How Do People Really Interact With TV? Naturalistic Observations of Digital TV and Digital Video Recorder Users

MICHAEL J. DARNELL
Microsoft TV, Mountain View, California

_____

To understand how people really interact with digital TV and digital video recorder (DVR) systems, 10 people were recruited, 6 with DVRs and 4 with non-DVR digital cable or satellite TV systems. Video-recording equipment was set up in their homes to record their TV viewing. The person viewing as well as the TV content being viewed were recorded as the person watched TV. Afterwards, the recordings were reviewed and analyzed by the experimenter together with the viewer, who then described his or her behavior. The results show that overall the most common way people interacted with TV was in avoiding advertisements and finding new programming to watch when the current show ended. These results can be used to find ways to improve how users experience TV systems.

_____

1. INTRODUCTION

It has been argued that the best approach for discovering people's needs when designing products for them is to observe them in their natural environment [Norman 2004]. This study was designed to observe people interacting with TV in their homes over time in as non-obtrusive a way as possible.

Most research about watching TV involves having people answer survey questions or filling out diaries to report their TV watching habits [Nielsen Media 2005; Lee and Lee 2001]. The TV may be instrumented to keep track of what channels people watch and which buttons they press on their remote controls. Although these methods have an advantage in that they can monitor large numbers of people, they also have disadvantages. Surveys and diaries rely on people's limited memory and insight into their

_____

Author's address: Microsoft TV, 1065 La Avenida, Mountain View, CA, 95125; email: mdarnell@microsoft.com

own behavior. Collecting data directly from the TV fails to capture what the person may be doing and what else is going on in the environment. There may only be partial information about what is showing on the TV at the moment. For example, one may be able to determine which program is on, but may not know whether the opening credits, apreview, or an advertisement is currently showing. So if the user fast-forwards, we don't know if he or she is skipping the boring parts of the show or the ads.

The current approach, naturalistic observation with retrospective review, overcomes many of these problems, in that it is able to correlate people's behavior with what is on TV and in their environment without being too obtrusive. It is like sitting in people's homes and watching them watch TV over an extended period of time. The method is also able to capture the goals of people's behavior because the TV viewer reviews the recordings of their recent TV watching and describes what they were doing and why. Thus, this method depends much less on people's unaided memory, and is minimally intrusive. This approach is ethnographic in nature, yielding rich data from a small number of participants [Microsoft 2005].

1.1 Previous Ethnographic Studies of TV Viewing

There have been previous studies using ethnographic methods to understand how people interact with TV.

Cornwell et al. [1993] installed video-recording equipment into people's homes to record their TV viewing and related behaviors. These authors focused on contrasting their methodology with other non-observational methods such as surveys and self-report methods. Their findings did differ from those of non-observational methodologies and were closer to the findings of other direct observation studies. Cornwell et al. were also interested in the amount of commercial (ad) avoidance during TV viewing. They found people changed channels on the average once per every 8.7 minutes of ad viewing versus once every 10 minutes while viewing program content; the most common interactions were changing channels sequentially (grazing) and changing channels during ads (zapping).

Eastman et al. [1995] conducted an ethnographic study involving 159 households. In the study, observers sitting in people's living rooms coded TV viewing behaviors over the course of one evening. Eastman et al. looked at differences in the use of the remote control in relation to gender, age, size of viewing group and program content. Overall the authors found that punching in channel numbers to be the most frequent way to change channels, followed by surfing, and finally by recalling the last channel. More channel changing occurred during program breaks on the hour and half hour than during programs. Eastman et al. concluded that ad avoidance was probably overstated in prior research. Overall, the participants used the remote control once or twice per half hour. However, when the participants were divided into groups according to how much they used the remote control, the high-activity user group behaviors were similar to those reported by Cornwell, et al. [1993].

More recently, Ritson [2003] conducted an ethnographic study of TV viewing by installing video cameras in people's homes (also see Ruzek [2004]). Ritson's study focused on what people did during ads; he compared his observations of the participants' behavior during ads to the answers participants gave in their interviews. He found large discrepancies between what people said they did and what they actually did. Ritson found that during ads people chatted, did other tasks, many of which involved leaving the room, read, jumped to other channels to watch until the ads were over, or simply surfed away temporarily. Finally, people did watch some ads, but it was a "minority activity."

