Interrupting People at Home*

Martijn H. Vastenburg, David V. Keyson, Huib de Ridder
Faculty of Industrial Design Engineering
Delft University of Technology
Delft, The Netherlands
{m.h.vastenburg, d.keyson, h.deridder}@io.tudelft.nl

Abstract – *Several factors were examined towards considering what influences users' acceptability of alerting and informational messages in a field study conducted at ten homes. Through the simulation study message urgency was found to be a better predictor of acceptability than the degree of user engagement during ongoing activities. A model is proposed for including a range of additional factors expected to influence acceptability. The factors will be examined in forthcoming studies.*

Keywords: Interruptions, messages, pervasive computing, ubiquitous computing, home, user experience.

1 Introduction

Information presentation and access in the home environment is expected to change in the near future due to technology development. Future homes could be able to sense situations, react appropriately and inform the user. Such aware systems will need to have a mechanism to determine when and how to interrupt the user. This paper describes the results of an exploratory user study in which user experiences of interruptions are analyzed.

1.1 Related work

In the field of mixed initiative interfaces, Horvitz and Apacible [1] [2] have studied the interruptibility of people in an office environment. In a user study test subjects performed 5 one-hour sessions in their offices. The sessions were taped on video, and system events were captured. After the session, the subjects were asked to tag and assess the video. Based on this data, a Bayesian model was constructed that predicts the desirability of different types of interruptions, ranging from a background notification to an intrusive phone call. In the papers mentioned, Horvitz and Apacible focus on the cost of interruptions, based amongst others on the state of the user.

In work by Hudson et al. [3] a different approach was taken in examining the interruptibility of office workers. Four staff members were monitored for 14-22 working days. Audio and video recordings served as a source for 'simulated sensors', which registered, for example, the number of people in the room. Subjects were asked to rate their interruptibility approximately two times per hour. Subjects had to hold up fingers to indicate the rating. This way the disturbance caused by the alerts and responses was minimized. The subjects were asked to give an in-situ self-report after each alert ("beeper study"). Based on the resulting data, a decision tree was constructed, with which 90% of the unwanted interruptions could be avoided.

Published research in the field of interruptions is focused mainly on the task-oriented office environment; the current study focuses on the home environment. Also, the studies mentioned did not consider the impact of the content of the interruption messages on the perceived acceptability of the interruptions. In our study the impact of the message on the acceptability is analyzed.

1.2 Aware homes

Interruptions may also occur in the setting of aware homes. Aware systems need to know when and how to communicate messages to the inhabitants. Compared to the office environment, in the home other aspects of the environment might be important for the cost-benefit analysis of the interruption. For example, when demands on attention are low in an office environment, an interruption might be acceptable. In the home environment however the system probably has to respect the moments of relaxation, so other ratings are needed to compute the optimal timing.

As a first step towards developing smart aware systems and gaining an understanding of how people experience interruptions at home, an exploratory study on interruptions in the home environment was conducted. Attention is paid to user activities in relation to interruptibility, taking into account social and environmental aspects, and perceived message urgency.

* **0-7803-8566-7/04/$20.00 © 2004 IEEE.**

2 User Study

The goal of the user study described here was to gain insight into user and environment aspects relevant to the acceptability of interruptions in a living room setting. The knowledge gained was used to reflect on an initial model for predicting when interruptions could best be given in an aware home system.

Horvitz and Apacible [2] described a cost-benefit approach to interruptions. We use a similar mechanism (Figure 1) to view interruptions in a home setting. The acceptability is assumed to depend on the urgency of the message and on the engagement in activities at the moment of interruption. A more urgent message leads to a higher perceived benefit; a higher level of engagement leads to a higher cost of interruption. In order to get a better understanding of the mechanism, we asked users to rate their engagement in activities, the perceived urgency of the message, and the acceptability of the interruption.

Engagement can be defined in terms of attention focus, curiosity, and intrinsic interest [4]. In this study, 'user engagement' indicates the involvement of the user in her current activity. The *user engagement* is measured using subjective ratings of concentration, urgency of the activity, and the annoyance of interrupting the activity.

The *message urgency* is a subjective rating of the urgency of the interruption message, regardless of the current context.

