User-Centered Design and Evaluation of a Concurrent Voice Communication and Media Sharing Application

David J. Wheatley

Motorola Labs, Social Media Research Lab,
1295 E. Algonquin Road, Schaumburg,
Illinois, 60196, USA
David.j.Wheatley@motorola.com

Abstract. This paper describes two user-centered studies undertaken in the development of a concurrent group voice and media sharing application. The first used paper prototyping to identify the user values relating to a number of functional capabilities. These results informed the development of a prototype application, which was ported to a 3G handset and evaluated in the second study using a conjoint analysis approach. Results indicated that concurrent photo sharing was of high user value, while the value of video sharing was limited by established mental models of file sharing. Overall higher ratings were found among female subjects and among less technologically aware subjects and most media sharing would be with those who are close and trusted. This, and other results suggest that the reinforcement of social connections, spontaneity and emotional communications would be important user objectives of such a media sharing application.

Keywords: User centered design, wireless communications, concurrent media sharing, cell-phone applications.

1 Introduction

The rapid and widespread acceptance of camera enabled cell phones has introduced new opportunities for the immediate and spontaneous sharing of still visual images using the cellular network and such capabilities are now almost ubiquitous. In fact, in 2004, camera phones outsold digital cameras by nearly 4 to 1 [11], with an estimated 60 million photos taken daily. There is also an increasing trend towards public photojournalism. The latter was well illustrated in the context of the London Underground bombings where still and video images captured by personal cell phones were, for the first time, extensively used by the broadcast media [3]. In fact, approximately 1000 photos and 20 pieces of video were sent in to the BBC News website [13]. This capability for the spontaneous creation and immediate sharing of visual images via wireless networks is changing the landscape of picture usage, both in terms of the content of images as well as the motivations for sharing [5, 12]. A next step would be to enable this real-time sharing of visual media (photos and video clips) to take place concurrently with a one to one or multi-point (group) conversation.

J. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2007, LNCS 4552, pp. 990–999, 2007.

Since these studies were completed, a number of concurrent media sharing applications have been, or will be introduced in 2007, such as Ericsson IMS "weShare" [1] and Motorola's European 3G Video-sharing solution. This paper describes how user centered design methods were applied to design, develop and evaluate an integrated application such as these, through a process of use case definition, paper prototyping, user interface design and finally, user evaluation of a handset based prototype.

2 Related Research

Evidence suggests that the rapid uptake and use of camera-enabled cell phones might have been driven by quite different motivations than was the case with traditional film or digital cameras. Image quality and careful photo composition appear to be of relatively low value while the fleeting capture of serendipitous and everyday events seem to be a fundamental user value created by the ubiquitous nature of the device. In a survey by IPSe Marketing in Dec 2002, the majority of camera phone users (42.4%) reported that they took photos of *"things that they happened upon that were interesting"*, this was followed by family members (39.5%), friends (36.6%), self (26.4%), pets (23.7%) and travel photos (21.5%) [4]. This same survey of 2007 Japanese respondents also concluded that *"nearly half"* had taken a photo in place of jotting a memo or sketching something on paper. The value of the photo and video enabled cell phone in capturing unexpected events has also not gone unnoticed by the news media, NBC News *"believes in it so much that they've begun equipping reporters and other staff members with video enabled cell phones….. [since] you never know where or when news is going to happen"* [4]. The BBC has also been formally evaluating such mobile imaging technologies [13].

In addition to the immediate and pragmatic usefulness of camera phones, sharing pictures with other people also frequently has a significant personal and emotional component [1, 9]. However, in a study by Kurvinen of four groups of 5 subjects sharing digital images, *"practically all of the messages sent … contained both images and text"* in order to fulfil these emotional needs and to assist the recipient in interpreting the visual image [6]. Kurvinen also found that the capability to fulfil this emotional need produced much of the value derived from sharing sequences of pictures in a turn-by-turn process of group communication. A number of papers have concluded that social/emotional communication is a key objective of mobile media sharing [4, 6, 9, 10]. One of the aims of these studies was to develop a prototype media sharing application and to evaluate how this user objective might be facilitated.

