Comparing Social TV approaches
A look into enjoyment, presence and awareness
Jarno Zwaaneveld
j.a.a.zwaaneveld@student.utwente.nl

ABSTRACT
Social TV is a rapidly growing area of research. Many people are already using Instant Messaging (IM) programs while they watch TV [1], chat options have been added to online video sites and soon there may be communication options for Media Centers or TV set-top boxes. This paper describes a research that looks into the different approaches in Social TV experiments currently being conducted with users. TV based and computer based approaches will be compared. It looked at the factors of enjoyment, presence and awareness. There were no significant differences discovered in these factors and recommends that systems should offer both approaches as options.

Keywords
Social TV, enjoyment, presence, awareness, text chat, audio chat, privacy.

1. INTRODUCTION
TV is less and less used as a common experience for people to form a social connection through. People would discuss last-night’s sporting event or episode of their favorite TV show. However, with more and more people watching alone and asynchronous (due to DVRs, on-demand video services, households having multiple TVs and a trend towards smaller households ([2, 3])) TV is no longer the ‘electronic hearth’ it once was [4].

Social Television is a term for systems that support communication while watching TV or similar video content. It already exists in some people’s lives in ad-hoc forms. For example, people IM or text-message each other while they are watching TV or other video content while others already consume live streaming media that offers text chat (for example: justinTV (www.justin.tv) and uStream (www.ustream.tv)). These options could reinvigorate and dramatically expand the social aspects of TV viewing to people that are geographically separated.

Most Social TV projects aim to augment the TV experience to make the watchers feel as if they are watching together with the remote viewers as if they are in the same room [5]. Systems offer either text or audio chat and many also provide buddy-lists, avatars and more to create connectedness between the users.

This is an active area of research for which a lot of papers have been published in recent years. A lot of the research is going in different directions; some studies add basic communication to a TV system [6] and others add internet TV to real-time chats on a computer [7]. Studies into these two approaches (computers-based and TV-based) will have different aims and encounter different problems and solutions. The question is if users have different experiences with each approach and if they are used in different ways. What are the positive and negative aspects of each approach? In what direction should development be focused to improve the user experience?

1.1 Factors
To be able to compare the approaches mentioned we must choose factors to look at. Enjoyment, presence and awareness are common but important factors.

1.1.1 Enjoyment
Enjoyment is essential in any form of social software and indeed important in many types of software in general. The point of social software is to make users freely interact with each other. Its importance is demonstrated in the fact that if the users do not enjoy interacting through a system they will find other methods of interaction that do fulfill their needs [8].

A good definition of enjoyment is: “A is enjoying E, if E is causing A to have a number of occurrent beliefs concerning E, which collectively add significantly to the pleasure (happiness) A is experiencing” [9] Rephrased: you enjoy something when you think it adds to your pleasure or happiness.

1.1.2 Presence and Awareness
Presence & awareness are to closely related factors in a Social TV context. Both terms have many possible definitions in different contexts and there is no consensus on a single one. Definitions relevant for Social TV systems are described here.

In [10] awareness is defined as “the state of knowing about the environment in which you exist; about your surroundings, and what they are doing.” In [11] presence is defined as referring “to the feeling of being socially present with another person at a different or remote location.”

In [12] and [13] a belief is expressed that awareness can lead to connectedness (“a positive emotional appraisal which is characterized by a feeling of staying in touch within ongoing social relationships” [14]). Connectedness is a part of the emotional experience of social interaction[15]. The importance of connectedness is highlighted by Rettie’s findings that “the need for connectedness is the most important factor in making a choice between communication channels.” [16]
These factors have some overlap. Connectedness and the feeling of presence often occur together while both require awareness to be present. These relations are demonstrated in the following Venn diagram (based on [16]).

Figure 1. Relationship between Awareness, presence and connectedness

A possible example for a Social TV system would be a status indicator or a buddy list to make a user aware of another user’s presence and availability for social interaction[17]. This awareness and presence do not necessarily create connectedness, for example, a user would not feel connected to a stranger. An example of connectedness without presence is a text message.

This makes the two related factors an important subject to look into. They are base parts of a socially used system.

1.1.3 Measuring factors

The factors mentioned are subjective and measuring them directly is impossible. The papers discussed used several methods. Most used a questionnaire with a Likert scale. A Likert questionnaire asks how much a respondent agrees with a certain statement (strongly agree, agree, neither agree nor disagree, disagree, strongly disagree). From the responses, a numerical score can be created. The studies here used these scores to compare different implementations of their systems. Some of the studies asked users to rate their enjoyment on a 100-point scale, where 0 was “not at all” and 100 was “very much”. The results from these were used in a similar way to the Likert questionnaire.