The *acceptability,* for the purpose of the current study, was limited to a subjective rating of the willingness to be interrupted with a given message at a given point in time. The perceived level of intrusion was also taken into account.

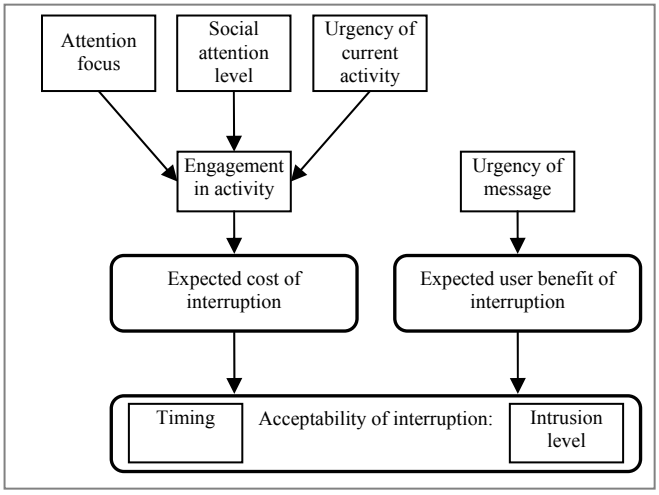

Figure 1. Limited model of the acceptability of interruptions

Hypothesis:

The cost-benefit ratio, based on the urgency of messages and the engagement of the user in her current activities, can be used as an indicator of acceptability of an interruption, as rated by users.

The two research questions in this study are: 1) Are users capable of distinguishing message urgency and their current level of engagement? 2) What is the relation between acceptability of interruptions, message urgency, and perceived user engagement? We expect a positive response to the first question. Acceptability is expected to be positively related to message urgency, but inversely related to user engagement (Figure 2). When people are highly engaged in their current activity, for example when a person is talking on the phone, we expect acceptability to be low. On the other hand, when the urgency of the message is high, a high acceptability is expected.

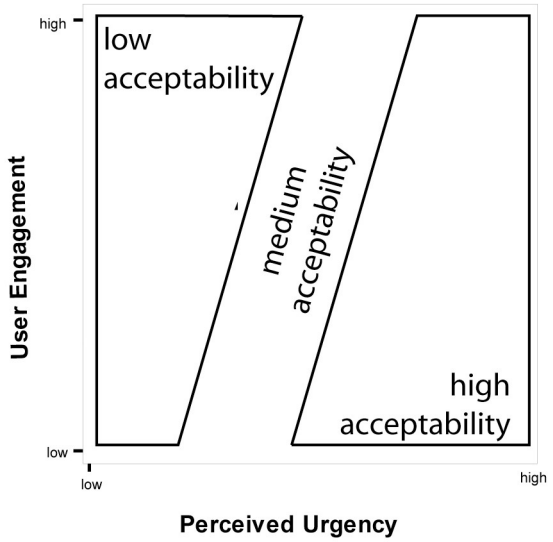

Figure 2: Expected acceptability of interruptions

3 Method

Test subjects participated at home. A laptop and a web cam were installed in the living room of the subject. The web cam was used to log the motion activity, and to capture the people present in the room at the time of interruption. A microphone was used to log the audio activity. The experimenter left the scene after placing the equipment and instructing the participants.

Participants were asked to do whatever they would do on a regular day, so the user engagement was not a controlled condition. Since the study took about 6 hours per session, a natural dispersion in user engagement was expected. Interruptions were given approximately two times per hour. When an interruption was activated, a sound was played, and the subject could read the message at the laptop. Subjects were asked to fill in a short questionnaire after each interruption. The study started

when the subject arrived home from work, and ended at bed-time.

3.1 Participants

10 subjects (4 women, 6 men) participated at home in the experiment. Ages ranged between 29 and 65, with a mean age of 45 and standard deviation of 13.9 years. Subjects were selected based on their home situation, being not living alone with no children at home.

3.2 Conditions

A set of 12 diverse interruption messages was created beforehand. In a prototype test the perceived urgency of these messages were observed to be diverse.