3 Methods

3.1 Phase 1 – Paper Prototyping

This process of user centered development, consisted of two phases. In the first phase, carried out in partnership with Purdue University, use case scenarios were developed to communicate the hypothesised functionality of a concurrent voice and media sharing application and were decomposed into seven, sequentially dependent user

tasks. Paper prototypes, representing the operations and screen flows required to complete each of these tasks, were designed based on the Motorola "Tactium" touch screen interface (see Fig. 1). Paper prototypes were specifically chosen in this phase so that they would be perceived as being very early in the development process such that subjects would be more willing to provide both positive and negative feedback to influence the development. If the prototype were perceived as being more finished then they might consider the qualitative feedback to be less influential in the development and would consequently be more reluctant to be critical.

Individual interview sessions were held in which these seven task scenarios were presented visually on a laptop PC, then carried out by subjects using the paper prototypes. These scenarios probed user values associated with the following functions;

- *Initiating a group voice call*
- *Availability of presence information*
- *Sharing still images concurrent with a group voice call*
- *Sharing live video concurrent with a group voice call*

After completing each task, subjects were asked a series of open-ended questions relating to each task and the functional capabilities represented. They also completed a short questionnaire (TAQ) to assess their level of technological awareness. This consisted of 12 alternative choice questions relating to the ownership and frequency of use of a number of representative portable media devices and services. Within this, subjects also self-reported levels of interest and capability in acquiring and learning about new electronic devices. The questionnaire was numerically scored, the highest score being 42. The mean score was 20.9, with values ranging from 1 to 42.

In order to reduce the qualitative data (from the open-ended questions) to a more readily analyzable quantitative form, a set of questions were generated from the response data, and the dichotomised answers to these questions coded. In this process 84 unique codes were generated, reflecting the 84 most frequently raised issues. The content of each paragraph unit of transcribed text was then coded, based on consensus by three or more of the experimental team, in order to facilitate quantitative analysis within a relational database, which reduced the qualitative data to a total of 1371 codes. The coded data was analysed using the Eztext/Answer qualitative analysis software suite, produced and widely used by the CDC (Centers for Disease Control) for analysis in medical projects [2].

Subject sampling was intentionally biased towards younger subjects, for whom using a cell phone and other communications and media rendering devices would be a familiar and integral part of their everyday lives. They also represented the early adopters for whom such an application would likely have relatively high value. The sample of 23 subjects consisted of 12 female and 11 male, (mean 20.7 yrs). The age range was from 18 to 28 yrs.

3.2 Phase 2 – Application Prototyping

The phase 1 findings were used to specify the functional capabilities of a handset based concurrent voice and media sharing application, the objective of this study

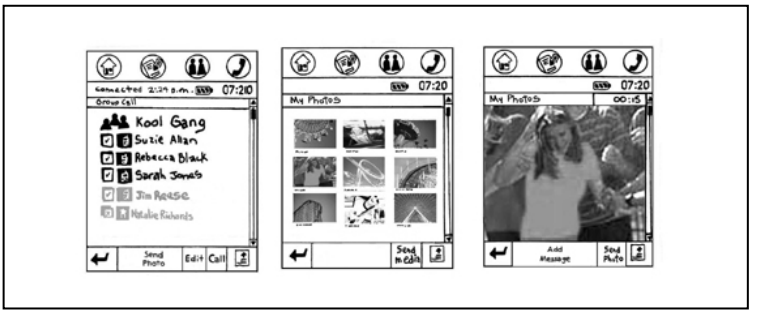

Fig. 1. Examples of touch-screen interface simulated in paper prototype

being to assess the relative impact of a number of functions and price, on subjective ratings of the application. This prototype was developed using J2ME on a Motorola E1000 3G handset and operated over a 3G test network. The application prototype represented the following functionality:

