

# Understanding Videowork

David Kirk<sup>1</sup>, Abigail Sellen<sup>2</sup>, Richard Harper<sup>2</sup> and Ken Wood<sup>2</sup>

<sup>1</sup>The Mixed Reality Laboratory  
University of Nottingham  
Nottingham NG7 2RD, UK  
dsk@cs.nott.ac.uk

<sup>2</sup>Microsoft Research Cambridge  
7 JJ Thomson Ave  
Cambridge, UK, CB3 0FB  
{asellen, r.harper, krw}@microsoft.com

## ABSTRACT

In this paper we elucidate the patterns of behavior of home movie makers through a study of 12 families and a separate focus group of 7 teenagers. Analogous to a similar study of *photowork* [13], the goal is to provide a deeper understanding of what people currently do with video technologies, balancing the preponderance of techno-centric work in the area with appropriate user-centric insight. From our analysis, we derive a *videowork* lifecycle to frame the practices users engage in when working with video technologies in the home, and uncover two broad types of video usage therein. This has implications for how we conceive of and devise tools to support these practices, as we discuss.

## Author Keywords

Video, Home Movies, Editing, User Research, Contextual Inquiry.

## ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## INTRODUCTION

One only has to look at online repositories of video such as YouTube to begin to understand how growing access to digital video is widening participation in a new culture of video production, exchange and viewing. As the capacity to video is being incorporated into increasingly diverse artefacts (such as mobile phones), the opportunities for non-professional video-makers to make, watch and exchange video have equally increased. Accompanying this rise in the prominence of video has been a surge in interest in providing editing tools. Despite this, and as one cynic has noted, “..far more amateur video is shot than watched, and people almost never edit it.” [7, p.3]. Efforts to introduce so-called user-friendly editing tools (and Apple’s I-Movie is usually brought to mind here) do not appear to be

addressing this issue. For some reason, editing seems a step too far for most users.

It seems to us that there are two possible courses of action that can result from this conundrum: one, the most common in the literature, is to devise tools to help *automate* the process of video editing [10, 12]. Here the view is that users need intelligent systems to do the work for them. In this conception, the role of the software becomes one of drilling down and distilling from users’ ‘raw footage’ a consumable or, if you prefer, a usable product: a video that delights. Given this goal it is perhaps not surprising that some researchers have even offered automating tools for the shooting of video, unlikely though it might sound [2].

If one follows this line of inquiry too literally, one might suppose that the user can be taken out of the loop altogether, of course. However, this is not the only route that follows on from the assertion that users don’t seem to edit. Another might be to discover what it is that users are doing when they take, share and watch video and thus perhaps find out why it might be that they don’t appear to edit as much as one would imagine. In short, one might go and look and see what it is that happens when people work with video.

We report a study in which we took this latter view. We took this view because too many researchers seem to have taken the other path, the technology one, showing in the process little or no interest with users, and even if they do, only appearing to acquiesce to *assumptions* about what users want to achieve with video. Evidence about user behavior is not the forte of these papers. The assumptions that they rely on may seem plausible, of course, and their adoption by the researchers in question was doubtless made with the best of will. The point is that few if any studies have inquired into what users do with video or what they might want from video tools.

More particularly, much as a recent paper on *photowork* sought to illuminate the working practices of users of digital pictures [13], we wish to similarly articulate the notion of *videowork*: a name for the practices in which users engage when ‘working’ (if that verb can be allowed) with video technologies. It is this work we have attempted to capture in our user studies. This paper, thus, seeks to provide a deeper level of understanding of what people do with video, documenting whatever capturing, editing,

*This space deliberately left blank for ACM Copyright notice. This version submitted to CHI 2007.*

archiving or sharing behaviors they undertake, and offering what we think is a useful interpretation of these behaviors.

Our case will be that one finds two broad practices of *videowork*. One is lightweight, or ephemeral, if you like, entailing the ad hoc capture and sharing of content on the capturing devices themselves, mostly, but with very little editing. The other is more planned for, more heavyweight, entailing a systematic capturing of content that is sometimes edited, and most often (if not always) prepared for sharing or viewing not on the capturing device itself but on TVs and PCs often via tangible products (such as DVDs). The main goal of the former kind, we will claim, is to support the spontaneous celebration of the ephemeral through visual means; the goal of the latter, the more serious, to underline the importance of the event captured, and to make an embodied product - a tape, a DVD - that stands as a physical embodiment of the visual traces of the event in question. In this latter case, editing often adds value to object produced, but the quality of that editing is of insignificance when compared to the importance of having the history captured in the first place. In this view, editing matters, but not as much as the video footage itself.