In the experiment, two sound signals were used to alert the user of a message:

- "B" (background): a non-intrusive soft bell
- "F" (foreground): an intrusive alarm bell

3.3 Schedule

A predefined schedule was used to activate messages. Activation time was either fixed or triggered by motion activity. The latter was used to ensure the users were present at the time of message presentation. A schedule consisted of 12 messages, with anticipated urgency levels varying from low to high. Six messages were linked to the background sound signal; six messages were linked to the foreground sound signal. The same set of 12 messages was used in all sessions. The order of the messages changed per session. If the participant was not present at the time of interruption, the questionnaire was skipped. Because of this, 20 out of 120 scheduled interruptions could not be used, leaving a total of 100 completed questionnaires.

Through a series of initial interviews with users matching the target group, messages were selected on the basis of plausibility, such that users could relate to the messages in terms of their living situation.

3.4 Questionnaire

The questionnaire consisted of three parts:

1. User engagement: people were asked to describe and rate their activity at the time of interruption. They had to give ratings on a -2 to +2 scale for:
   Q1) I was concentrated
   Q2) I was doing something urgent
   Q3) I was doing something important
   Q4) The interruption was disturbing from an emotional perspective
   Q5) The interruption was disturbing from a practical perspective

2. Message urgency:
   Q6) The message was important
   Q7) The message was urgent

3. Acceptability of the interruption:*
   Q8) The notification had to be shown immediately (vs. the notification had to be postponed)
   Q9) I did want to see the notification (vs. the notification had to be skipped)
   Q10) The sound signal for the message had to be more in the foreground (vs. more in the background)

*Note: though not stated, subjects were instructed to consider current activity when answering this question.

The questions were related to the model depicted in Figure 1 as follows: Q1 refers to Attention Focus, Q2-Q3 refer to Urgency of the Current Activity, and Q4-Q5 generally refer to Social Attention Level. Q6-Q7 refer to Urgency of Message. Q8 refers to Timing, Q9-Q10 refer to Intrusion Level.

4 Results

Table 1. Sample of data collected during one session. Messages I1-I12 are described in Figure 3

time	18:27	18:50	19:13	19:36	20:03	20:18	20:49	21:24	21:35	21:57	22:17	22:35
activity	unpacking groceries	discussing day	watching tv and talking	watching tv	watching news on tv	reading		working on computer	working	working	working	reading mail
Q1	1	1	1	2	2	2		2	2	2	2	1
Q2	-1	1	0	0	0	2		1	0	0	0	-1
Q3	0	2	1	1	0	0		1	0	1	0	-1
Q4	0	2	0	2	1	1		2	1	1	1	0
Q5	0	0	0	0	-1	-1		1	-1	0	1	-1
Q6	2	1	0	0	1	-2		1	2	2	2	1
Q7	-1	-1	0	0	-1	-2		1	2	2	2	-1
Q8	0	1	2	1	2	0		1	2	2	2	2
Q9	2	2	2	0	-2	-2		1	2	2	2	-2
Q10	0	0	2	-1	-2	-2		1	2	2	2	-2
message	I11	I9	I3	I6	I12	I10	I8	I5	I2	I4	I1	I7
signal	B	F	B	F	F	F	B	B	F	B	B	F

Table 1 shows as an example a condensed view of the data collected during one session. The session started approximately 18:00, and finished approximately 22:45. Twelve interruptions were scheduled. For each interruption, the user wrote down her current activity, and rated the 10 items on the questionnaire. At 20:49 no measurement was taken because the user was not present at the time of message presentation. Video and audio activity data was also collected during the course of the sessions for later analysis.

4.1 Message urgency and user engagement

In order to determine whether message urgency and user engagement could be considered independent factors, and to what degree they may influence acceptability, a factor analysis with principal components using SPSS [5] was conducted on the results of the questionnaire. Results of a Varimax rotation on the questionnaire data and the resulting factors are depicted in Table 2, with all values <.10 suppressed. The first component explained 37% of the variation, the second component 30%. A third component explained only 7% of the variance. Hence, the variance in the data can be attributed mainly to two components.

