

Concept and Usability Testing Methodologies for Interactive TV

D.J. Wheatley

Motorola Applied Research Center
1295 E. Algonquin Rd.
Schaumburg, Illinois, USA
+1 847 538 1577

David.j.Wheatley@motorola.com

S. Basapur

Motorola Applied Research Center
1295 E. Algonquin Rd.
Schaumburg, Illinois, USA
Telephone number, incl. country code
sbasapur@motorola.com

ABSTRACT

The evolution of TV from a passive media consumption to a more complex interactive experience requires that greater attention be paid not only to the UI design but also the approaches to user evaluation. Furthermore, time/errors based methodological approaches traditionally adopted in a task-based productivity context may be inadequate to effectively evaluate this more social and entertainment use context. This paper describes two quite different user studies, with different objectives, carried out to evaluate a TV-based video communications system.

Categories and Subject Descriptors

H.4.3 [Communications Applications]: Computer conferencing, teleconferencing, and videoconferencing; H.1.2 [Models & Principles]: User/Machine Systems - Human factors; H.5.2 [User Interfaces]: User-centered design, Evaluation and Methodology

General Terms

Experimentation, Measurements, Human Factors, Usability.

Keywords

Video mediated communications, system usability scale, video telephony, user study, usability evaluations.

1. INTRODUCTION

Television viewing has traditionally been a passive media consumption experience with little interactivity beyond that of changing channels and more recently, DVR trick play. The introduction of more complex interactive functionality is not only changing and expanding the experience, but is creating more complex user-TV interactions which demand more design development and user evaluation. Two evaluation methodologies are described which were used in the development of a TV-based videotelephony system and which had quite different objectives. The first study set out to evaluate the user value of the concept of videotelephony via the domestic TV, including when, why and for what purposes such a system would be used and what features and functions might be required. A prototype system was developed based on the findings of this concept study and, in the second study, two variants of the system were evaluated for usability.

2. CONCEPT STUDY METHODOLOGY

The concept study set out to evaluate the user value of a TV based video telephony experience by comparing it to a PC/webcam experience and a face-to-face communications experience [1]. Equipment was set up to enable bi-directional audio/video communications between pairs of participants. In the PC and TV conditions, the equipment was set up in separate rooms so that no face-to-face communication was possible between the subjects except via the system. In the face-to-face condition (F2F) the subject pairs were seated at a table, opposite one-another. Each subject pair carried out three activities in each of the three conditions (F2F, PC and TV), these being controlled for order effects. Though these activities were purpose-driven, the aim was to create a social rather than a work related interaction. In keeping with the social/domestic usage context of the system, these activities were a Taboo game, a Charades game and a Lego building task. These were chosen in order to engender a fun and relaxed ambience consistent with the expected context of use. A further aim was also to encourage both verbal and non-verbal (visual/gestural) communication between subjects.

In the Taboo game, one subject was given the name of a major US city and using verbal and non-verbal communications, had to get the second subject to guess the name. The first subject was also given a list of words which could not be used. Three different city names were used and allocation to the conditions was controlled. In the Charades activity, one subject was given the name of a movie and using only visual communication methods, had to get the second subject to guess the name of that movie. These were also chosen to be reasonably challenging such that some creativity in visual communications might be required.

The third activity was a Lego building task. In this, the first subject was given a set of Lego blocks (but no instructions), while the second had only a picture of the final assembly. The objective was for the second subject to provide instructions to the first subject to build the assembly, using verbal and/or visual forms of communication and without showing the picture to the other person. As the aim was to provide the context for a goal oriented interaction between participant pairs, the success or failure of each activity was not relevant as a quantitative measure and hence was not recorded. At the beginning of each study condition, subjects also spent two to three minutes in general discussion in order to get accustomed to the medium of communication being evaluated. Each of the other activities then took between 1 and 5 minutes each, for a total "in condition" time of around 10 to 20 minutes. The order of presentation of the three activities was also controlled

for order effects. For each of the three activities, each subject independently completed a pre and post-study questionnaire to quantify their expectations of the communication method prior to the experience and their value judgments immediately following. Some of the expectations rated by users were comfort with the communication medium, quality of communication, privacy and security etc. They also individually completed a system usability scale [SUS] to enable quantitative comparison of the three conditions. In the TV condition alone, a special “features” questionnaire was also tendered to identify the relative value of some potential future features. Following each set of three conditions the subject pair was brought together for a semi-structured discussion about their experiences.