A few of the papers used a process based on ‘grounded theory’ (GT). In GT, researchers generate a theory from collected data. From this data, multiple researchers individually marked and coded what they considered important data points. Researchers extracted these coded points from the texts and then grouped them into concepts. From the concepts, they formed categories, which were then the basis of a theory. Here, the collected data included things such as interviews, automated usage logs, and voice mail diaries.

1.2 Paper organization

The next chapter describes the research questions used in this paper. Next, I will explain the method used and the search for the papers used in this study. The following two sections will explain each of the paper, a chapter per approach, and a summary for each. In the last section, I will present conclusions followed by a closing section regarding future work.

2. RESEARCH QUESTIONS

- Are there differences in the experience users have when using systems with different Social TV approaches?

We will focus the research on the factors enjoyment, presence and awareness. The main question will be answered through the following sub-questions:

- What, if any, difference exists between Social TV approaches in terms of enjoyment?
- What, if any, difference exists between Social TV approaches in terms of presence and awareness?

3. METHOD

This paper is a literature study about Social TV, so first relevant papers must be found. Most of the search for papers was done through the ACM Digital Library with ‘social TV’ as a search term. The references used in these results were also looked up using Google Scholar. User studies discussing the factors of enjoyment, awareness and presence were selected from these results.

Five papers discussing a TV based approach were selected along with four discussing a computer-based study. A short description of the approaches used in these papers follows in the next chapter.

Next the results presented in these papers will be discussed compared, to see what differences exist. Then the research questions will be answered and conclusions will be drawn.

4. COMPUTER-BASED

4.1 How text and audio chat change the online video experience [18]

This paper presents an example of a lab experiment with a computer-based system. In the study, groups of friends were presented YouTube videos while being able to use either a web-based text chat, an audio chat (using voice-activation), both or neither. They selected highly rated videos for use in this study. Some groups of users watched the videos in the same order while others watched in a different order. They measured enjoyment of the video, the chat and overall enjoyment with a Likert questionnaire. They also measured distraction by asking them with another Likert questionnaire. They compared the amount and content of chat between audio and text chats.

The results did not show significant differences in enjoyment of videos between text, audio, both or no chat. Furthermore, the researchers found that enjoyment of chat was independent of enjoyment of video. However, they also found that distraction had little correlation with enjoyment of videos or overall.

Audio chat resulted in more shared laughter (double the amount of text chat). This suggests an increase in amusement. The paper also indicates that audio chat might lead to more
connectedness and presence than text chat due to audio chat users sharing laughter more directly.

They also recommend reserving audio chat for people that have already established relationships because of its higher degree of intimacy. For strangers a text chat affords a higher degree of selective self-presentation.

4.2 CollaboraTV: Making Television viewing social again [2]
This paper presents a field trial of the CollaboraTV system. During the four-week trial, users could access a buddy-list and a program-guide through their computers. The program-guide listed all the shows available through the system, highlighting those that have been watched by a user’s buddy. It also offers separate lists of popular shows and shows that buddies have watched.

CollaboraTV presented shows with a ‘virtual audience’. This virtual audience is displayed using simple avatars in a movie-theatre-like setting below the content.

While watching synchronously (at the same time as other users), chat content was displayed within speech bubbles. The system also recorded this chat content as temporally linked annotations. Users watching the same show later (asynchronously) will see the conversations unfold while they are watching as if they were watching at the same time and can add to them if they want. Users can also create two other types of annotations: interest points (indicating positive or negative reaction to the shows content) and expressions (expressed by their avatars). These annotations are only shared with users that are in their buddy lists due to privacy.

The researchers logged the interactions between users (for use in a GT process) and had the users fill in a Likert questionnaire.

In the results users said the system was fun to use. Over half agreed that, compared to traditional TV, the virtual audience was more enjoyable.

The paper also reported that users were very engaged when watching content with a virtual audience. It enriched the viewing experience and provided a sense of social presence. Some users explained that sharing programs that elicit strong emotions with friends or family brought them closer together, which indicates connectedness. Users did complain that there were not many people online and that the experience would be better with more people.

4.3 Watching Together: Integrating text chat with video [7]
This paper presents two studies of chatting behavior while watching video online. Researchers showed members of MovieLens (an online movie recommendation community) movies while being able to use an IRC-based chat feature. The members were then asked to fill in a Likert questionnaire and to comment on their experience.

