ELSEVIER

Perceived usefulness and performance of human-to-human communications on television

Hokyoung Ryu *, Aaron Wong

Institute of Information and Mathematical Sciences, Massey University, New Zealand

Available online 12 September 2007

Abstract

A key assumption of future television (TV) environments is that the future TV viewing experiences will be more active and interactive. Currently several TV technologies based on networked computing, e.g., IPTV (Internet Protocol TV) or Mobile TV, have made it possible for people to interact with their TVs, or even with other viewers through their TVs, by allowing them to access additional functions, for example actively participating in a quiz show, instantly sharing other viewers' opinions and sending or receiving emails while they are watching a TV programme.

To ensure uptake of these new TV technologies, it is essential to match the performance of the novel systems to both current TV viewing experiences and future user needs, since the characteristics of the additional tasks that TV viewers will perform will lead to other substantially different TV viewing experiences. This paper reports on the usefulness and performance evaluation of a novel TV-based human-to-human interaction environment, where audiences of a specific TV channel can exchange public and private text messages. It combines the two popular analogies (Internet chat and SMS) and embeds them into the viewers' TV experience.

Our results showed that this TV hosted human-to-human communication environment could integrate well with the current TV viewing experience. This might lead to applications such as TV-based personal messengers and/or live bulletin board community for fans of a specific TV show. They also revealed several issues that need to be addressed in the development of new TV technologies.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Television; Evaluation; Texting; SMS; Verbal conversation; Perceived usefulness; Performance

* Corresponding author. Tel.: +64 9 4140800x9140; fax: +64 94418181.
  E-mail address: h.ryu@massey.ac.nz (H. Ryu).

1. Introduction

Of the many activities carried out at home, viewing TV is arguably the most common experience that might be considered as informative, enjoyable, and relaxing. It frequently seems to be the default activity that one engages in when one does not have anything better to do. Such experiences influence our behaviour. For instance, TV seems to encourage people to interact with their families and/or their friends, either talking about TV programmes they have viewed, or retrieving conversational motives from them to explain their own attitudes or social actions (Steuer & Hustedt, 2002). Of course, TV can also have the opposite effect whereby people watch TV and do not interact with each other. These behaviours have long been recognised as individualistic aspects of the general TV viewing experience (Kubey, 1986; Kubey & Csikszentmihalyi, 1990).

The new TV technologies, e.g., Web-based TV, Interactive TV (ITV), and Internet Protocol TV (IPTV), are leading to a convergence of the traditional TV environment and the Internet. New services based on these technologies will have a major impact on the current TV viewing experience. This leads us to question how much of current attitudes and behaviour will persist in the new TV environments, and how much the newly developing technologies can extend the current TV viewing experience. We expand on these ideas in the following section.

1.1. The TV viewing experience

Gauntlett and Hill (1999) illustrate the current TV viewing experience with the following two excerpts:

> Television is the biggest source of information in my life. I treat television as my window to the world (15-year-old schoolboy, Gauntlett & Hill, 1999, p. 114).
> With my husband, we talk about all the programmes we watch – some only by a passing comment, but many more deeply... With our family, when we phone each other – we often mention programmes we have enjoyed to each other. Also if we spot anything forthcoming of interest we point it out (58-year-old female school teaching assistant, Gauntlett & Hill, 1999, p. 128).

The first quote from a 15 year-old schoolboy illustrates a use of TV as a way of enhancing the viewer's individual awareness or knowledge of the world. On the other hand, the second passage represents a social use context. Clearly, the female school teacher is very willing to consolidate intimate relationships with her social contacts, i.e., family members, friends, or partners, through communication triggered by her past or present TV viewing experience. Lull (1980, 1982, 1990) defines this TV viewing experience as a way of establishing common grounds among viewers for their *relational uses* or *social uses* of TV, emphasising that viewers employ their TV viewing experience to facilitate communication, or at least, as a means to open up conversation. Himmelweit, Oppenheim, and Vince (1958) also stressed the social use of TV, identifying that those interviewees who had richer social ties would be more attracted to the type of content that they could share later. Although technology, behaviour, environment and the family unit have changed since then, the social use of TV seems to be still prevalent (Robinson, Kestnbaum, Neustadtl, & Alvarez, 2000).

Other studies have presented the TV viewing experience as an effective human-to-human communication tool. For instance, Kubey and Csikszentmihalyi's study (1990) revealed that TV viewers are talking with their family members at home, or their friends over telephones, for about 60% of their TV viewing time. Furthermore, to some extent, the content of this communication is heavily reliant on what they have seen or heard on TV. The researchers claimed that the current TV viewing experience mainly includes getting pleasure from the TV content they are watching (i.e., individual use), but also makes them engage in their social interaction situations (i.e., social use). Following on from this understanding, Gray (1992) observed, based on several ethnographical studies, that a very important part of the pleasure of TV programmes is to "gossip about them the following day" (p. 14). Morley (1992) and Palmer (1986) also concluded that the communication activity which is derived from the current TV viewing experience is a critical part of our everyday lives. In effect, one of the key aspects of the current TV viewing experience is thus to be able to communicate and share media experiences between members of a social network.

An important note should be made here. Most early studies on the TV viewing experience, apart from Kubey and Csikszentmihalyi's study (1990), have only considered the nature of the communication activity as being retrospective. In other words, the communication activity happens later, after watching the TV. However, more recent surveys on media uses (e.g., Miller, 2005) show that SMS (Short Message Service) texting and other user interactions in response to TV programmes have gained a lot of popularity recently as part of the viewing experience. For example, some TV channels in New Zealand use SMS chat which involves sending text messages to a phone number, so that the messages are shown on the TV a short while later. Using SMS, however, leads to difficulty-performance trade-offs. That is, a viewer's visual modality must be shared to both process the information from the TV and compose text messages on mobile phones, so that one inevitably misses some of the TV content. Texting is more difficult than simply watching TV with subtitles, because one needs to look back and forth between the TV and the mobile phone. Nevertheless, this does not seem to be a major constraint, since using SMS has been a commercial success with many popular TV programmes, such as "Big Brother". In this context, a practical question of our research was to explore whether this text-based communication activity might be integrated with new TV technologies such as Web-based TV and Internet-based TV.