1.2 Generalizing Previous Findings to Digital TV and DVR Systems

These previous studies were conducted on traditional analog TV systems where the users could only change channels by surfing, punching in channel numbers or by using a recall last channel button. More recent TV systems have more functionality, including on-screen interactive guides and digital video-recording features (DVRs). The new features are believed to have a significant impact on people's TV viewing behaviors. Recent survey research shows that on average 50% of what people with DVRs watch is recorded [Choate 2003]. Also, 90% of DVR users claim to watch less than 25% of the ads when playing the programs back, and 64% claim to watch none of the ads!

The aim of the current study is to use the naturalistic observation method to see how people use TV systems with DVR and on-screen interactive guides. A specific goal is to see how people interact with their TVs most of the time, which will help us find opportunities for enhancing the TV viewer's experience.

In the United States, people have three popular TV systems in their homes with interactivity that is beyond simply changing channels: digital cable, digital satellite, and digital video recorders (DVRs). Digital cable and digital satellite services are in about 40% of US households; they are characterized by an on-screen interactive guide which shows a programming schedule for each channel. DVR systems also have an interactive on-screen guide, but in addition viewers are able to easily schedule and play back a large number of recordings from a hard disc; and DVR features are being added to digital cable and satellite receivers. It is expected that the penetration of DVR, currently at about 8% of households, will increase rapidly to a point where most households will have it.

2. METHOD

2.1 Participants

Ten people, 7 females and 3 males participated in the study, 6 with DVRs and 4 without, but with digital cable or satellite TV; see Table I. Experienced users were selected, that is, those who had their system for at least 6 months and watched at least 21 hours of TV per week; personally used the remote control while watching TV; and were not employed in a TV-related business. They typically watched TV during primetime; were between the ages of 31 and 60; lived in multiperson households; used their on-screen interactive guide at least 4 days per week; and mostly did not watch TV in a bedroom or where observation would not be welcome.

| Table I. Participants' Characteristics					
Participant	Condition	Married	Gen.	Age	System
1	DVR	Yes	F	39	TiVo S/2 w analog cable
2	DVR	Yes	F	48	TiVo DirecTV
3	DVR	Yes	F	39	TiVo S/2 w digital cable
4	DVR	Yes	F	37	TiVo DirecTV
5	DVR	No	F	37	TiVo S/2 w digital cable
6	DVR	No	F	41	Comcast DVR
7	Non-DVR	No	M	39	DirecTV
8	Non-DVR	No	F	59	Digital cable
9	Non-DVR	Yes	M	31	Dish Networks
10	Non-DVR	Yes	M	39	DirecTV
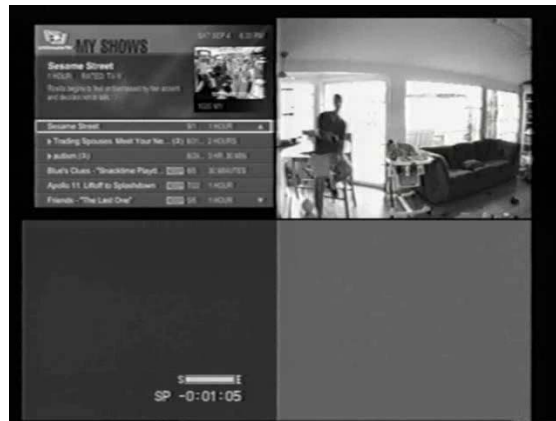
Fig.1. Output video stream with 4 quadrants showing clockwise from top left: (1) set top box output; (2) video camera view of living room from top of TV; (3) blank quadrant; (4) VCR output (not playing a tape).

2.2 Equipment

In each household, a wide-angle video camera with microphone was positioned on top of the main TV with a view of the seating area in front of the TV (see top-right quadrant in Fig. 1). The video output of the digital cable, satellite, or DVR device was captured, as well as the video output of the VCR and DVD player. The video camera and video outputs were fed into a video processor which scaled the video feeds to ¼ their normal size and arranged them into 4 quadrants on a single output video stream (see Fig. 1). The output video stream and camera audio were input to a VCR to record onto 8-hour VHS tapes.