- *Separate individual and group contacts listings*
- *Presence/availability information presented within the contacts lists*
- *The ability to select and send a still image concurrent with a one-to one-voice call and with no voice call (similar to MMS)*

Where possible, existing screens and menus were used in the design of the prototype while new screens were designed to have a similar look and feel (Fig. 2). Specific task flows were developed to be as intuitive as possible in order that subject responses would be focused primarily on the overall functional capability and value of the application rather than the specific UI implementation. Subjects were presented with scenario storyboards to set the context for each task and to illustrate the functions and capabilities of the prototype. They then carried out each of six tasks using the prototype. Each task demonstrated a different permutation of the key functions within a conjoint analysis experimental structure. In order to control for order effects, two sequences of presentation of the task scenarios were defined and used with alternate subjects. After completing each task, subjects were asked how much they would rate that version of the service if it was free, and also what their rating would be if it added $10 to the monthly cost of their cell phone bill. In each case, they gave their rating on a 7-point Likert scale. After each of the two task groups (sending and receiving media), subjects were asked a series of open-ended questions to qualitatively explore the expected contexts of use. The task completion and open-ended question responses were recorded on both video and audio for later analysis.

Each task consisted of a combination of attributes on three dimensions. These dimensions were:

Contact List: Calling using Group phonebook (“G”) vs. individual phonebook (“Ind”)
Presence Information: Available before calling (“PI”) vs. not available (“No PI”)
Media Sharing: Concurrent media sharing with a call (“CMS”) vs. media sharing only outside a call (“No CMS”)

Fig. 2. Examples of prototype application screens

Subjects were first presented with the highest and lowest versions of the service, with the highest including all three attributes (G, PI and CMS), and the lowest including none of these attributes. These initial versions were intended to convey to the subjects the extreme ranges of the service, so that they would initially use the high and low ends of the response scale. Then four other versions were presented that combined the attributes in a partial factorial design. The sample of 12 subjects was selected using similar criteria to phase 1 and consisted of 6 males and 6 females. Two of these subjects did not report their ages, but the reported average age of the remaining 10 was 23.4 (s.d. 2.1), and their median age was 23. The youngest subject in this sample was 19 and the oldest was 26.

4 Results and Discussion

Results from Phase 1 indicated that concurrent media sharing was very positively regarded by subjects: with 91% positively evaluating photo sharing and 87% positively evaluated video sharing. In addition, 26% ranked media sharing as their favorite function of those presented in the study; interestingly, all of these were low and medium technology awareness users – a trend which was to be repeated in the second phase of the study. The positive evaluation was relatively equal across genders, and slightly correlated with technology awareness.

In terms of the intention to share photos concurrent with a call, 78% stated that they would actually do it but here there were clear differences between genders and technology awareness levels. 92% of the females (compared to 64% of the males) and 100% of the highly sophisticated users (compared to 66% of the low and medium users, combined) indicated that they would actively use (in addition to value) photo sharing. Nonetheless, technologically sophisticated subjects did not support the concept without qualification. One of them felt that although photo and video sharing was *"a good idea. . . basically [it] depends on technology at the moment, transfer speeds, wireless connections, processing power and battery life."*

As well as technical limitations, image quality was a recurring theme for technologically sophisticated subjects. Many were concerned about the resolution

needed to provide acceptable image quality. Many also indicated that they would use media sharing for reasons of convenience, particularly that this would enable the sharing of daily and unexpected moments with friends or family. One subject concluded, however, that despite technical drawbacks, *"It would be cool. It would be an interesting way to share random things with your friends, attach some text, 'Oh, something cool happened.' "* This seems to confirm the value of spontaneity in mobile media sharing predicted by previous research.