This research question, however, (text-based communication on TV) might be regarded as trivial, even though many commercial applications, e.g., WebTV™ or AOLTV™, have seen the necessity of text-based communication as a key driver of future TV technologies. In the current TV viewing experience, no one is suggesting that a person should be viewing TV to communicate with their friends or family, or even to consolidate their social ties with them. Loosely speaking, people watch TV to satisfy their enjoyment first and, if motivated, only share their experiences later. Choosing between text-based and verbal communication relies heavily on the availability of media and contexts of use. However, a recent survey on TV viewers' general attitudes (Knowledge Network survey, cited in the Madison Avenue Journal December 6, 2004) showed that many TV audiences are now searching for more ways to control their TV viewing experience and make a place for TV as part of their increasingly hectic lifestyles, implying a roadmap for the future TV environment. The survey also reported that the proportion of the viewers using SMS has dramatically increased. Therefore, gaining an understanding of the benefits and limitations of text-based commu-

nication in the current TV viewing experience seems to be important and timely as it has implications for the development of future TV technologies. This paper does not intend to comprehensively investigate all the possible communication activities, since this is rather too broad a scope. Instead, we narrow down this issue to explore the benefits and limitations of text-based communication, comparing it with the most common communication activity, i.e., verbal conversations on phones, in the current TV viewing experience. This will give an insight into how effectively the future TV environment may cope with human-to-human communication, extending our social interaction in the TV viewing experience.

1.2. Potential usability issues of communication activity in TV viewing

The verbal conversations described above seem to be possible because the current TV viewing experience is a very low-demand or low-involvement activity. Many psychological studies on the TV viewing experience (e.g., Armstrong & Greenberg, 1990; Armstrong & Sopory, 1997; Salame & Baddeley, 1989) showed that the two different modalities (i.e., auditory modality for phone calls and visual for viewing TV) would not conflict substantially, so the performance of viewing TV would not deteriorate much. However, this is not the case for text-based communication activity, since the viewer's visual modality would be shared between processing the information from the TV and composing or reading text messages on a mobile phone. This leads to the first question of this paper, regarding how difficult text-based communication really is.

One of the early Internet-based TV environments has made it possible for viewers to do text-based communication more easily. See Fig. 1. WebTV™ (recently re-branded as MSNTV™) first developed this concept through a high-speed home network, so that

Fig. 1. Chatting on WebTV™, *Courtesy from msntv.com.*

viewers can text-chat with their friends or relatives on their TV sets. This text-based human-to-human communication via TVs is being widely used in other new TV technologies, e.g., AOLTV™. However, text-chatting on WebTV™ or AOLTV™ does not allow people to continue to view the TV programme, leading them to change mode from viewing to communicating. Also, an early usability study performed by Nielsen (1997), even though it was based on an early version of WebTV™, showed that using the text-chatting facility was not a great user experience. Indeed, WebTV™ is still perceived by most people as being another form of PC rather than a TV, in the sense that to use it properly, users should sit closer to the TV set (less than one metre is normal) with a wireless keyboard and mouse. The current TV viewing experience is basically "lean back" interaction, which means that one does not want to watch TV as if using a PC. Any communication activity via the TV should consequently address this issue. Another practical concern of this paper was therefore to evaluate this text-based communication in a realistic television viewing situation being distinct from PC use, in order to suggest some design guidelines for future Internet-based TVs.

As a new way of supporting human-to-human communication and reducing the usability problems noted above, the authors have worked on a concept to send text messages between TVs, i.e., TxtTV (see more details in Section 2). Indeed, this idea is not new. As shown in Fig. 2, a Korean company recently introduced an application, synchronising text messages between mobile phones and TVs. For example, a mother could send a text message to both a child's mobile phone and TV set, because she may expect that her child is watching TV when the text message is sent. Further, it allows the recipient (the child in this example) to send a text message back to the sender via the TV set. While the convenience of communication is evident, it is less certain that the concept is socially acceptable within the current TV viewing experience. Only large scale usability testing can show its feasibility, which is the main concern of this paper.

Fig. 2. Message TV: Text messages from TV to mobile phone or from TV to TV. *Courtesy from TelecomKorea.com.*
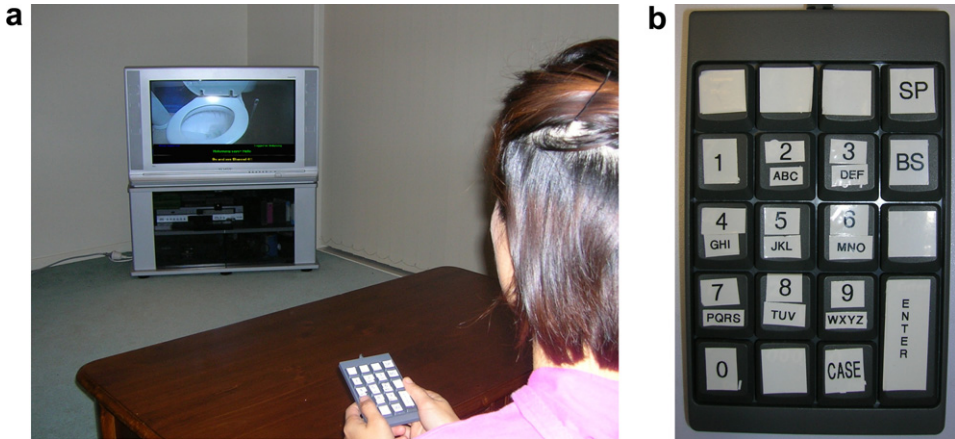
Fig. 3. A Prototype of TxtTV (a) and the remote control for TxtTV (b).

2. Texting on TV: TxtTV

None of the new TV environments reviewed previously is able to support both TV viewing and human-to-human communication at the same time. The concept of TxtTV is based on two analogies. First, a future Internet-based TV environment may be used to bring people together, enabling them to make use of shared reference points while they are watching the same TV programme (AOL patents, cited in the New York Times June 28, 2004). It may help people to constantly stay in contact with their friends or family over a period of TV viewing time. Second, the concept owes much to the vast success of SMS on mobile phones. Of course, its success cannot be taken for granted in other domains, but several sociological accounts propose potential applications of text messages for Internet-based TV. For example, Ling (2004), Licoppe and Heurtin (2001) and Ryu (2005) maintained that asynchronous text-based communication (i.e., SMS) is of value to enhance a broader and ongoing relationship between senders and recipients rather than highly attentive phone calls or chatting on PCs, so that subsequent communication is more easily reciprocated in their attitude. Further, it seems to be the most favoured communication style in relatively casual and informal situations that are closely equivalent to the current TV viewing experience.