They also conducted a more controlled lab experiment using a similar video and IRC-chat setup with specific cartoon content. They used groups of friends and groups of strangers to be able to measure the difference in impact of chat between them. They also studied the difference between groups with and without chat. They used a series of highly rated short cartoons as video content and used different structures. Of the chat group, half had short breaks between the videos while the other half had an extended time after all the videos had played. However, both groups could chat whenever they wanted.

The MovieLens members were very divided. They reported having fun and that having others helping them understand the movies increased their enjoyment of them. They were very positive about having others share their experience, which for some led to a sense of community and awareness of others, though others thought it a burden. A few users said they were not interested in chatting during a movie or that it was distracting and potentially annoying, though one participant said they might get used to "multitasking like that".

In the cartoon experiment, chat did not seem to have a significant impact on the enjoyment of the cartoons. The researchers attributed this lack of difference between chat and no chat in good cartoons to a possible ceiling effect (i.e. the cartoons were already fun, so they could not be rated higher). However, users rated poorly rated cartoons higher with chat than without it.

Users chatted freely about personal things unrelated to the cartoons. Sometimes they confessed to paying more attention to the chat than to the cartoons. It also elicited feelings of closeness and users felt awkward and unsure about what to say to strangers. This feeling of closeness indicated connectedness and presence.

This paper presents research into practices of college students using the internet to consume television content. These practices included viewing TV content while connected to remote users by way of IM.

The researcher conducted semi-structured interviews that were transcribed. They then analyzed these interviews using a GT process.

Participants reported occasionally consuming the same content while IM-ing, mostly for sporting events or linking video clips, sharing their excitement, which suggests enjoyment. They indicated they mostly watched what others were watching when in a communal setting, such as a shared living room. However, they regarded watching TV through the internet as a mostly solo experience, suggesting that they are not looking for connectedness when they consume TV content their computers.

4.5 Summary
In all papers, the users reported enjoying using the various Social TV systems. There did not seem to be a significant difference in enjoyment between the various implementations, apart from the increased amount of laughter in the audio chat system, which could indicate increased enjoyment over text chat. The data from the questionnaires did not reflect this though.

Three of the papers reported feelings of connectedness, community and presence. People were freely chatting as in normal conversation and felt close to their chatting partners. One of these papers did report a lack of people using the system.
The fourth paper, while reporting that people liked sharing content, indicated that users that regularly use internet-based TV-content regarded it as a solo experience, strongly contrasting with the social context in which users consumed regular TV content.

5. TV-BASED

In this section the first two papers present evaluations of two versions of related prototype Social TV systems connected to a TV screen. Users performed interactions with the system using a standard remote control. It uses pop-ups and a buddy list to make users aware that their friends and family are logged in and are watching TV. It allows predefined messages to be sent to buddies watching the same program. Along with the buddy list there was a color changing lamp (“orb”) set in the room that can convey the number of online buddies when the TV is not turned on.

5.1 Ambient Social TV: Drawing People into a Shared Experience[19]

For this study, the researchers collected data from semi-structured interviews, usage logs and voice mail diaries. They analyzed the data using a variation on the GT process.

In this paper, users reported liking the orb and its effect of stimulating their curiosity, drawing them into turning on their TV to communicate.

They added that presence lead to an expectation that someone was available for interaction. When people did not respond it was unclear if someone was having technical problems or just not replying. However, the peripheral awareness from the orb did lower the barrier to communication. The awareness that the orb conveys is peripheral because it can convey awareness while the user is doing something else, while the orb is literally in the peripheral vision of the user.

5.2 Examining Presence and Lightweight Messaging in a Social Television Experience[6]

The researchers collected data for this study from interviews, usage logs and voice mail diaries as well. They analyzed the data using GT.

In this study, users said they were disappointed to be unable to have a real conversation, saying the canned messages were too impersonal, sometimes working around this by calling each other on the phone (and sometimes just using the system to be able to tell when someone was home and then calling them).

The orb often prompted people go turn on TV and see who was on and what they were watching. Some users also reported they ended up viewing more programs that others were watching instead of just watching what they felt like. Many reported learning more about the viewing habits of other people in their group, even between siblings. One user did express that the orb did not offer an advantage over IM, because “through IM you do know when they are near the computer”.

Users reported that the system pulled them into the TV viewing experience, and feeling like others were around. However, they considered the limited messaging capability of the system a drawback, and often called each other as a workaround.

Participants in the study expressed a need for richer communication to be able to express themselves fully. The predefined messages prevented them from really watching TV “together”. This influenced a negative evaluation of the system, users claimed they liked to talk and would want to hear the other’s voice.

5.3 The Uses of Social Television [20]

In this paper, researchers first presented the results of an initial focus group interview. The users only had a moderate response to these interviews. Some users were enthusiastic while others said it would be distracting and interfering with their ‘down time’.