2.3 Procedure

Three 2-hour visits were made to each participant's home. On the first visit, the participant was interviewed about his or her TV viewing, the recording equipment was set up in the home, and the participant and experimenter scheduled the recording of up to 8 hours of TV viewing over the coming week, usually in 3 to 4 sessions, based on the participant's typical viewing habits, he or she was told that if they missed watching TV during a session, that was OK; they should just watch TV when they would normally.

On the second visit, the recordings were reviewed and analyzed for two hours by the experimenter and participant. The experimenter fast-forwarded through the recordings to every point where the person interacted with TV or where a significant event happened (e.g., the participant left the room). The participant was asked to describe what he or she was doing and why. At the end of the second visit, the participant and experimenter scheduled the recording of up to 8 more hours of TV viewing over the coming week.

On the third visit, the recordings were viewed and analyzed for two hours and the recording equipment was removed. Participants were compensated $240 US for taking part in the study.

3.   RESULTS

Up to 16 hours of TV watching were recorded for each participant. However, the 2-hour duration of the visits only allowed from 2 to 4 hours of the recordings to be reviewed during each visit; and some participants were only recorded watching TV for a few hours in total. The following results include 132 minutes of reviewed video for each participant.

This number is based on the participant with the smallest amount of reviewed video. For the other participants, the reviewed sessions were randomly ordered and the first 132 minutes were included in the results.

Data comprising the reviewed sessions is composed of a series of time-stamped observations, including specific sequences of interactions with the TV system and the participant's high-level goals. The sequences were delimited in time by relatively long periods of TV viewing.

An example observation follows. The participant watched a recording of the TV program "Frasier".

Frasier ending. Skips ads until final scene during credits, watches some, then stops the program, deletes the program now, goes to now playing, chooses Seinfeld and plays it.

For purposes of analysis, the observations were coded at a high level of abstraction, which allowed comparisons to be made across different TV systems. This is the high-level behavior for the example above:

End of show, Find another

Observations were also analyzed by listing the interactive TV functions, those in the "Frazier" observation include:

- 30-sec. skip
- stop playback
- delete recording
- go to recording list
- play recording

3.1 Results for Participants with DVRs

Because participants with DVR systems spent the majority of their time viewing recordings, the data from those with DVRs and those without were analyzed separately.

The 6 participants with DVRs spent an average of 36% of their time watching live or delayed TV and 64% of their time watching recordings; two of them *only* watched recordings. Previous survey research reports that overall DVR users watch about 50% live and 50% recorded programs [Choate 2003]. However, Bernoff [2004] reports that TiVo users claim that they watched only about 35% live and 65% recorded programs.

Women aged 18 to 44 watched only about 35% live TV programs; the current sample includes 5 women in this age range. Five participants with DVRs watched programs alone during the session; one participant watched TV with another person 88% of the time.

Figure 2 shows the most frequent high-level behaviors for the DVR participants, including the number who exhibited the given behavior.

Skipping ads while watching recordings or delayed TV was by far the most frequent interaction DVR users had with their TVs. Three of the participants had their TiVo set up to use the *30 sec. skip* feature. The other three participants did not, and used *fast forward* to skip ads, mostly using speeds 2 or 3, and skipped nearly all the ads they encountered.

Out of about 55 pods (i.e., clusters) of ads, participants were observed to watch (with full attention) only 1 complete pod and 13 individual ads. Of the 13 ads, 8 were previews, 5 of which were the first ads in their pods. Watching with full attention was defined as watching the TV at playback speed with no chatting, no interaction with the TV, no leaving the room, no muting the sound, no talking on the phone, and no reading.
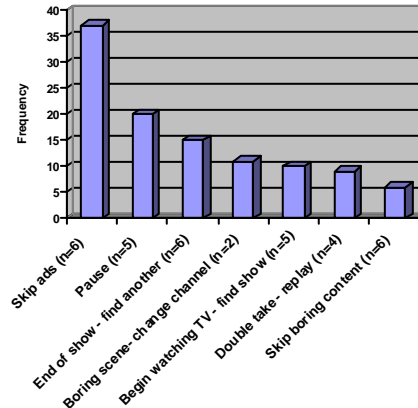
Fig. 2. DVR: Most frequent high-level behaviors and number of participants (of 6) exhibiting the behavior.