To implement TxtTV, the WebTV™ and AOLTV™ environments were considered first. However, since it was not possible to develop new code in these platforms, an open-source interactive Java™ TV emulator – xleTView[1] was used for the development of an experimental apparatus. To provide a realistic TV viewing setting, two 25-inch LCD-TV units, as shown in Fig. 3a, were set up in two separate rooms. Also, the remote control (Fig. 3b) for texting was developed, with similar dimensions to the general remote controls used in traditional TV sets and the same layout with common mobile phones. In fact, while a cordless keyboard is adopted in most Internet-based TV environments the remote control was used here, since the experiments described below considered a naturalistic TV viewing

---

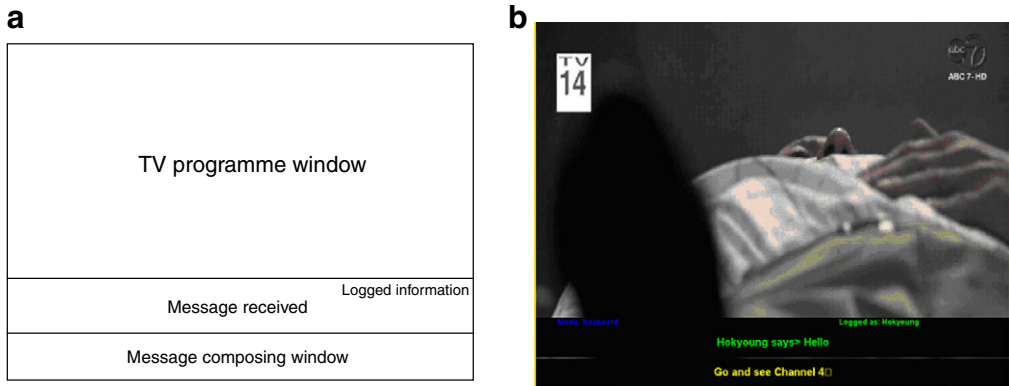[1] See more details at xletview.sourceforge.net.

Fig. 4. Components of TxtTV (a), and screenshot of TxtTV (b).

situation that is distinct from using a PC. The keypad layout on the remote control was similar to that of common mobile phones, so that one could compare SMS use on mobile phones with TxtTV. A pre-recorded TV programme was synchronised between the two TV sets through both telephone-based home network systems and a server for controlling the TV sets, ensuring that participants always saw the same TV contents.

To exchange text messages on both TV sets, a simple mechanism was implemented over TCP/IP protocols. Once a text message from one TV set was ready to pass to the other TV, it was passed through the server based on the Ethernet switching system.

Fig. 4a shows the screen layout of TxtTV. There are two message windows, i.e., the window for text messages received, and the window for composing text messages. In addition, there is logged information on the right, which shows the people who have logged on to the Ethernet switching system. It is very similar to the "buddy list" in most chatting software. For editing text messages, TxtTV considered the multi-tap method. That is, the user presses each key one or more times to specify the input letter, e.g., the '2' key is pressed once for the letter 'a', twice for 'b', and three times for 'c', which is currently the most common text input method for mobile phones. 'Times New Roman', as a font, is used, and the font size of the messages is 20 points for easier reading at a distance of 2 m from the TV set.

The intention of this research was to explore whether text-based communication activity on TV could be effectively associated with the current TV viewing experience without any significant difficulty-performance trade-offs. If this is the case, then we seek to answer how a new TV environment can best serve this communication activity in terms of relevant usability issues.

3. The evaluation (experiment 1)

To do so it was necessary to examine two aspects of text-based communication activity in a naturalistic TV viewing situation: *perceived usefulness* and *performance*. The former represents the kinds of behaviour expected of a user within the activity, whereas the latter provides analytic accounts by focusing on the routine behaviour of doing the activity (Simon, 1988). These two aspects allow one to identify how a new technology achieves

critical user experience in the near future, and eventually how the new TV technology may be accepted. To explore the perceived usefulness of TxtTV, after exposing TxtTV or the other control conditions to participants, we obtained the participants' perception of each system. This approach has been well validated through many usability studies for predicting user acceptance and explaining the use of newly developing systems (Dadayan & Ferro, 2005; Davis, 1989; Mathieson, 1991). Performance was assessed using a wide variety of measures, depending on the contexts of the experimental task. We decided an appropriate measure for not disturbing primary task performance was the NASA-TLX scale (Keinonen, 1998). More details are described in each section below.

The evaluation described in this section was suggested by one of the ideas from Section 1, that the text-based communication activity in the current TV viewing experience might be useful to capture most TV viewers' interests in future TV environments. In this sense, texting on TV (TxtTV) was predicted to provide more enjoyable viewing, with less distraction from the TV programme viewed.

The data was collected in a relatively naturalistic TV viewing context, with three experimental settings: "SMS use in viewing TV", "Telephone use in viewing TV", and "TxtTV". The first setting was a control condition to enable us to interpret the size of effects in the other two settings. The experiment involved exposing participants to one of the experimental settings, controlled for TV contents and viewing time, and then asking them to give seven ratings on four questions. The experimental task involved pairs of participants, who were asked to communicate with their experimental partners while watching TV. The participants who used telephones were asked to hang up or call whenever they wanted to. Likewise, the users of SMS and TxtTV were told to send short text messages if they wished to do so. In this way we were able to examine how much each experimental setting was perceived to be useful and usable, simulating a naturalistic TV viewing circumstance.

The seven ratings scales assessed different aspects of the communication activity. Fig. 5 depicts the questions used in this experiment. The first three questions assessed how much participants thought that the communication activity they did while viewing TV enhanced their TV viewing experience. The other question collected their feeling towards the communication activity.

The purpose of this evaluation was to determine under what circumstances this sample of participants found the communication activity useful. To answer this question the ratings were analysed in terms of their absolute level, i.e., did the mean rating indicate agreement or disagreement. Indeed, a more plausible measure to address the question, e.g., how many times the participants communicated, was not considered in this experiment, since it was planned to examine in Experiment 2. Analysis of variance was used to compare the size of the effects of the independent variables.

1.　The system added enjoyment and entertainment to my TV viewing 2.　The system made me feel closer to her or him 3.　The communication with him or her made me miss some TV contents 4.　The system was very comfortable and easy to use

Fig. 5. Questions for identifying perceived system use.

3.1. Method

3.1.1. Participants

Fifty-six pairs (102 participants) were assigned at random to one of the three experimental setting (''SMS use in viewing TV'', ''Telephone use in viewing TV'', and ''TxtTV''). The participants were recruited from Massey University students, and 41 pairs knew their partner prior to the experiment, whilst 15 pairs did not. In particular, 17 out of 41 pairs defined themselves as close (or best) friends. None of them had any experience of using Internet-based TV. Table 1 summarises the experimental settings and the allocation of the participant pairs. All the participants had a higher level of English proficiency in order to ensure their understanding of the TV contents and their communication activity in English. For all the statistical analyses presented for all the experiments, the sampling unit was the participant pair.