They also did a field trial with a prototype that had an open audio connection (instead of the predefined messages of the previous two studies). The data collection consisted of videotaping the users during the trial along with the programs they watched. Immediately afterward, the researchers also conducted interviews. They analyzed the data using a variation on GT.

During field trials with this prototype, there were much more enthusiastic responses. Users claimed it was “a blast” and were much more positive after actually using the system. They reported enjoying small talk during commercials or discussing the show during breaks and shared excitement over common interests and joking around. Researchers reported the social aspect of it gave pleasure to their participants. It relieved loneliness and enhanced the intensity of experiences.

Additionally users reported feeling like others were in the room and forgetting they were “talking through the television” which helped some to feel less alone.

Researchers also reported that the flow of the conversations was very natural. The participants noted the difference from a phone conversation in that they did not have to really divide their attention between two things. One problem with larger busy groups was that the audio being drowned out, making conversations almost impossible.

The users did note that they missed the physicality of actual presence, like a high-five after a sports team performs well.

5.4 Media Center Buddies: Instant Messaging around a Media Center [21]

The paper presents a study adding IM to a media center connected to a TV. Based on MSN Messenger (http://services.nl.msn.com/messenger/) users had access to their existing buddy list while using the media center functions. Multiple users could be logged in on the same media center. It also discusses the possibility of displaying the content a user is consuming in their buddy list.

The researcher asked users to fill in a questionnaire (with yes/no questions and 100-point scale questions) along with a more open-ended interview.

The study reported an increase in enjoyment when both participants were logged into IM (opposed to one participant logged in) while interruptions did not significantly interfere with enjoyment.
The systems IM features were not as enthusiastically accepted. Having others share the TV while receiving IM messages proved unfortunate in some cases. One participant had to patch up his relationship with his girlfriend after receiving an IM from an ex while another mentioned blocking some people on his buddy list prior to the study. Others reported wanting only a sub-set of their buddy list to be able to contact them while watching TV.

The reaction to sharing their current viewing choice was not much better. All male participants had reservations, they did not want their mothers/girlfriends know they were watching pornography nor their male friends to know they were watching home decorating guru Martha Stewart. Strangely, none of the female participants shared similar reservations.

5.5 Exploring Social TV [22]
This paper presents research using TVs in two rooms connected with an audio feed. Alternatively it offered a persistently displayed avatar whose look can be adjusted along with the possibility to send graphic symbols (such as emoticons) to the remote users. This to provide a sense of awareness and provide a non-verbal means of communication. It used users consuming the same content in the same room as a control group.

Users were asked to fill in questionnaires (using a Likert-like scale) and were interviewed.

Most users said they supported the idea of social TV in preliminary interviews. On average they said they would use the system for 3.5 hours a week (while the TV is on for almost 7 hours each day in the average American household[23]), with 7% claiming they would have the system turned on all the time while watching TV.

The study reported that face-to-face and audio connected users conversed for a very similar amount of time. User behavior in the face-to-face sessions suggested that visual gestures and facial cues were not very important when people are communicating while watching TV.

Additionally it was indicated that there is a comparable degree of “joint TV experience and social presence” between co-located TV sessions and remotely communicating using the audio feed while watching TV.

Users rated the presence and joint TV experience much lower for the graphic symbols.

5.6 Summary
In all studies, users reported at least interest in social TV. The field studies all reported users thoroughly enjoying use of the system. The only reported problems were with predefined messages not giving enough freedom of expression and a lack of privacy in one study.

The orb was very successful in drawing people into watching TV and communicating with their friends or family.

An open audio connection through a Social TV system afforded more natural communication than a phone conversation.

Watching TV remotely connected with an audio connection afforded a similar feeling of presence to watching TV while present in the same room.

The feeling of presence when using the systems relieved loneliness and helped lower the barrier for engaging in communication. Users often joined their friends that were already watching a program.

6. DISCUSSION
6.1 Enjoyment
There were no indications to suspect that there is a significant difference in enjoyment for TV or Computer based Social TV systems. All the different implementations reported that users enjoyed using the system, with any differences reported specific to that implementation.

The fact that both approaches offer the same basic functionality would explain this lack of difference. Their only difference is in the way the communication functionality is implemented.

It seems, however, that a TV-based system is probably less suited for text communication. Its users almost universally criticized the predefined messaging system, while implementations with a free text chat were successful.