When sending and receiving media, subjects were generally trusting in the honesty and discretion of their content recipients, an overwhelming majority (91%) believing that the media recipients should be able to forward and/or save that media. Interestingly, males were, relatively speaking, slightly less trusting. Only 82% of the males, compared to 100% of the females, indicated that they would have no objection to the media being forwarded by the recipient. In the end, 35% expressed an interest in having some kind of media "locking" function, which would enable control over whether it could be further forwarded. These results suggested that the application prototype should focus on enabling still image sharing concurrently with a group voice call, the ability to initiate a call from either an individual or group contacts list and presence information.

In phase 2, subjects showed high agreement on the top-rated scenarios, but less agreement on lower-rated ones. When the service was free, the highest rated scenario included all three attributes—G, PI, and CMS—with an average of 6.6 (median 7.0, s.d. 0.79) in fact 11 of the 12 subjects gave this their top rating. The scenario that excluded all attributes had the lowest average rating (ave. 5.5, median 5.0, s.d. 1.5). When there was a monthly charge of $10 for the service, the highest-rated scenario included PI and CMS, but excluded G. It had an average rating of 4.5 (median 4.5, s.d. 1.4).

The impact of the service attributes on the conjoint ratings was assessed using ANOVA, the Attribute factor having 3 levels: type of contact list (G vs Ind), Presence information (PI), and Concurrent Media Sharing (CMS). The results showed significant main effects of Price ($F(1,11) = 34.7$, $MSe = 0.95$, $p < .001$, $\eta^2 = 0.76$) and a trend toward a significant effect of Attribute Present ($F(1,11) = 8.94$, $MSe = 0.17$, $p < .05$, $\eta^2 = 0.45$). Ratings were significantly higher when the service was free than when it cost $10/mo (5.96 vs. 3.87), and were higher when the attributes were present than when they were absent (5.13 vs. 4.69). A series of *t*-tests did show some significant or near-significant effects of the individual attributes. When the service was free, there were marginally significantly higher ratings for CMS than for no CMS ($t(11) = 1.91$, $p < .05$; Fig. 3). When the service cost $10 per month, there were significant preferences for CMS ($t(11) = 3.28$, $p < .01$) and marginally significantly higher ratings for PI ($t(11) = 2.25$, $p < .05$).

In order to assess the impact of the grouping variables on ratings of the functional permutations, a stepwise regression was performed on the transformed conjoint data. The overall ANOVA for the regression was statistically significant ($F(13, 77) = 82.3$, $p < .001$, $MSe = 0.021$, $R^2 = 0.93$). There was a significant effect of Cost ($t(77) = 15.7$), with the free versions of the service being given higher ratings than the $10 versions. The following were some of the significant 2-way interactions:

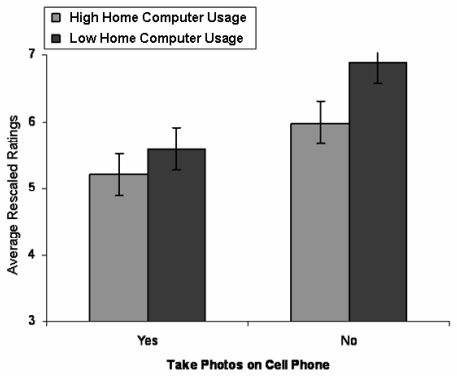

Fig. 3. Average Rescaled Ratings for each Attribute and Price

Gender X Technology Awareness. There was no significant difference between males and females who were high in technical awareness, however, of those who scored low on this scale, females gave significantly higher average ratings than males ($t(11) = 3.67$, $p < .01$).

Home Computer Usage X Cell Phone Photos. Subjects who did not take photos on their cell phones and were low on the computing usage variable gave significantly higher ratings than users who did take photos on their cell phones and were high on computing usage ($t(11) = 5.26$, $p < .001$; Fig. 4).