3.1.2. Experimental design

The experimental design was a 3 by 3 between-subjects factorial. The medium used for the communication activity was the main between-subjects independent variable (SMS, Telephone, and TxtTV), and the social relationship of pairs (close, acquaintance, or never met before) was a between-subject nuisance independent variable. In particular, SMS was a control condition. The nine experimental groups were formed by factorially combining the possibilities, given in Table 1. The dependent variables used the seven ratings discussed above.

3.1.3. Apparatus

A news programme was pre-recorded (see Fig. 3a), and two Philips™ 25-inch LCD-TVs were used. Also, two PCs and a server were used to run the codes for TxtTV, and the pre-recorded TV programme was viewed for about 15 min on both TV sets. The telephones (Panasonic™ KXTG5632M) and mobile phones (Samsung™ MT200) were located with the participants in the separate rooms. Each participant sat 200 cm from the TV set. Two Olympus™ S723 portable recorders were used to record the telephone conversations for subsequent analysis. Also, all the text messages exchanged on the mobile phones and TxtTVs were collected. Figs. 3 and 4 give an impression of what the apparatus looked like.

3.1.4. Procedure

Each pair of participants was first provided with the instructions regarding the experiment. These gave information about the experiment, the purpose of the study, and the data protection policy. Participants were told that this was an experiment examining how people could communicate over a TV set. To ensure that both participants were

Table 1
The nine conditions and number of participant pairs in Experiment 1

	SMS + TV	Telephone + TV	TxtTV	Total
Close friend	5	6	6	17
Acquaintance in the class	7	7	10	24
Never met before	5	5	5	15
Total	17	18	21	56

comfortable with the given experimental conditions a pre-trial session was undertaken and they were given the option of withdrawing from the experiment.

The participants were told that the task was to communicate with their partner while watching TV, and their partner had the same view as they did. Also, they were told that they could send text messages, call or hang up the phone whenever they wanted to. After each member of the pair had entered their respective rooms, the experimenters simultaneously turned on the two TV sets. As the pair finished viewing the TV programme, they were asked to rate each of the four questions, as shown in Fig. 5.

3.2. Results (Perceived system use)

Results are presented separately for each rating scale. Two questions were asked of these data: (i) under what conditions did the mean rating indicate general agreement, i.e., when was the mean rating above the neutral rating of 4.00, and (ii) how did the two independent variables, communication medium (Telephone, SMS, and TxtTV), and the level of intimacy (close, acquaintance, and never met), affect the ratings. The former question was answered by inspecting the means. T-tests were carried out to assess whether each mean was significantly higher or lower than the neutral value 4.00. The latter question was answered by carrying out a two-way between-subjects analysis of variance. Levine's test for heterogeneity of variance was used to assess whether the assumptions of the analysis of variance were met.

3.2.1. Scale 1: "The system added enjoyment and entertainment to my TV viewing"

Scale 1 was to ask whether they had enjoyed the TV viewing experience with the communication activity, with an expectation that they would give the highest ratings to TxtTV, as it presents a novel approach. Table 2 gives mean ratings for the three experimental settings across the three personal relationships. Looking at the figures in the total row and column in Table 2, it can be seen that there is an effect of communication medium and intimacy. TxtTV (mean 4.26) gives the highest rating rather then the others, and the close friends (mean 4.29) seemed to gain more enjoyment. A mean rating above 4.00 indicates that on average people agree rather than disagree with a scale.

A two-tailed T-test assessed whether the mean ratings were significantly different from 4.00 (neutral value), revealing that none of the means are significantly different from 4.00 except for the mean ratings for TxtTV. This was further analysed by a two-way between subjects analysis of variance, giving main effects for medium and intimacy ($F_{2,47} = 3.88$, $p < .05$; $F_{2,47} = 6.29$, $p < .01$), but no higher order interaction effects. As noted above,

Table 2
Mean ratings for Scale 1, 'The system added enjoyment and entertainment to my TV viewing'

	SMS + TV	Telephone + TV	TxtTV	Total
Close friend	4.60 (1.51)	4.08 (1.62)	4.25 (1.42)	4.29 (1.50)
Acquaintance in the class	2.79 (1.42)	3.71 (1.27)	4.50 (1.54)	3.77 (1.57)
Never met before	2.40 (1.26)	2.90 (1.52)	3.80 (1.23)	3.03 (1.43)
Total	3.61 (1.50)	3.21 (1.65)	4.26 (1.43)	3.73 (1.57)

A rating 1 indicates the participant strongly disagrees with the statement, a rating of 7 that they strongly agree (mean/s.d).

ratings were significantly higher for TxtTV (mean 4.26) than the other conditions (mean 3.21 for 'SMS + TV' and 3.61 for 'Telephone + TV'). A Tukey test (at $p \leqslant .05$) also supported these observations.

It is striking that TxtTV is favoured over both SMS use and verbal conversations. A possible explanation for these results would be the novelty effect of TxtTV. Another interpretation would be that the difficulty of using SMS might be correlated with these ratings. Also, the lower ratings for the phone conversations while viewing television imply that our participants were more attentively conversing with their partner rather than both getting pleasure from the TV contents and being engaged in their social interaction. Further accounts of these findings are considered in the discussion.

3.2.2. Scale 2: "The system made me feel closer to her or him"

Scale 2 tested whether communication activities in the TV viewing experience enhanced the intimacy of the relationship between partners. Our prediction was all the media would be likely to have the same effect in that this question only reflected their general perception of communicating while viewing TV. Indeed, this scale seems to be somewhat inappropriate, since it may be too subjective and often affected by other factors not considered in this evaluation, e.g., the pairs who are in love or the actual messages exchanged. However, it might tell us how much the participants saw the benefits of the communication activities in the current TV viewing experience. Table 3 depicts these means. All of them were greater than 4.00 (two-tailed *T*-tests), which might reasonably be said to indicate that the communication activity *per se* made people believe that they were getting more intimate communication with their partner.

A two-way between subject analysis of variance was not appropriate for these data as Levene's test for heterogeneity of variance was found to be significant ($F_{8,47} = 3.49$, $p < .01$). A Kruskal–Wallis test was performed to compare the 21 pairs using TxtTV against the other pairs, i.e., collapsing across intimacy. This was not significant (K–W = 1.38, n.s). The same test applied to the intimacy also indicated a non-significant difference (K–W = 5.58, n.s). These results support the general benefits of human-to-human communication in the current TV viewing experience irrespective of both their relationship and the medium used. Also, the novelty effect of TxtTV does not seem to be sensitive in this scale.

