Second, user engagement varied greatly depending on the person and on the circumstances of device introduction. Most patients seemed reluctant, even afraid, to use the device, particularly when it was placed in a public location with a lot of patient traffic, even if they understood how to use the device. Patients said that they were fearful of breaking the device. This is consistent with previous findings on interfaces for the illiterate, where inexperience with electronic devices creates a barrier for first-time users [28]. These fears were not allayed by audio prompts from the device.

On the other hand, when given private tutorials, some users would all but monopolize the device, appointing themselves both teacher and technology expert for their peers. Between these two extremes, there was a range of engagement patterns – one point of note was a tendency for crowds to gather around a person who was willing to operate the device, and then to disperse when the person stopped.

Third, it seems clear that human mediation of some kind is essential for the device to be used, at least initially. The public poster itself was largely ignored, and the device was used only when our translator encouraged people to try it. On the other hand, peer tutorials were very effective in encouraging use.

Fourth, the brochure was more effective than the larger poster. This may have been because patients perceived it less like hospital property, or because it was easier to pass around.

Finally, we observed several usability problems with the smartphone application, but once patients were taught how to overcome the problem, they continued using the device. One patient slid his finger between numbers while another pressed too hard. There were also mis-registration problems with the touchscreen, causing numbers to be mistakenly pressed twice or the adjacent number being pressed instead. Most of these issues would be remedied by using a large, physical numeric keypad instead of a touch-screen virtual keypad.

VII. FUTURE WORK

We hope to better understand how to integrate featherweight multimedia into development-focused programs and organizations, as well as to measure their educational impact in comparison to other forms of media.

We are also exploring a possibility of a low cost “kit” that can be used to construct featherweight multimedia devices in which the audio and visual content are customizable. This kit

can be easily modified to create devices of varying sizes, form factors (from A5 to poster sizes), and costs. The kit would consist of a board with an array of buttons that trigger audio playback. Sheets of paper could be overlaid onto this board with playback icons positioned over the underlying buttons. This kit would cost US\$20-25 for an A4-sized “audio tablet” with a 1GB SD memory card. However, the kit would cost only US\$0.50 to make an A5-sized “audio card” that contains 10-15 seconds of audio content.

VIII. CONCLUSION

In this paper, we presented a broad exploration of featherweight multimedia – combinations of electronic audio and static visuals – that can be put together for significantly lower cost than even very low-cost PCs, but can nevertheless provide a rich, interactive multimedia experience.

Within the contexts of agriculture extension in rural villages and healthcare information dissemination in an eye-care hospital, we tried exploratory studies with five different configurations of featherweight multimedia devices, all with illiterate and semi-literate users. Over 75 people directly interacted with the devices, and more than 150 people participated in formal or informal learning sessions in which they observed the use of the devices.

The results from this preliminary exploration suggest that most illiterate users can become immediately comfortable with the simple interfaces provided by featherweight multimedia, although fears of breaking the technology and such can create barriers to casual usage. At the other extreme, some users, especially when given short, private tutorials on usage, seem to delight in using these devices and showing them off to peers. We also note that, consistent with much of the literature on technology for development, that social support for the technology is an essential component of impactful usage.

IX. ACKNOWLEDGEMENTS

Thanks to Joice Sister, M. Saravanan, and Dr. R.V. Ramani at Sankara Eye Hospital; Mahalaxmi and Srikantamurthy at Digital Green/GREEN Foundation; S. Sivaraju, Ajay Khanna, Kalai Arasi, Meera Lakshmanan, and Avinash Setty for translating; and Sid Vishwanathan & family, Leah Findlater, and Gursharan Singh.