With regard to each type of videowork, this characterization might seem to be leading us away from the original topic, the design of editing tools. Skeptics of user studies may think that this study will offer little more than a description of what people do. However, we will show that this kind of study undertaken to reveal the “user perspective” can help frame what kind of editing users want to undertake, and thus point towards design issues about how to help, aid or even automate specific editing practices. We shall argue, also, that there are some more fundamental issues about the overall experience of videoing that the research raises, issues we think have been completely neglected by designers but which provide interesting possibilities for new design solutions.

## BACKGROUND

The extant literature in ‘working with video’ can be broadly split into two main themes. The first has to do with the construction of software tools to support casual browsing of large collections of video. Several such studies [4, 14, 20] have attempted to find novel ways of parsing video clips/movies to chunk and display them using representative key frames which enable the critical content of video clips to be found more expediently. This implicitly assumes a model of video browsing as the same as digital photo browsing. Previous research in the digital image field has demonstrated, however, that such implicit assumptions about the form of browsing and searching behaviors can be shown to be misleading when actual user research is undertaken [13].

By and large the majority of work in the video technologies arena has, however, focused on the role of software in the automating of the video editing process. One particularly prominent strand has been the development of the ‘Hitchcock’ system at the FX Palo Alto Lab [8, 10, 11], the development of which is guided by a concern that (with

reference to home videos) “the interesting parts of the video are usually buried in long sections of boring or bad quality video, and only a few close friends and family are willing to sit through it all” [8, p.660]. Consequently they have developed their system to automatically parse raw home videos, cutting out sections that the computer has deemed to be of poor quality. The system then prepares clips of video and organises them into what are referred to as “meaningful piles” (either on time sequence or colour similarity), and from these the user is able to construct their new video. Comparable systems have been developed by Yip et al [19] and Hua et al [12] whose work in particular points to a future when even more of the process could be fully automated by using more sophisticated software to mine the raw footage to uncover shots which better expose the intended semantics of the footage.

These efforts have highlighted, however, a difficulty with editing out so much of the raw footage. As Yan and Kankanhalli argue, since “home videos constitute footage of great sentimental value, such videos cannot be summarily discarded” [18, p.107]. The problem, in other words, is not solely related to how to edit; the problem, if that is what it is, starts with the content in the first place. No editing tool can make ‘poor’ content good.

Such an understanding has led to the idea that even more automation is in order. This automation would help the users gather better footage in the first place.

For example, Adams et al [2] constructed the Integrated Media Creation Environment (IMCE) in an ambitious attempt to distil key concepts of film theory into a software package. Used in conjunction with a PDA (personal digital assistant), the IMCE guides the amateur ‘videographer’ in planning a series of shots before the point of capture. It uses pre-defined concepts of film genres and events around which the user can ‘construct’ his or her own video.

Unfortunately, there is simply insufficient evidence to know whether such tools aid or hinder users. Reports of video editing software, from the fully automated versions [12], through to the less directive systems such as Silver [5], all suffer from a common lack of concern with user behavior, with what one might call usefulness (as against a more pedantic concern with whether users can use the software in question). All they offer are limited evaluation studies designed to ratify that the system built can be used, or used after a period of initial training [5, 10]; not whether it is useful. In the case of the more automated systems, the only concern seems to be whether the tools make video capture quicker than a manually edited film, not whether it is better video [12].

The closest the HCI literature comes to the study of actual existing home video practice is work from Georgia Tech on the Living Memory Box project [1, 15]. However, the development of the Living Memory Box focused exclusively on the use of media to record memories, ignoring the creative aspects of the technology which are integral to a full understanding of videowork. Additionally,

the more video-specific aspects of the research [1] have deliberately been chosen to avoid the issue of video editing, considering that the browsing, annotating and retrieving of centrally stored home video snippets is of more importance.

Perhaps in an effort to acknowledge this, Abowd et al conclude, when attempting to discuss how such video technologies might come to be implemented in real users' homes, that:

“We have not spent enough time understanding the needs of the wider population of potential users of such a system. Deeper consideration of the needs of both casual and devoted family movie archivists should be studied to inform the further evolution of this system” [1, p.07].