3.2.3. Scale 3: "The communication with him or her made me miss some TV contents"

Scale 3 was intended to assess general negative feelings about the communication activity in the current TV viewing. Table 4 depicts these mean ratings. Most of the ratings were

Table 3
Mean ratings for Scale 2, 'The system made me feel closer to her or him'

	Telephone + TV	SMS + TV	TxtTV	Total
Close friend	6.00 (0.74)	6.10 (0.74)	5.83 (0.58)	5.97 (0.67)
Acquaintance in the class	4.93 (1.77)	5.64 (1.60)	5.20 (1.58)	5.25 (1.63)
Never met before	5.10 (0.74)	4.50 (1.78)	6.10 (0.74)	5.23 (1.33)
Total	5.33 (1.31)	5.44 (1.56)	5.60 (1.23)	5.46 (1.36)

A rating 1 indicates the participant strongly disagrees with the statement, a rating of 7 that they strongly agree (mean/s.d).

Table 4
Mean ratings for Scale 3, 'The communication with him or her made me miss some TV contents'

	SMS + TV	Telephone + TV	TxtTV	Total
Close friend	5.90 (2.13)	6.42 (0.90)	4.58 (1.78)	5.62 (1.79)
Acquaintance in the class	6.50 (0.65)	6.14 (1.09)	4.50 (1.47)	5.56 (1.47)
Never met before	6.70 (0.48)	6.50 (0.85)	4.00 (1.21)	5.73 (1.55)
Total	6.38 (1.26)	6.33 (0.96)	4.40 (1.52)	5.63 (1.58)

A rating 1 indicates the participant strongly disagrees with the statement, a rating of 7 that they strongly agree (mean/s.d).

very high, i.e., above rating 4.00, indicating general agreement with this statement. Only the TxtTV condition consistently gives a low rating around 4.00 indicating neutrality with this statement (two-tailed T-tests).

Again, a two-way between subject analysis of variance was not appropriate for these data as Levene's test for heterogeneity of variance was found to be significant ($F_{8,47} = 3.49$, $p < .05$). A Kruskal–Wallis test was performed. Despite these much higher ratings, the non-parametric test shows TxtTV is perceived as a significantly less distracting from the TV contents (K–W = 12.41, $p < .01$). Yet, there is no intimacy effect on Scale 3 (K–W = 4.22, n.s) either.

It is very striking that TxtTV is believed to be a less interfering medium for performing the given task, compared to the telephone conversation. It is generally thought that using the telephone is a common medium that people normally use in their current TV viewing experience (Kubey & Csikszentmihalyi, 1990; Steuer & Hustedt, 2002). A possible view of these results would be that the telephone conversation requires more attention, so that participants might feel that its use is more disruptive to TV viewing. On the other hand, the participants with TxtTV were less sensitive to the communication activity, probably due to its asynchronous nature of communicating with a partner. However, this explanation cannot be applied to SMS use, because it seems to be the most challenging activity, which took their attention from the TV contents. More explanations of these findings will be discussed later.

3.2.4. Scale 4: "The system was very comfortable and easy to use"

Scale 4 was to directly test how comfortable the system was to use. The ratings for this question were uniformly low, indicating that none of the systems resulted in means ratings indicating agreement with this statement (see Table 5). The system coming nearest to a

Table 5
Mean ratings for Scale 4, 'The system was very comfortable and easy to use', for the three media of communication in the three relationship of the participants (mean/s.d)

	SMS + TV	Telephone + TV	TxtTV	Total
Close friend	2.10 (1.29)	3.00 (1.71)	3.17 (1.64)	2.79 (1.59)
Acquaintance in the class	2.79 (1.37)	2.64 (1.95)	3.35 (1.35)	2.98 (1.55)
Never met before	1.70 (1.25)	1.90 (1.66)	3.40 (1.35)	2.33 (1.58)
Total	2.26 (1.36)	2.56 (1.80)	3.31 (1.41)	2.75 (1.58)

mean rating 4.00 was TxtTV (mean 3.31), but a *T*-test indicated a non-significant difference.

Despite the much lower absolute ratings, the pattern of results is similar to that of Scale 3. A two-way between-subjects analysis of variance gave main effects for medium ($F_{2,47} = 5.11$, $p < .01$), but not for intimacy ($F_{2,47} = 1.40$, n.s). As these results parallel those for Scale 3, the lower levels of agreement appear to have reduced the sensitivity of the test to a degree where the effect of intimacy is no longer significant. Even though TxtTV was generally better than the others, our participants perceived the communication activity itself as being less comfortable. These findings present something of a paradox. In both Scale 1 and 2, our participants clearly saw the benefits of the communication activity along with TV viewing; however, scale 4 response showed that they still felt that the activity was uncomfortable or difficult to use. One possible explanation of these results may be that the communication activity as the secondary task gives rise to negative attitudes toward the primary task (TV viewing), but it may be acceptable in terms of usefulness (Scale 1 and 2). This might be interpreted to mean that critical user experience with future TV technologies seems to go beyond usability. Another possible explanation is discussed later.

3.3. Discussion and conclusions

The main outcome to be drawn from these results is that certain levels of social communication activity can be effectively integrated with the current TV viewing experience. That means that highly attentive social communications like telephone conversations would not be well matched with the watching TV. This seems to contradict the early studies, e.g., Kubey and Csikszentmihalyi (1990), which identified telephone conversations as an effective social interaction tool regarding TV. However, verbal conversation was the normal way to communicate at the time of that study, so it is not appropriate to compare it directly with the results from this paper. This therefore implies a practical contribution of this experiment. Further, it can be seen that text-chatting on TV, as adopted in MSNTV™ or AOLTV™, might be less useful, since it appears to be a highly attentive and synchronous communication activity.