X. REFERENCES

- [1] One Laptop Per Child. [Online]. laptop.org. [Accessed Aug 31, 2008].
- [2] Classmate PC. [Online]. classmatepc.com. [Accessed Aug 31, 2008].
- [3] International Team Aims to Develop a \$12 Laptop During MIT Summit. [Online]. chronicle.com/wiredcampus/article/3219/international-team-aims-to-develop-a-12-laptop-during-mit-summit. [Accessed Aug 31, 2008].
- [4] Lourie, S. (1990). World literacy: where we stand today - One billion non-literates. Editorial, *UNESCO Courier*. July 1990.
- [5] Medhi, I., Sagar, A., Toyama, K. (2006). Text-free user interfaces for illiterate and semi-literate users. *Proc. ICTD*, p. 72-82.
- [6] Books of Hope. [Online]. booksofhope.com. [Accessed Aug 24, 2008].
- [7] Q & A about Speaking Books. [Online]. booksofhope.com/pdf/Q%20%20A%20about%20Speaking%20Books%20.pdf. [Accessed Aug 24, 2008].
- [8] SADAG's school based curriculum evaluation. [Online]. booksofhope.com/research/SADAG%27s%20School%20Based%20Curriculum%20Evaluation.pdf. [Accessed Aug 24, 2008].
- [9] LeapFrog. [Online]. leapfrog.com. [Accessed Aug 24, 2008].

- [10] VTech WhizKid . [Online]. vtechwhizkid.com. [Accessed Aug 24, 2008].
- [11] Fisher-Price PowerTouch [Online]. fisher-price.com/us/powertouch/default_flash.asp. [Accessed Aug 24, 2008].
- [12] HHS to provide new interactive book of health information to women of Afghanistan and their families. [Online]. hhs.gov/news/press/2004pres/20040803.html. [Accessed Aug 24, 2008].
- [13] PicTalk. [Online]. edindiasolutions.com/index-4.html. [Accessed Sep 20, 2008].
- [14] Literacy Bridge. [Online]. literacybridge.org. [Accessed Aug 24, 2008].
- [15] Tour-Mate. [Online]. tourmate.com. [Accessed Aug 24, 2009].
- [16] Prentke Romich Company. [Online]. prentrom.com. [Accessed Aug 24, 2008].
- [17] Stifelman, L, Arons, B., Schmandt, C. (2001). The Audio Notebook. *Proc. CHI*. p. 182-189.
- [18] Livescribe. [Online]. livescribe.com. [Accessed Aug 24, 2008].
- [19] Daisy. [Online]. daisy.org. [Accessed Aug 24, 2008].
- [20] Dolphin. [Online]. yourdolphin.com. [Accessed Aug 24, 2008].
- [21] Tang, A., Finke, M., Blackstock, M., Leung, R., Deutscher, M., Lea, R. (2008). Designing for bystanders: reflections on building a public digital forum. *Proc. CHI*. p. 879-882.
- [22] Vogel, D., Balakrishnan, R. (2004). Interactive public ambient displays: Transitioning from implicit to explicit, public to personal, interaction with multiple users. *Proc. UIST*. p. 137-146.
- [23] Plauche, M., Nallasamy, U., Pal, J., Wooters, C., Ramachandran, D. (2006). Speech recognition for illiterate access to information and technology. *Proc. ICTD*. p. 83-92.
- [24] Sherwani, J., Ali, N., Mirza, S., Fatma, A., Memon, Y., Karim, M., Tongia, R., Rosenfeld R. (2007). HealthLine: Speech-based access to health information by low-literate users. *Proc. ICTD*. p. 131-139.
- [25] Parikh, T., Javid, P., Sasikumar K., Ghosh, K., Toyama, K. (2006). Mobile phones and paper documents: Evaluating a new approach for capturing microfinance data in rural India. *Proc. CHI*. p 551-560..
- [26] Medhi, I., Prasad, A., Toyama, K. (2007). Optimal audio-visual representations for illiterate users of computers. *Proc. WWW*. p. 873-882.
- [27] Digital StudyHall. [Online]. dsh.cs.washington.edu/info/papers.html. [Accessed Aug 24, 2008].
- [28] Medhi, I., Toyama, K. (2007). Full-context videos for first-time, illiterate users. *Proc. ICTD*. p.140-148.
- [29] Gandhi, R., Veeraraghavan,R., Toyama, K., Ramprasad, V. (2007). Digital Green: Participatory video for agricultural extension. *Proc. ICTD*. p. 21-30.
- [30] Talking Greeting Card. [Online]. simaproducts.com/products/product_detail.php?product_id=632. [Accessed Aug 31, 2008].
- [31] μ MP3 Playback Module. [Online]. roguerobotics.com/products/electronics/ump3. [Accessed Aug 24, 2008].