Table 2. Rotated Component Matrix

Goal variable	Question	Component 1	Component 2
User Engagement	Q1		,642
	Q2		,859
	Q3		,845
	Q4		,808
	Q5		,690
Message Urgency	Q6	,908	
	Q7	,912	
Acceptability of the Interruption	Q8	,888	
	Q9	,768	,124
	Q10	,798	

Questions Q6-Q10 have strong loadings on component 1 whereas Q1-Q5 correlate strongly with component 2. Component 2 can be referred to as *user engagement*, since Q1-Q5 all related to the user engagement at the time of interruption. The interpretation of component 1 is less clear. Questions Q6-Q7 have the highest loadings on this component, suggesting message urgency as the most appropriate labeling. Questions Q8-Q10 were intended to measure acceptability, but also imply a sense of urgency, namely urgency of the notification. Therefore, component 1 will be referred to as *perceived urgency*.

The component matrix indicates a strong positive correlation between message urgency and acceptability, which is consistent with our expectations. We expected a negative correlation between user engagement and acceptability; however, the results from the factor analysis do not support this claim. One could conclude the best way to predict acceptability of interruptions, in terms of message urgency and user engagement, would be to consider only the urgency of the messages. We will have a closer look at this in the analysis below.

4.2 Perceived urgency

The subjective ratings on the messages varied in perceived urgency. The vertical axis in Figure 3 shows the interruption messages ordered by the means of the perceived urgency per message. Please note the messages were originally in Dutch. The horizontal axis shows the perceived urgency. There was no significant effect found of the sound signal (B/F) on perceived urgency.

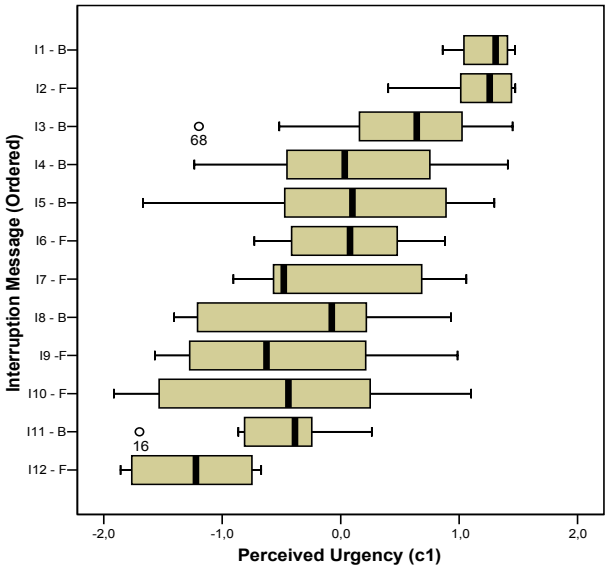

Figure 3. Perceived urgency of the interruption messages

An overview of the interruption messages:

I1 - The fire alarm in the shed has detected smoke
I2 - Alert: Chemical leak - please close all windows
I3 - The mailman is at the front door
I4 - Your appointment tomorrow 08:30 has been cancelled
I5 - The bath is ready
I6 - Coffee is ready
I7 - The front door has not been locked properly
I8 - Plants need water
I9 - Toilet paper has run out - shop closes in 10 minutes
I10 - The movie you selected starts in 5 minutes
I11 - To save energy, please lower the thermostat
I12 - Weather forecast: tomorrow will be rainy

For the messages at both extremes of the perceived urgency axis, all participants were consistent showing little variability. In the middle range there were subjective differences. However, the general ordering of messages across subjects was consistent in terms of perceived urgency.

4.3 Acceptability

Figure 4 shows the acceptability of interruptions based on the components that resulted from the factor analysis. Each item in the graph is a single interruption of a single subject. The interruptions are labeled by acceptability level as rated in Q8, reduced to 3 levels. In this analysis we used Q8 as the single best indicator of acceptability.

As can be seen in Figure 4, the 100 data points of judged interruptions are distributed evenly over the 2-dimensional space, suggesting user engagement had no influence on the perceived message urgency.

The three symbols in Figure 4 indicate the different levels of acceptability. High urgent perceived messages tend to be highly acceptable, following medium urgent perceived messages, and lastly low urgent perceived messages. No relation between perceived user engagement and acceptability was found.