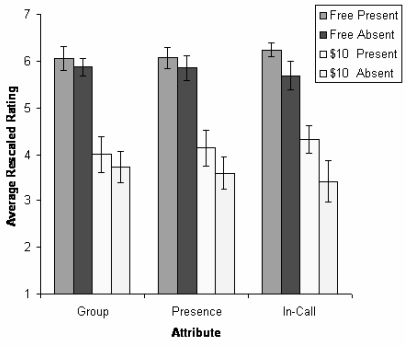

Fig. 4. Average Rescaled Ratings for each level of Home Computer Technology Usage and use of Cell Phone for taking Photos

Technology Awareness X Cost. Among subjects who were low in technical awareness, the ratings were significantly higher when the service was free than when it cost \$10/mo ($t(11) = 2.73$, $p < .01$).

Gender X Cost. For males, ratings were significantly higher when the service was free than when it cost \$10/mo ($t(11) = 2.86$, $p < .01$). Females gave significantly higher ratings when the service was free than males did when the service cost \$10/mo ($t(11)$

= 3.22, $p < .01$). However, there was no significant difference in ratings between females with the $10/mo service compared to males with the free service.

5 Conclusions

Phase 1 demonstrated that the concept of concurrent media sharing was of significant user value and that photos were slightly more valued than video. It was found that many subjects did not fully grasp that the video sharing was "live"; but defaulted to a file sharing mental model with the expectation that a video clip would be captured, saved and then shared as a video file, possibly contributing to the lower score. The more technologically aware subjects were also conscious of potential technical limitations of such media sharing, which may also have been a factor in this group rating the function somewhat lower than the less technologically aware subjects. Gender differences were also found in that female subjects generally rated photo sharing higher than males; suggesting a social element also found in phase 2. The group calling capability also revealed concerns about the accuracy of presence information and the unambiguous knowledge of who was in the group call. The latter was found to be most important when sharing visual media.

On the whole, subjects were willing for recipients to be able to save and/or forward media, but there was also an interesting process of self-censoring, in that they would simply not share media which was sensitive or private and/or they would not share media with those whom they did not trust. Despite this, there was also interest in the concept of 'locking" media to control or limit whether it could be saved or forwarded. There was also mixed reaction to adding text messages to shared media, particularly within a concurrent voice call, which contrasts with that found in [6]. This negative reaction seemed to be based on two factors; the difficulties inherent in text entry using a mobile device and the redundancy of a text message during a concurrent voice call. As one subject described it; *"adding text would take too long, it would be such a hassle, especially if I could tell them on the 'phone what the caption would be"*. For the same reasons, the ability to personalize or modify the media (with borders, word balloons etc.) was also felt to be somewhat irrelevant. There were also concerns about the physical operation of the handset, arising from a necessity to hold the device in the hand to look at the screen and/or operate the touch screen (to select and share media) while simultaneously involved in a voice call (with an expectation of holding the device to the ear).

Results from phase 2 further confirmed that incorporating concurrent media sharing was likely to add significant value to wireless communications services. While there was a trend suggesting that adding *any* of the attributes tested would increase the value for users, concurrent voice and media sharing were the only individual attributes that significantly increased subjects' ratings. The other attributes, Presence and Group Calling, did not significantly increase subjects' value judgments. A second finding was that value for the service was inversely related to subjects' technology usage and awareness. As in phase 1, there were higher value judgments among low technology aware users. This also appeared to play a role in the results, interacting with gender. While females gave higher value judgments than

did males (as found in phase 1), this effect was limited to subjects who were low in technology awareness, as measured by the TAQ.