### FIELD STUDY

With this “deeper consideration” in mind, we undertook to conduct a study comprising of two parts. The first consisted of a series of 18 contextual inquiry interviews [3] with 12 families in their own homes (see Figure 1). The participants (12 males and 6 females) ranged in age from 11-65; there were young couples, singles, new families with young children, families with older children and older adults whose children had left home. The second was a focus group designed to elicit a different perspective on issues raised in the contextual interviews and to explore how distinctly younger age groups orient to videoing. Subjects here were drawn from a local high school and consisted of 7 students, all aged 17.

Recruitment of all participants was restricted to people who had at least some experience of creating their own videos, whether it be on the fly on their video phones, or with more sophisticated, bespoke devices with associated software on Macs and PCs.

To give focus to our enquiry we structured our interviews and discussions around a previous framework for photowork [13] making the assumption that there would be sufficient conceptual similarity between the two practices. Considering that both activities are forms of media capture, this approach allowed us to examine as a point of interest those aspects of videowork which differ from photowork. As with photowork, this approach is driven by the goal of unpacking the details for the whole lifecycle of activities that constitute videowork whatever that life cycle might be.

Interview sessions began with a discussion of what technology our participants possessed for video capture and why and when they used some devices rather than others. Motivating factors for purchasing video capture technology were also explored. From this point, participants were encouraged to describe their typical activities when creating their videos. To ensure validity of responses, participants were asked to ground their descriptions in actual examples of practice, by describing exactly how and why they had gone about creating their last three video projects and, where possible, by showing examples of them. This element of the interview facilitated the discussion of how

and why the films had been edited, if at all. Discussion then covered what happened with the participant's movies once they had been created including how they had been archived and stored and how they had been shared.



Figure 1. A participant's home video editing suite.

### VIDEOWORK LIFECYCLE

Figure 2 shows the lifecycle that the contextual inquiries and focus group materials produced. In broad outline, it has four stages: a “pre-capture”, “at capture”, “post-capture” and an “end use” stage. Each stage entails a different set of activities, with key issues motivating or limiting behavior in each. We found, further, that any potential route through the videowork lifecycle was heavily dependent on the technology used for video capture, i.e. used at the commencement of the cycle. Our analysis demonstrated that there are principally four common pathways through the life-cycle corresponding to video capture with analogue or digital camcorders, digital cameras and mobile phones, each offering a unique set of activities. Figure 2 however, encapsulates the whole process to aid in this comparison and to provide an overview. As we shall see, these pathways in fact reflect two distinctly different types of practice, with camcorders representing one type, mobile phones another, and the digital camera falling somewhere in between.

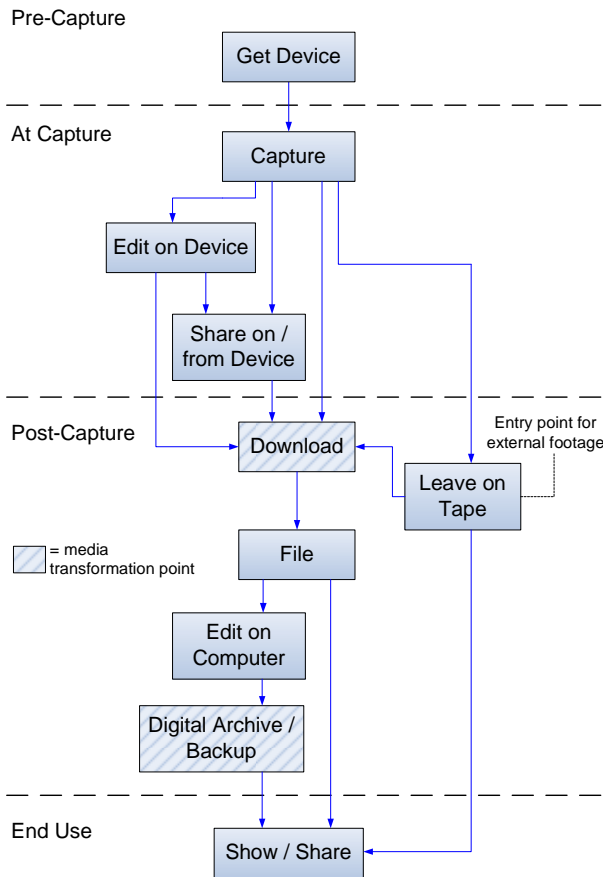
#### Pre-capture

The significance of these different types of practice, and more importantly, the character of the videowork entailed in each, can be seen from the outset. Some of the older participants (interviewed at home) told us that the introduction of video technology into their lives often coincided with a major life change. That is to say, video did not change their lives, it was their changing lives that made video relevant. For one couple it was their wedding and ensuing honeymoon; for eight other families it was the arrival of children. As one participant said:

“We got it actually for ‘M’, when my little girl was born, a month before she was born and it was for that reason mainly.” (Participant 5)

In other words, the first major insight that our videowork lifecycle concept helps identify is the importance of an

event that leads users to identify a value in video. A video is a way that they can document, or if you prefer, celebrate special events.