3. SUBJECTS

In order to ensure comfortable interactions between subject pairs, and to emulate the social context being investigated, same-gender pairs were recruited together to ensure that they knew one-another. Eighteen pairs participated in the study, with six pairs in each of the three age groups. The age categories were not chosen with the intent of being equal in range, but to increase the likelihood of age related differences being found in the subjective data by sampling from three disparate segments of the general population.

4. RESULTS

The post-test questionnaires indicated a significant effect of condition for all aspects, with TV rated significantly higher than PC, and F2F significantly higher than PC. However, there was no significant difference between the ratings for F2F and TV, consistent with the data from the SUS. In other words, the TV condition was rated very close to face-to-face communication. The features questionnaire indicated the two most highly rated items to be related to control; control over the sharing of video and control over who could use the system. Many subjects stated that the TV interaction felt more personal with a greater sense of connection with the other person because it was easier to see facial expressions and to sense feelings and emotions. Also, the head to waist view enabled body language to play a greater role, allowed subjects to be more expressive and there was a much greater sense of eye contact. However the self-view image generated some negative responses, partly due to being presented as picture-in-picture. Some subjects often looked at themselves and many characterised it “weird” to see themselves talking and the expressions they made, causing some to pay undue attention to their posture, hair, weight etc. This was readily explained by the fact that, in normal face-to-face communication, you do not see yourself so they were not used to it and most described it as a distraction. Subjects accepted the value of the self-view for camera set up but most would want to turn it off during a video call.

5. USABILITY STUDY METHODOLOGY

Based on these findings, a prototype application was developed and two different user interfaces designed. A usability study was carried out to evaluate the ease of use, learn-ability and usefulness of features offered. 24 subjects (mean 33 yrs) participated and each completed a total of 19 task sequences using each of the prototypes. They rated each

prototype on the criteria of ease of use and the usefulness of the function demonstrated. The relative value of the features and functions was also rated in terms of the extent to which the presence/absence of the features might impact on a decision to purchase or acquire the system, one aim being to further test the subjective findings of the concept study. All ratings were done using six-point Likert scales. The tasks included making and receiving video calls, camera adjustments (digital pan, tilt, zoom, camera view presets, self-view on/off and camera auto-aim) and changing of video settings to make audio only calls. The task sequence within each prototype condition was fixed but the conditions were controlled for order effects. The study environment was set up to have the look and feel (including ambience and lighting) of a typical domestic sitting room and the subjects sat on a couch with a coffee table in front much as they would do at home. Prototype user interfaces were operated using a standard TV remote control at a viewing distance of about 10 feet.

6. DISCUSSION

These two user studies were undertaken at different stages of the project and used quite different methodologies because they sought to address different research questions. The concept study was conducted at a very early stage of the product concept development to identify the user value of the *concept* of TV-based video telephony and the social/contextual benefits and drawbacks relative to some existing communication methods. This was used to steer the system design in terms of features, functions and market positioning. Usability was not an element of this study, in fact at this time; the UI had not even been designed; the concept evaluation study actually guided the design team in the UI design. One key lesson learned was that we may have gained a better understanding of how subjects felt about the concept if we had used some kind of “emotional state” assessment tool. We chose not to use such a tool in the study since they are time consuming to tender and to analyse the data. There are tools such as AttrackDiff™ which address some of those issues [3]. Where similar study objectives exist, we would likely try out a tool such as this to assess the emotional experience of users. The more traditional usability study was carried out later in the development with two objectives; to confirm that the features and functions identified in the first study had been integrated in a way which was useful, and secondly to evaluate the usability of the interface design. In both cases, alternatives were evaluated to enable subjects to make comparative rather than absolute judgments.

7. REFERENCES

- [1] Wheatley, D.J., and Basapur, S., 2009. A Comparative Evaluation of TV Video Telephony with Webcam and Face to Face Communication. Proc. of EuroITV. Leuven, Belgium, pp 1-8, Jun 2009.
- [2] Harboe G., Massey N., Metcalf C., Wheatley D.J., & Romano G. “Perceptions of Value: The Uses of Social Television” Proc. of EuroITV 2007, Amsterdam, Netherlands, 24-25 May 2007
- [3] Hassenzahl, M. (2002). The effect of perceived hedonic quality on product appealingness. International Journal of Human-Computer Interaction, 13, 479-497.
<http://www.attrakdiff.de/en/Home>