After skipping ahead, participants used *rewind* about 30% of the time because they either overshot the point where the program resumed or because they missed something among the ads, such as a preview.

Pausing was second to skipping ads as the most frequent DVR behavior. When participants paused, it was not generally associated with events on TV, but with internal events (e.g., they felt hungry) or events in their environment. Participants generally paused the program to do something else in the room (e.g., concentrate on the computer, chat) or to leave the room. Pauses were most often brief (2 minutes or less). Only 2 of 20 pauses involved live or delayed TV.

The end of the current TV show was the third most common event that caused participants to interact with the TV. When the current recorded TV show was over, people almost always skipped ads; watched the last scene or preview of the next episode; stopped the playback; deleted the recording; went to their list of recordings and played another recording. Occasionally, they went to the guide and chose a live show to watch. Navigating the recording list included both arrowing (using the arrow buttons on the remote to move down or up, one line at a time) and paging (using the page buttons on the remote to move down or up a screen at a time). Occasionally, this was accompanied by getting program information to see if the program had already been seen.

Navigating the guide included arrowing, paging, and punching in channel numbers in the guide. In the guide, navigation generally involved *island hopping*, that is – punching in a channel number to get to a section of the guide, arrowing (and sometimes paging) in that section, occasionally looking at info for a program, then punching in another channel number to get to another section of the guide and arrowing around in that section. Typically, this procedure continued through two to three islands. When the participants got program information, it was generally to read the first few lines of the synopsis of the program.

The fourth most common high-level behavior usually occurred while watching live TV; that is, people *switched channels* to avoid boring (non-ad) content, or to look for better content. Most frequently, they switched channels by using the recall button on the remote to jump to the previous channel. The recall function has a wide variety of labels
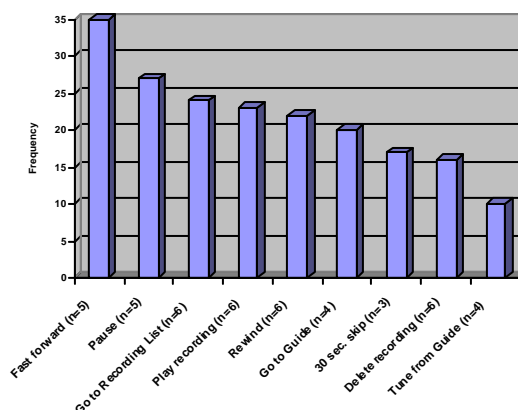
Fig. 3. DVR: Most frequently-used functions and number of participants (of 6) using the function.

on various systems, including last, previous, recall, and jump. Less frequently, people went to the guide or to the list of recordings to find a show to watch.

The fifth most common high-level behavior occurred when participants entered the room and *began to watch TV* during the session. They mostly went to the recording list and played a recording; less frequently, they went to the guide to find a live show to watch. The recording sessions were scheduled so that each participant knew what time his or her TV viewing would be recorded. Sometimes recording started while the participant was already watching TV; at times the participant started to watch TV during the session; and sometimes he or she never watched TV during the session.

The sixth most common high-level behaviour, *a double take,* occurs when a participant watching TV realizes that he or she missed something and rewinds to the part missed. Generally, participants who did this were also doing something else while watching TV (e.g., using a computer, reading a magazine). This level of attention was characterized as level 2 by Lee and Lee [1995]. The viewer is in front of the TV but divides his attention with another activity.

The seventh most common high-level behavior was *skipping boring content* by fast forwarding or 30 sec. skip, and generally meant skipping a single scene at a time.

Figure 3 shows the frequencies at which participants used the most common functions together with the number of participants using the given function. A function such as 30 sec. skip was treated as a single use, since it behaves like an unbroken sequence.

3.2 Results for Participants without DVRs

The four participants without DVRs were three men and one woman. One of them watched TV alone during the session and the other three watched with another person most of the time (97%, 89% and 89%, respectively).

The participants without DVRs had digital cable or satellite TV but no DVR capabilities. Digital cable and satellite systems provided an on-screen interactive guide, which has a schedule of TV shows for each channel, as well as information about each program, most notably a synopsis; they also have favorite channel and reminder features.