One solution would be to just add a full keyboard to the TV-based systems but it is not clear if users would want this in their living room. Other possibilities to give users a way to freely enter text would be to add a text-messaging style of entering text by repeated button presses on the remote control, or a very small keyboard added to a remote control. One last possibility would be speech recognition.

6.2 Presence and awareness
In presence and awareness, there were a few differences. One of the computer-based systems had very few people online. This is possibly due to people not having a tendency to turn on the computer the minute they come home (as opposed to turning on the TV). TV’s are turned on almost 7 hours each day in the average American household[23], while individual computer usage is around 10 hours per week[24]. The likelihood of finding other people online increases of course with the length of time the system is turned on.

The researchers suggested an invitation mechanism to work around this but this would still only reach people that are using their computers at the time. An always-on solution like the Orb [6] would be a possible solution.

The TV based systems seemed to be aimed at connecting groups of people to other groups (as opposed to connecting individuals for the computer based systems). This leads to a number of possible issues.

Firstly, a lack of presence-information could result in sharing possibly private information with someone you did not know was present remotely. In some cases, just being in contact with a person could result in a conflict with a significant other. This presents a serious privacy issue. Possible solutions would include only using a sub-set of a person’s buddy list for Social TV purposes, or even having a separate TV-buddy list for the household. A system connecting two individuals or individuals to a group does not share this problem. One-to-one they are aware of whom they are talking to and one-to-group they are aware they are talking to a group and would likely be more careful in discussing private information.
In addition to this, for systems that share information on what the users are watching, users should be able to censor or at least turn off this capability. Systems that want to provide recommendations based on what others are watching should collect those statistics optionally and anonymously. Viewing habits are private information and protection of this privacy should always be a priority.

In terms of feeling the presence of others and being aware of their actions, apart from the previously mentioned issues, no difference was found between computer and TV-based approaches.

One last point raised by the difference in group-to-group versus the one-to-one implementations is that it seems to reinforce a point made by one of the reviewed articles. People use their computers as an individual. Even content usually watched on a TV is consumed alone when viewed on a computer. The studies mentioned in this paper all seem to follow this assumption implicitly.

A possible explanation for this is the place computers and TV’s have in a household. People commonly place TV’s centrally in the living room. Even when there are multiple TV’s in the house, at least some of the TV content would be communally consumed in the living room. Computers are often placed in a separate room, with one person at a time using it.

### 6.3 General

In terms of answering the question “what approach gives the best experience?”, I think the answer depends on the kind of household the system will be placed. In a home with multiple people and a living room, a TV based approach seems the most logical. Multiple users can use it at the same time and the living room is a social context.

For people living alone both approaches could work equally well. One consideration might be young singles and students. Living in student-housing or similar situations they might not own or have room for a television set. Here a computer-based approach would be ideal.

Taking both situations into account I think that a system that can do both would be ideal. Considering that some of the systems mentioned were implemented on Media Center PC’s there shouldn’t be any technological problems creating this possibility.

The scenario of a student living abroad watching TV on his laptop with his parents and siblings in their living room back home behind their TV is a simple illustration of the use of a system that can do both approaches and indeed the basic concept of Social TV.

### 7. CONCLUSION

I found no significant difference in enjoyment between approaches. Only implementation-specific differences were found.

I also found no true difference in awareness and presence. Both approaches can offer similar implementations and logically seem to afford similar experiences. The only difference noted was in how both users and researchers looked at the TV and computer mediums. The computer is viewed as something one person uses, while the TV can be shared.

Now, to answer the main research question: “Are there differences in the experience users have when using systems with different Social TV approaches?”

I think the answer is no. If implementations offer similar functionality then the experience will likely be similar as well, regardless of the approach used.

### 8. FUTURE WORK

One of the issues raised earlier was text input on a TV based system. The proposed solutions should be compared to each other in the context of a living room. This way an optimum solution could be found for text-based Social TV systems.

A gender difference in concern for privacy about viewing habits might be an interesting psychological study from which the results may or may not be relevant for Social TV.

A direct comparison between TV and computer based approaches might explain more about user preferences. i.e. while the approaches themselves might offer similar experiences users might still prefer to use one over the other. For example, a user might prefer a TV-based approach because he has his TV in the living room and feels it is a more relaxing and social context. This can only be concluded after a direct comparison.

Extended studies on how having a Social TV system in the home might change TV viewing and social interaction, could be worthwhile. Indeed, while I was working on this study Huang et. al [25] published another follow-up study on their Social TV system, looking deeper into how people integrate Social TV into their lives.

### ACKNOWLEDGMENTS

I would like to thank my supervisors H. Koppelman and Y. Cao and my fellow students in the referaat for all their feedback and suggestions.

### REFERENCES