The pattern of results was broadly in line with the predictions made. In particular, TxtTV gains more favour from our participants. There are a number of possible explanations for these results. One is the novelty of the technology. The participants who used TxtTV might have been attracted by the new technology, and enjoyed it more as a result. Yet, this novelty effect of TxtTV was insensitive in the other questions, i.e., Scale 2, 3, and 4. Also, our participants saw that TxtTV could be a less distracting way of communicating while viewing TV, in that it was easier to use and less distracting than a telephone conversation. Indeed, it is true that SMS use in the current TV viewing experience would be the most challenging activity, but it is very striking that the telephone conversation is also perceived as challenging. This account seems to be contrary to what the multiple resource theory (e.g., Allport, Antonis, & Reynolds, 1972; Kantowitz & Knight, 1976) predicts. According to this theory, our telephone users can use two modalities to separately process two types of information (auditory modality for telephone conversations and visual modality to mainly process images from TV). Of course, while the auditory modality should process the verbal information from both the telephone conversation and the TV programme, the interference should be minimised over that of SMS and TxtTV, which

mainly employ visual modality for both text messages and TV contents. The experiment described in this section cannot explicitly tell us how these results should be interpreted. It is only possible through more controlled experiments, and is beyond the scope of this paper. Another explanation could be the experimental task. In this experiment, our participants were allowed to use telephones at any time, which intended to simulate a realistic TV viewing situation. Paradoxically, this naturalistic TV viewing situation made many of them (10 pairs out of 18 pairs) hold on the phone for the whole experimental session. This meant that they were more engaged in the communication activity (secondary task) than the TV viewing activity (primary task). This may account for the lower ratings for Scales 3 and 4. To clarify these results, the evaluation described below used a different experimental task.

Another empirical contribution of this research is that our participants appear to use the TV contents as a motive for their communication. Actually these data were not described above, because the telephone conversations were very poorly recorded, so that a further analysis of them was not possible. However, both the SMS data (16 out of 17 pairs) and messages on TxtTV (20 out of 21 pairs) clearly show that many users are using the TV contents as their communication motive, which is very much in line with the early studies on TV uses (Gray, 1992; Himmelweit et al., 1958; Morley, 1992; Palmer, 1986).

...
Participant A: *YEAH HEY I WAN NA C THIS*
Participant B: *HECK YES DA VINCI CODE*
Participant A: *A WHILE AGO COUPLE WKS..*
Participant B: *DA A LOT OF BKS*
Participant A: *THIS IS SUCH FUSS OVA NOTHING*
Participant B: *LOL*
Participant A: *YEA WELL ITS GONA B DAT WAY ALL OVA A WORLD*
Participant B: *I ACTULY WATCHD DIS DA FILM REVIEWR DOESN'T LIKE THIS MOVIE*
...

Fig. 6. Part of the verbal protocol from one of the pairs who used SMS while watching TV.

...
Participant A: *WHAT SHOW R U LOOKING NOW*
Participant B: *tv1 news*
Participant A: *SAME*
Participant B: *TALKING ABOUT DAVINCI CORD*
Participant A: *I SAW THE DAVINCI CORD*
Participant B: *IS IT NICE*
Participant A: *DID U SAW THIS MOVIE*
Participant B: *NO*
Participant A: *I THINK IS GOOD*
Participant B: *THE TICKET IS TOO EXPENSIVE*
...

Fig. 7. Part of the verbal protocol from one of the pairs who used TxtTV.

Fig. 6 contains part of the verbal protocol from a particular pair who used SMS while watching TV, and an excerpt from TxtTV users is depicted in Fig. 7.

Even if this evaluation seems to be sufficient to guide a potential application for adoption in the future TV environment, there would be difficulties in interpreting results such as Scale 4. The following experiment directly assessed this issue, removing the problems found in this experiment.

4. Experiment 2

While the previous evaluation identified the relative usefulness of texting on TV over conventional ways of communication, it was still difficult to tell if texting would achieve an essential level of usability in Internet-based TV environments. The evaluation in this section was designed to address this question, setting out three purposes.

First, we wished to examine the mental workloads for each communication activity used in the previous experimental settings. The contention is that a user who uses TxtTV could have a lower workload, based on the results of Scale 4 in the previous evaluation. To test this contention we replicate the same experimental settings but make use of the NASA TLX measure, not disturbing primary task performance (Hart & Staveland, 1988). Second, it is evident that to some extent the communication activity *per se* makes TV viewing hard. Scale 3 in the previous experiment indicated that TxtTV might be less distracting than the other two conditions. To reconfirm the assessment, we used ten questions about the TV programme participants had viewed, so that we could measure how much the communication distracted their TV viewing. Finally, the level of engagement with the communication activity while viewing TV was analysed. In this test, the experimental task forced the participants to comprehend the TV contents, but allowed them to communicate with their partner. Therefore, one can see how this sample of participants recognise the communication activity in such a limited purpose, i.e., either whether they only concentrated on the TV programme; or whether they would like to communicate with their partner about the TV contents, or simply do their communication activity only. To do this we were interested in measuring (i) how many participants engaged in the communication activity, (ii) whether messages exchanged were relevant to the TV contents, and (iii) how many messages were exchanged. These measures can partially support Scales 1 and 2 of the previous evaluation. An important note should be made here. Telephone use seems to be the most convenient rather than SMS or TxtTV. This means that the simple analysis of the absolute figures may be of little interest to understand human behaviour. Instead, a relative figure, i.e., the number of messages relevant to the TV contents out of the total messages exchanged, was investigated. Also, the number of participants who were engaged in the communication activity was only used to give an insight into participants' attitudes toward the communication activity.

This experiment differed from the first one in three ways. First, the measures are not the same, which paid more attention to usability rather than usefulness. A second difference was the experimental task. In the previous experiment, the participants were not forced to watch the TV programme. However, in this experiment, they were told that they would have a quiz of the TV contents. For the quiz questions, we used a news programme with five episodes. The last distinction was that there is no consideration of intimacy between participants, since we had already identified that the effect of social ties or relationships was minimal to determine the benefits of the communication activity in the current TV

viewing experience. By manipulating these three aspects in this experiment, we hoped to increase our understanding of the issue in TV viewing.

4.1. Method

4.1.1. Participants, apparatus and procedure

One hundred participants were used in 50 pairs and assigned at random to one of the three experimental settings (17 pairs in "SMS use in viewing TV", 12 pairs in "Telephone use in viewing TV", and 21 pairs in "TxtTV"). As described above, the same apparatus with Experiment 1, except a different TV programme (a talk show), was used. Also, the same procedure as the previous experiment was followed, except they were told that they should comprehend the TV contents, but they were allowed to communicate with their partner. As they finished watching the TV programme, the NASA-TLX scale was measured using software developed by the Naval Research Laboratory[2], and then the quiz questions were administered.

4.1.2. Experimental design

The experimental design was a one-way between-subjects factorial. The medium used for the communication activity was the between-subjects independent variable (SMS, Telephone, and TxtTV). The dependent variables were the number of messages relevant to the TV contents out of the total messages exchanged, quiz performance, and NASA TLX scales. For all the statistical analyses presented in all the experiment, the sampling unit was the participant pair.