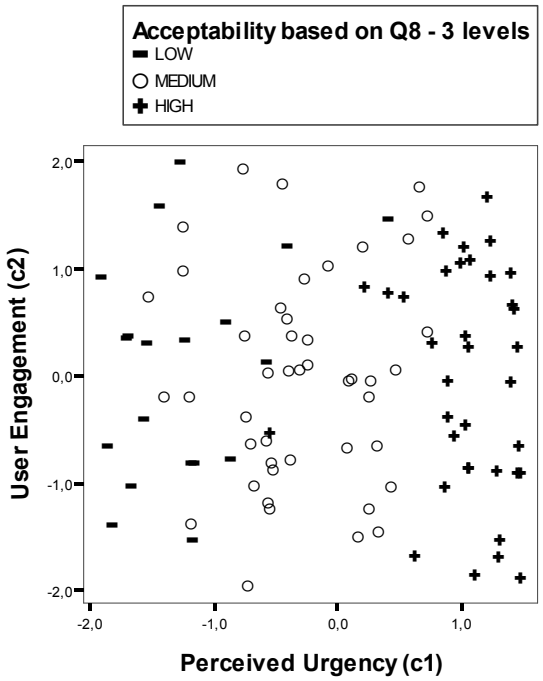

Figure 4. Scatter plot of data items

4.4 Alert sound level

Similar to the acceptability of messages according to urgency, the preferred sound signal level also positively correlated with Message Urgency. This holds for both the background condition and the foreground condition.

4.5 Engagement and activities

In the questionnaire, subjects were also asked to type in their current activity. Table 3 shows a selection of the activities mentioned, clustered by perceived engagement level.

As can be seen in Table 3, the watching TV activity ranges from low to high on perceived engagement scale. Apparently, in the current setting, the physical activity alone is not a sufficient predictor of user engagement, given that the underlying goals behind an activity may contribute to the perceived engagement.

Table 3. Activities vs. User Engagement

User Engagement	Activities
HIGH	watching TV news, watching favorite TV show, gluing photo on a card, eating
	editing text on computer, serving dinner, walking to the bathroom, discussing day, on the phone, sending an SMS message
MEDIUM	watching TV, cooking, reading, working, eating
	watching TV, cooking, chatting, cleaning, watching birds, drinking tea, reading mail
LOW	drinking wine, computer stuff, entering the room, reading newspaper, relaxing, nothing
	watching TV, relaxing, washing the dishes

5 Next steps

The lack of a relation found between perceived urgency and user engagement in terms of interruption acceptability, could be attributed to a number of factors, including:

- While users were verbally informed to consider context when answering questions on acceptability, in the questionnaire this was not explicitly communicated to users. This may have resulted in a low-level interpretation of acceptability, i.e. subjects might have judged the message urgency and presentation rather than taking into account their own current engagement and activities.
- Interruptions were only given in the vicinity of the messaging system. Therefore, the range of activities in which the user could have been engaged in at the moment of interruption was by definition limited. This may have reduced the influence of user engagement on message acceptability.
- The assessment was confined to one evening. Users might be highly engaged with the message system itself because of its novelty.

While the measurement of user engagement was limited in the current study, it might be interesting to consider user acceptance of a smart messaging system whereby message presentations are managed on the basis

of urgency alone, without attempting to measure and account for user engagement. Ideally, such a study should be conducted in a realistic setting with real messages over a longer period of time. Given a system which utilizes the level of message urgency to manage, one could consider using different displays for messages depending upon the classified level of urgency. A system could display all low urgency messages via a non-intrusive interface in the background, for example a display next to a door. The high urgency messages would have to be communicated via an attention-demanding alert. The medium urgency messages could then be classified by an intelligent system in order to select the best interface and intrusion level.

Acknowledgements

The authors thank Caroline Hummels and the other colleagues from the ID-StudioLab for their advice and support.

References

[1] Eric Horvitz, "Principles of Mixed Initiative User Interfaces", Proceedings of CHI 1999, City, pp. 1-3, 1999.

[2] Eric Horvitz, Johnson Apacible, "Learning and Reasoning about Interruption", ICMI 2003.

[3] Scott E. Hudson, James Fogarty, Christopher G. Atkeson, Daniel Avrahami, Jodi Forlizzi, Sara Kiesler, Johnny C. Lee, Jie Yang, "Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study", Proceedings of the conference on Human factors in computing systems, Ft. Lauderdale, pp. 257 – 264, April 2003.

[4] Peter Chapman, Sanjeebhan Selvarajah, Jane Webster, "Engagement in Multimedia Training Systems", International Conference on System Sciences, 1999.

[5] SPSS 12.01, © SPSS Inc, www.spss.com.