Figure 4 shows the most frequent high-level behaviors for participants without DVRs, including the number of participants who exhibited a given behavior.
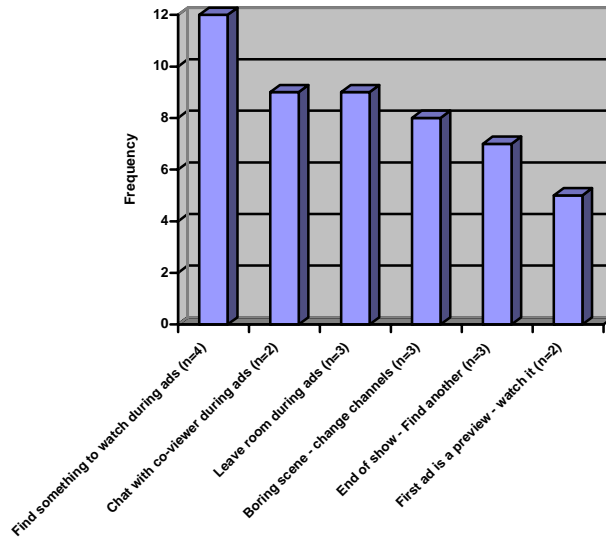
Fig.4. Non-DVR: Most frequent high-level behaviors and number of participants (of 4) exhibiting the behavior.
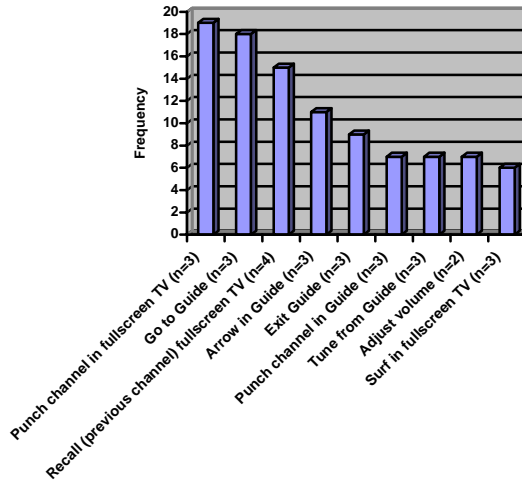

Fig. 5. Non-DVR: Most frequently-used functions and number of participants (of 4) using the function.

The appearance of a pod of ads was the event that most frequently led a participant to interact with the TV. The most frequent high-level behavior was to *find something else to watch during the ads,* with the intention of returning to the program when the ads were over. Most often, this was done by punching in a number for another channel, usually followed by *surfing* in the vicinity using the channel +/- buttons on the remote (i.e., island hopping), or using recall to jump to the previous channel. Some participants went to the guide to find another program to watch.

Island hopping commonly occurred in full-screen TV as well as the guide, that is, punching in a channel number to get to it and then *surfing* in the vicinity of that channel, then punching in another channel number to get to another one, and surfing in that vicinity.

The second and third high-level behaviors that occurred during ads were, respectively, *chatting with a co-viewer* or *leaving the room*

Out of about 47 pods of ads, the participants were observed to watch, with full attention, only 2 complete pods and 6 individual ads. Of those 6 ads, 5 were previews and were the first ads in their pods.

The fourth most frequent high-level behavior was *changing the channel because the show is boring* or the participant thinking that there was something better on another channel. Most typically, this was done by using recall to jump to the previous channel, or by using guide (or favorites) to find another program to watch.

The fifth most frequent high-level behavior was *finding something else to watch when the current program ended*. Most often this meant going to the guide to find something to watch; less frequently, it meant island hopping in full-screen TV.

The sixth most common high-level behavior was *watching the first ad in a pod when it was a preview* for the current show, for another show on the same channel, or for a movie.   Subsequent ads in the same pod were sometimes watched and sometimes avoided.

Figure 5 shows the most frequently used functions by the participants without DVRs.

4. CONCLUSIONS

4.1 Summary

Participants with DVRs mostly watched recordings; their most frequent interactions with the TV were skipping ads, pausing, and finding something else to watch when programs ended. The functions used the most had to do with controlling the video stream: fast forward, 30 sec. skip, pause, rewind and play. Other highly used functions were finding and playing recordings and finding and tuning to live TV programs.