4.2. Results (Performance)

4.2.1. Number of messages relevant to the TV contents

Table 6 represents the data obtained from this evaluation. The dependent variable of primary interest is the proportion of messages relevant to the TV contents out of the total messages exchanged.

Looking at the first row, which represents the number of pairs who communicated with their partner more than once, it was clear that all the participants were engaged in the communication activity. This gives an insight into one side of human behaviour in the current TV viewing experience, even though the primary task was to comprehend the TV contents. This was not statistically analysed due to the ceiling effect.

Comparing the figures in the second row reveals there seems to have a considerable difference, i.e., text-based communication (SMS and TxtTV) tends to be more oriented to the TV contents they had watched. This was analysed by a one-way analysis of variance, revealing that they were significantly different. A Tukey test showed that the telephone condition was significantly lower than both SMS and TxtTV uses, which were not significantly different from each other. A possible explanation for these results would be cost-benefit trade-offs. It is true that the effort to use TxtTV and SMS is much higher than for verbal conversation, so that the participants sought to maximise the use of text-based communication activity to support their primary task, comprehension of the TV contents.

---

[2] See more details at nrl.navy.mil.

Table 6
Communication performance to the three experimental settings (mean/s.d)

	SMS + TV ($n = 17$)	Telephone + TV ($n = 12$)	TxtTV ($n = 21$)	$F_{2,47}$	p
Number of pairs who communicate more than once	17	12	21	–	n.s.
Messages relevant to the TV contents/total messages exchanged (per cent)	70.23 (15.62)	45.24 (11.54)	65.75 (17.29)	4.21	<.05

F is the statistic computed to compare the three experimental conditions. p gives the statistical significance of this test.

4.2.2. Comprehension of TV contents

Table 7 gives the mean performance of the 10 questions about the TV programme that the fifty pairs had viewed. Comparing the figures reveals that the verbal conversation seems to be slightly higher than the other experimental conditions, but a one-way analysis of variance test showed this difference was not significant.

In conjunction with the previous results, even though many pairs were engaged in the communication activity, their comprehension of the TV programme was not significantly hampered by the activity. This empirical data indicate that the task – watching TV – is not a highly cognitive task, so it allows other activities to take place at the same time, which is very much in line with the early studies (e.g., Gauntlett & Hill, 1999; Kubey & Csikszentmihalyi, 1990). A practical contribution of this paper empirically supports the early studies. It also suggests that future TV technologies might consider communication activities as usable functions to associate with the current TV viewing experience.

4.2.3. Workloads

NASA TLX is a multi-dimensional rating scale that derives an overall workload score based on a weighted average of ratings on six subscales (Hart & Staveland, 1988): *mental demand, physical demand, temporal demand, effort, performance*, and *frustration* level. It can be generally used for an estimate of the attentional demands of a task, considering the demand a task imposes on a user's resources (Proctor & van Zandt, 1994). Workloads could thus be conceived of as the inverse of usability (Fairclough, 1991). Table 8 summarises the data, together with the mean scores for the six workload items.

Overall, TxtTV seems to have the least workload and SMS use while viewing TV has the highest. A Tukey test showed that TxtTV was distinct from the others, but that they were not different from each other. In each workload item, a one-way between-subjects analysis of variance was applied, indicating that there are significant differences in terms of mental, temporal demands, and frustration. A Tukey test showed that, in terms of mental demand, TxtTV was significantly lower than the other two conditions, which were not

Table 7
Quiz performance to the three experimental settings

	SMS + TV	Telephone + TV	TxtTV	$F_{2,47}$	p
Performance (max: 10)	7.41 (1.83)	7.77 (1.34)	6.86 (1.57)	2.76	n.s.

F is the statistic computed to compare the three experimental conditions. p gives the statistical significance of this test.

Table 8
Mean ratings of workloads for each measure

	SMS + TV	Telephone + TV	TxtTV	$F_{2,47}$	p
NASA TLX (Total, max:100)	64.46 (15.64)	57.89 (13.55)	51.18 (13.85)	8.00	<.01
Mental demand	68.82 (16.00)	70.00 (19.89)	53.57 (20.43)	8.47	<.01
Physical demand	43.24 (21.84)	35.63 (17.90)	45.71 (22.51)	1.75	n.s.
Temporal demand	67.94 (21.78)	53.96 (20.16)	50.48 (17.87)	7.73	<.01
Performance	59.71 (20.48)	58.33 (18.81)	56.19 (20.74)	1.04	n.s.
Effort	59.41 (23.48)	52.71 (20.38)	53.10 (13.92)	1.28	n.s.
Frustration	62.21 (25.50)	47.92 (21.96)	48.69 (18.84)	4.40	<.05

F is the statistic computed to compare the three experimental conditions. p gives the statistical significance of this test.

different from each other. Both the temporal demand and frustration were also analysed by Tukey tests, revealing that SMS use was significantly worse than the others, but that the others had no difference between them.

5. Discussion and design implications

Newly developing TV technologies, e.g., Interactive TV, MSNTV™, AOLTV™, and IPTV, have been drawing attention to the connection between the Internet and the traditional TV environment, as a primary direction of the future TV environment (Miller, 2005). Most of them stressed the convergence as a fundamental need from the perspective of human-to-human interaction, allowing us to be able to chat or email while viewing TV. However, the paradox of the convergence is that it may increase opportunities to create and maintain human-to-human interaction (social use of TV), but tends to overlook other TV use, i.e., individual use of TV. That means that communication activities on TV demand a set of engagements at various levels that differ substantially from the current TV viewing experience. This is the main motivation for this research. Accordingly, we developed and evaluated a prototype called TxtTV.

5.1. Lessons learnt

The apparatus devised for this paper provided a naturalistic TV viewing situation, but in the context of the different communication activities. Comparisons of the three configurations in this paper allow us to assess the potential value of the communication activities available in this experimental context.

The first evaluation in this paper suggested texting on TV could enhance human-to-human communication, without excessively disturbing the current TV viewing experience. This was mirrored in Scales 1 and 3 of Experiment 1. This extended TV viewing experience can let TV viewers interact with one another or with contents associated with the TV programme they have viewed by sending text messages that are displayed on the TV screen. Scale 2 could be seen in this light. A real challenge here is to meet user's needs and usability, which was reflected in Scale 4. In terms of usability, the second experiment quantitatively confirmed the significant advantages of TxtTV over the other conventional communication mechanisms. Neither of these possibilities had been demonstrated empirically before, which implies a practical contribution of this paper.