These results could arise because individuals who do not use technology as much may consider more of the social implications of in-call media sharing than its value as a new technology. The interaction with gender also fits in with this hypothesis. This interaction may be influenced by males basing their judgments on the perceived usefulness of the technology for job performance [14]. Low technology-aware females, on the other hand, may primarily consider the social possibilities that would be afforded by the technology [14]. If this is the case, then the capability of enhancing social contacts could contribute positively to the value that these females place in the application. These conclusions about gender differences may also account for an interaction between gender and price. While females' average ratings for the free service were significantly higher than the males' ratings for the $10 service, their ratings for the $10 service did not significantly differ from the males' ratings for the free service. One possible conclusion from this result is that females are more willing than males to pay to receive the social benefits of this technology.

Overall, these results suggest that providing the capability of sharing media concurrently with a group voice call does enhance the value of mobile phone services for some users. However, this increased value may depend on those users' goals and it seems to provide added benefit mainly for users who are interested in using media content to supplement the social aspects of their communications. The results also suggest that more technologically sophisticated subjects may have been less impressed by the functional capabilities and that this led them to assign lower ratings.

Acknowledgments. Thanks are due to Prof. Sorin Matei and Wendy Zeitz-Anderson of Purdue University, IN, for undertaking the phase 1 data collection and analysis, and to Lynne Ferguson and Jim Wolf of Motorola Labs and the staff of Motorola Networks, Arlington Heights, IL, for permitting the use of the 3G test network for phase 2.

References

1. Bjorling, M.E, Carlsten, J., Kessler, P., Kruse, E., Stille, M.: Sharing Everyday Experiences. Ericsson Review No 1 (2006)
2. CDC EZ-Text; CDC Home: www.cdc.gov/hiv/software/ez-text.htm
3. Gough, P.J, Marlowe, C.: Cell Phone Video first from London Bombing Scene, Hollywood Reporter.com (8 July 2005) http://www.hollywoodreporter.com/thr/new_media/article_display.jsp?vnu_content_id=1000975698
4. IPSe Marketing Inc.: The Mobile Phone Morphs into Camera-equipped email Terminal, Online Report (Feb 21 2003) http://www.ipse-m.com/company/release/release_02_e.htm
5. Ito, M., Okabe, D.: Camera phones changing the definition of picture-worthy. Japan media Review, (29 August 2003) http://www.ojr.org/japan/wireless/1062208524.php
6. Kurvinen E.: Emotions in Action: A Case in Mobile Visual Communication. In: Proceedings of the Design + Emotion Conference, Loughborough, UK (1-3 July, 2002)
7. Lehtonen, T-K., Koskinen, I., Kurvinen, E.: Mobile Digital Pictures – The Future of the Postcard? Findings from an Experimental Field Study. In: Laakso, V., Östman J. Postcard in the Social Context. Korttien talo. Hämeenlinna (2002)

8. Matei, S., Zeitz-Anderson, W., Wheatley, D., Ferguson, L.: CEC/Ensembles User Requirements Study – Final Report, Motorola Internal Report (October 4, 2004)
9. Pering, T., Nguyen, D.H, Light, J., Want, R.: Face-to-Face Media Sharing Using Wireless Mobile Devices. In: Proc. Of 7th IEEE Int'l Symposium on Multimedia (ISM '05) (2005)
10. Sarvas, R., Viikari, M., Pesonen, J., Nevanlinna, H.: MobShare: Controlled and Immediate Sharing of Mobile Images. ACM Multimedia, New York (October 10-16, 2004)
11. Strategy Analytics, Camera Phone Sales Surge to 257 Million Units Worldwide in 2004 (April 14, 2005)
12. The Pitch: The Cameraphone could be the Next Little Big Thing, The Pitch: Special cameraphone Supplement, Issue 9.1 (December 2003) www.the-pitch.com
13. Twist, J.: Mobiles Capture Blast Aftermath. BBC News World Edition (8 July, 2005)
14. http://news.bbc.co.uk/2/hi/technology/4663561.stm
15. Venkatesh, V., Morris, M.G.: Why Don't Men Ever Stop to Ask for Directions?, Gender, Social Influence and their role in technology acceptance and usage behavior. MIS Quarterly 24(1), 115–139 (2000)