Participants with digital cable or satellite but without DVR interacted with TV mostly by finding something else to watch when ads or boring content occurred or when their program ended; they also took advantage of ad breaks by chatting with co-viewers or by leaving the room. The most highly used functions were changing channels by punching in channel numbers, recalling the previous channel, and going to the guide to find a show.

4.2 Comparisons to Previous Research

One of the aims of this study was to see how previous ethnographic research on analog TV systems generalized to digital TV and DVR systems.

The results of previous ethnographic studies can be most easily compared for participants without DVRs in the current study, since participants with DVRs spent so much time watching recordings. The present results are largely consistent with the Cornwell and Ritson studies, in both of which small numbers of participants were videotaped. In those studies, the occurrence of ads was a major reason for channel changing and other activities; actually watching ads was a minority activity. The current study corroborates their findings for users of digital TV systems with interactive on-screen guides. The on-screen guide provides another way to find something else to watch during ads.   The results are also largely consistent with the Eastman et. al. [1995] study, in which the participants were high-activity users.

4.3 Contrasting DVR and non-DVR Behaviors

For both DVR and non-DVR systems, interacting with the TV was mostly driven by ads. DVR users playing back recordings mostly skipped ads; non-DVR users mostly changed channels, chatted, or left the room.

For both DVR and non-DVR user/participants, interacting with TV was also driven by the ending of the program they were currently watching. Participants with DVRs typically deleted the recorded programs after watching them and then found another recording to watch. Participants without DVRs typically used the guide to find other shows.

Interaction was also driven by boring content or the belief that there was something better to watch. Participants with DVR skipped scenes while watching recordings; participants with DVR and those without DVR who watched live TV used the recall function to jump to the previous channel or the guide to find something else to watch.

4.4 Advertisements

This study didn't focus on ads, but it turned out that ads were the most common cause for people to interact with their TVs. For DRV users, this almost always involved skipping ads while watching recordings. The results are consistent with the findings from surveys by Choate [2003] and Bernoff [2004]: people report skipping most of the ads when playing back recordings.

Users without DVRs frequently changed channels while the ads were on and watched something else, but returned when the ads were over. Ritson [2003], who observed TV viewers in London, found similar behavior.

The types of ads watched most often were previews or promotions for TV programs, showing that people do watch ads that are relevant to them. If ads could be better *targeted* to match the interests of viewers, they would probably be watched more often.

The first ad in a pod is more likely to be seen, and when it is a preview, watched more frequently. After watching a preview, viewers are more likely to watch the entire pod of ads. This corroborates Ritson's [2003] research which shows the superiority effect for the first ad in a pod (but see TV Dimensions 2006).

The results support the idea that DVR users may actually be exposed to more ads than non-DVR users; see Roberts [2005]. Although DVR users skip almost all the ads in their recordings, they must monitor them with full attention to do so. In contrast, non-DVR users typically change the channels when ads come on or leave the room, completely avoiding exposure to whole pods of ads.

If TV ads contained some time- and position-constant information (e.g., a logo or simple message), they would probably be more effective while being skipped. People monitoring the ads in order to skip them would be likely to perceive that information.

Partial-screen user interface panels, partially-transparent screens, or screens with picture-in- picture (PIP) whose function is to allow users to watch TV while getting additional program information, expose people to more ads. People generally use program information screens during breaks for ads or at the end of programs when ads are shown.

People frequently change channels during ads, intending to return when the ads are over. A feature that allows people to easily monitor the ads on another channel in a PIP, while watching something else, would not only expose people to ads on the other channel but make it easier to return to the original channel in a timely manner when the ads end.

Pausing was frequent on DVR systems; people would often remain in front of the TV while it was paused. This could be an opportunity to display ads.

4.6 Implications for Design and Development

One of the purposes of this study was to find ways to enhance the user experience of new TV systems.

The most frequently used DVR features were those to skip ads and boring content (fast-forward and 30 sec. skip). Some features were used for remedial purposes--for example, rewind was used about 30% of the time following skipping. Future studies could identify optimal fast-forward and skip speeds and behaviors to make these processes as efficient as possible.