Of course, other samples in other contexts may give different results and the generality of these results can only be tested by further studies, which are now being planned. For instance, people would probably prefer to use the phone to communicate, rather than sending messages to their friends via TxtTVs, when they do not know who may be watching and reading their messages. Therefore, perhaps, users in some situations (such as when conveying private and personal feelings) would be very unlikely to use TxtTV. Furthermore, the sample size in these two experiments might not reliably ensure the interpretations made in this paper. However, taken at face value the results are encouraging. They suggest that as text-based communication on TV becomes more familiar, and television users learn to use the communication tools, the benefits will be of much interest in the light of the widespread use of SMS on mobile phones. That is, as texting on TV has entered popular consciousness and use, it may follow the same track as its development on mobile phones. One can view the measures taken by Scales 1 and 2 in Experiment 1 in this light. Finally, design and testing the right attributes of newly developing systems means leaving the laboratory and considering the system and its functions in real world environments. To do so we extended Monk, Carroll, Parker, and Blythe (2004) to review the usefulness and performance of TxtTV in a naturalistic situation. The experimental method (combination of perceived usefulness and performance) demonstrated here provides a way of testing the new inventions and exploring the hypotheses set out above. It is to be hoped that, by applying the method to other conditions and contexts, researchers can provide the insights needed to make the user of the new technology more useful and usable.

5.2. Using results for a future TV design

The results of this paper imply that social interaction seems to be a key factor that influences critical user experiences with new TV technologies (Roibas & Johnson, 2006). The assumption underlying this development is that TxtTV can be made more usable by paying careful attention to both the social use of TV and the current TV viewing experience. In this light, texting on TV would have a place in future development. However, current developments in Internet-based TV, e.g., MSNTV™ and AOLTV™, seem to focus on a synchronous communication method similar to that seen on PCs, i.e., text-chatting. A tentative answer for this approach would be negative because it intrinsically asks us to pay too much attention to the communication activity at the expense of missing the TV contents, as in the verbal conversations in the two experiments described above. Of course, to confirm this answer, many other studies that could be performed to increase our understanding of human behaviour are undoubtedly needed, and these are planned for the near future.

References

Allport, D., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology, 24*, 255–265.

Armstrong, G. B., & Greenberg, B. S. (1990). Background television as an inhibitor of cognitive processing. *Human Communication Research, 16*, 355–386.

Armstrong, G. B., & Sopory, P. (1997). Effects of background television on phonological and visuo-spatial working memory. *Communication Research, 24*, 459–480.

Dadayan, L., & Ferro, E. (2005). When technology meets the mind: A comparative study of the technology acceptance model. In *Electronic Government, Proceedings* (Vol. 3591, pp. 137–144).

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Fairclough, S. (1991). *Adapting for the TLX to measure driver mental workload (1017)*. Loughborough, UK: HUSAT Research Institute, University of Loughborough.

Gauntlett, D., & Hill, A. (1999). *TV living: Television, culture and everyday life*. London, UK: Routledge.

Gray, A. (1992). *Video playtime: The gendering of a leisure technology*. London, UK: Routledge.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). Amsterdam, The Netherlands: Elsevier.

Himmelweit, H. T., Oppenheim, A. N., & Vince, P. (1958). *Television and the child: An empirical study of the effect of television on the young*. London, UK: Oxford University Press.

Kantowitz, B. H., & Knight, J. L. (1976). Testing tapping time-sharing: I. Auditory secondary task. *Acta Psychologica, 40*, 343–362.

Keinonen, T. (1998). *One-dimensional usability*. Helsinki, Finland: Helsinki University.

Kubey, R. (1986). Television use in everyday life: Coping with unstructured time. *The Journal of Communication, 36*(3), 108–123.

Kubey, R., & Csikszentmihalyi, M. (1990). *Television and the quality of life: How viewing shapes everyday experience*. Hillsdale, NJ: Erlbaum.

Licoppe, C., & Heurtin, J. P. (2001). Managing one's availability to telephone communication through mobile phone; a French case study of the development dynamics of mobile phone use. *Personal and Ubiquitous Computing, 5*, 99–108.

Ling, R. (2004). *The mobile connection: The cell phone's impact on society*. Morgan Kaufmann Publishers.

Lull, J. (1980). The social uses of television. *Human Communication Research, 6*(3), 197–209.

Lull, J. (1982). How families select television programmes: A mass observational study. *Journal of Broadcasting and Electronic Media, 26*(4), 801–811.

Lull, J. (1990). *Inside family viewing: Ethnographic research on television audiences*. London, UK: Routledge.

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behaviour. *Information Systems Research, 2*(3), 173–191.

Miller, S. (2005). Taking on the masses with mobile messaging TV. *ACM Computers in Entertainment, 3*(2), 1–7.

Monk, A. F., Carroll, J., Parker, S., & Blythe, M. (2004). Why are mobile phones annoying? *Behaviour and Information Technology, 23*, 33–41.

Morley, D. (1992). *Television, Audiences and Cultural Studies*. London, UK: Routledge.

Nielsen, J. (1997). WebTV usability review. Retrieved 20, July, 2006, from the World Wide Web: <http://www.useit.com/alertbox/9702a.html>.

Palmer, P. (1986). *The lively audience: A study of children around the TV set*. Sydney, Australia: Allen and Unwin.

Proctor, R. W., & van Zandt, T. (1994). *Human factors in simple and complex systems*. Boston, MA: Allyn and Bacon.

Robinson, J. P., Kestnbaum, M., Neustadtl, A., & Alvarez, A. (2000). Mass media use and social life among internet users. *Social Science Computer Review, 18*(4), 490–501.

Roibas, A. C., & Johnson, S. (2006). Unfolding the user experience in new scenarios of pervasive interactive TV. Paper presented at the conference on human factors in computer systems, Montreal, Canada.

Ryu, H. (2005). Multi-modal messaging services: Design factors of text-based communication via mobile phones. Paper presented at the HCI (Human–Computer Interaction) international, Las Vegas.

Salame, P., & Baddeley, A. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology, 41A*(1), 107–122.

Simon, T. (1988). Analysing the scope of cognitive models in human–computer interaction: A trade-off approach. In D. M. Jones & R. Winder (Eds.), *People and computer IV* (pp. 79–93). Cambridge, UK: Cambridge University Press.

Steuer, F. B., & Hustedt, J. T. (2002). *TV or no TV? A primer on the psychology of television*. Oxford, UK: University Press.