Participants with DVRs typically watched multiple recordings in a sequence during a session. When a show ended, significant navigation was required to play the next show. To the extent that people know in advance what programs they would like to watch, a playlist feature would make this process more efficient.

Participants without DVRs used the recall function quite frequently to jump between the last two channels. However, they often monitored more than two channels, especially when they pressed recall and found that ads were being shown on the previous channel as well. A recall feature that involves more than two channels might make this process more efficient, especially if it allowed people to monitor ads in a PIP.

The most frequently used functions and sequences indicate which features should be made most efficient to use. For example, having a separate recording list button on the remote control would be more efficient than having to navigate through a menu to get to the recording list.

One of the benefits of DVR is to save time by allowing users to skip ads and boring content, which could be taken to an extreme by allowing people to play recordings back with audio at a slightly greater than normal speed (e.g., 1.2x – 1.4x) This could save time without noticeably degrading the video experience.

4.7 Implications for Usability and Quality Assurance

Knowing the most common user behaviors and patterns in their interactions with TV will improve the validity of our usability testing of TV products, since we can incorporate these frequently occurring scenarios into our usability tests. These scenarios should also be included in quality assurance test cases and be given an appropriate level of priority.

4.8 Future Quantitative Studies

Although the ethnographic methodology of the current study is able to provide rich and detailed data, we do not know how representative the results are. Such knowledge requires a larger sample size whose relationship to the population is known. However, using this ethnographic methodology with a large sample size is probably prohibitive due to time constraints.

But this study can help us generate hypotheses that could be tested using other methods such as surveys or data obtained by computer from STBs in people's homes. For example, the frequency with which the *pause* button is pressed or the guide is accessed by a large random group of users is data that could be obtained from cable operators and compared to the results of this study. This data, together with an independent means of assessing how representative it is, would give us an accurate understanding of how people really interact with TV. Finally, it would be interesting to do further research on the probable effects of gender, age, the size of the viewing group, and TV content on the interaction with digital TV and DVR systems.

ACKNOWLEDGMENTS

REFERENCES

BERNOFF, J. 2004. *The Mind of the DVR User*: *Media and Advertising.* Forrester Research, Cambridge, MA.

CHOATE, J. 2003. *The DVR Monitor Wave IV*. C Cubed, LLC.

CORNWELL, N.C., EVERETT, S., EVERETT, S.E., MORIARTY, S., RUSSOMANNO, J.A., TRACEY, M., AND TRAGER, T. 1993. Determining RCD use: Methodology matters. In, *The Remote Control Device in the New Age of Television,* J. Walker and R. Bellamy, Jr. (eds.), Praeger, New York, 43-56.

EASTMAN, S. T. AND NEWTON, G. D. 1995. Delineating grazing: Observations of remote control use. *J. Communication 4,*1,77-95.

LEE, B. AND LEE, R.S. 1995. How and why people watch TV: Implications for the future of interactive television. *J. Advertising Research 35*, 6, 9-18.

MICROSOFT. 2005. Making technology conform to peoples' lives. http://www.microsoft.com/presspass/features/2005/apr05/04-04Ethnographer.mspx. Retrieved Nov. 29.

NIELSEN MEDIA. 2005. What TV ratings really mean (and other frequently-asked questions). http://www.nielsenmedia.com/whatratingsmean/. Retrieved Nov. 29.

NORMAN, D.A. 2003. *Emotional Design: Why We Love (or Hate) Everyday Things.* Basic Books, New York.

RITSON, M. 2003. Creative business: Talking, reading, tasking. *Financial Times* (Feb. 4).

ROBERTS, J. L. 2005. Play it again, TiVo. *Newsweek.* http://www.msnbc.msn.com/id/10073890/site/newsweek/. Retrieved Nov. 29.

RUZEK, P. 2004. Ignoring TV ads is the norm, research shows. *Marketing Update.* http://www.ami.org.au/AMIMU/04Oct/0410_advertising_webcast5.htm. Retrieved Feb. 17, 2006.

TV DIMENSIONS. 2006. *In-Pod Commercial Positioning.* Media Dynamics, New York, p. 335.