**Figure 2. The videowork lifecycle**

A second finding, contrasting and yet paradoxically confirming the first, is that those who do not have such special events do not seek out specialized video devices. Most of the teenagers we spoke to, for example, had little interest in using a video camera. Their principal capturing method was to use their mobile phone, which was something that they already possessed. Video for them was a benefit that came from the need for another technology (namely telecommunications), though having got video they made it clear that they were certainly not ones to avoid using it. Thus one cannot say that this segment of user actively engaged in the pre-capture stage, it being more of a stage that they participated in by dint of doing another task: getting a mobile phone.

The contrast here is between types of people in different life stages and how that affects their relationship to video capturing devices and in particular to the purchasing of such devices. This is not to say that one type of person uses video and the other doesn't. Rather, it is to say that the reasons for video use are different. Put colorfully (and

perhaps a little unkindly), the distinction here is one between teenagers who have only interest in themselves, and people who are not only older, but who have in addition a burden, albeit self-imposed, of documenting history. This history is not merely their own, but others too: their children's, their children's children, and so on.

### At-Capture

#### Capture

This distinction--between life stage and technology, between purposes, needs and hopes--follows through all the subsequent stages of the lifecycle. Thus, when it came to the capture of video, our family participants, those who were of a life stage that imposed some kind of responsibility on them, principally had one main tool for capture, despite often possessing various video-enabled technologies. This would appear to be because these subjects felt that only some tools 'were up to the job'. In their view, not any old video device would do, particularly things that were designed for something else. This manifested itself in a variety of ways, one being a resistance to so-called converged or multifunction devices.

"I don't want a kettle that makes toast, like I don't want a microwave I can watch telly on...I want things to do things they're supposed to do." (Participant 15)

For most of this group, the quality of capture was a deciding influence, as was surety of storage of the video for posterity. They did not want to buy a device that 'let them down'. The point here is not simply that their creative efforts would thereby be lost (although there is that), it is more that the moment in question that demanded videoing would be lost. Video cameras had to be 'just' video cameras and 'good to boot' because of the work they needed to do.

Given these motivations, it is hardly surprising that several of our participants told us that the quality of the video recorded on some of the digital cameras they had owned, the capacity of some of these devices, their reliability, and so on, had caused disappointment. The issue here is not, however, simply a question of usability. The point is more complex. It is this: that getting these things right, getting reliability, quality, image and so on done well, is bound up with the importance of the event being captured. This importance was not to be undermined by poor technology, whatever 'poor' meant in any particular instance (difficulty in use, poor focus, sound, color, ability to replay and so on).

The design implications of this are immediately clear, though curiously not mentioned in the literature we have read. They have to do with such things as having sufficient feedback on the device to know, with certainty, that video images are being captured; to indicate the quality of that capture so that it may be adjusted in situ as necessary (i.e. before it's too late) and to reflect the fact though the event is so important that it must be videoed, the videoing of it must not affect the event itself, or excessively handicap the experience of it that the person taking the responsibility to video has.

This last point, a design one we suggest, was something that several of our subjects mentioned as a paradox. For many, it was clear that they perceived costs inherent in the capture of video. They commented on having to remember and plan to take a video recording device along with them, knowing that they would then be encumbered with an awkward device when they wanted to be engaged in what was going on.

“The reality of it is like at ‘L’s birthday party I was so busy doing everything else and in fact when the cake came out somebody shouted ‘video it’ and I went and got my video camera.” (Participant 15)

Not only does this imply something about the level of engagement required when using a device (the users don’t want to be so absorbed in ‘working’ the video that they become like producers of a film; they want to act in the event too!). This casts light, too, on why it was that for most of these subjects, video cameras were not taken with them all the time. In their lives, most of the time, there was no need to video. Video was not necessary to document the routine, the mundane, or the everyday.

The use of digital cameras - and as we shall see also mobile phone cameras - provides a different perspective on issues of user experience and design. One individual explained his delight at the realization that a new digital camera he had bought was robust enough and had sufficient picture quality and memory to store any occasion that he was likely to record, even if those weren’t important occasions. In his view he could risk taking it with him all the time. His behavior underlines, it seems to us, the contrast between video cameras and the purposes behind their use, and the purposes behind digital cameras and camera phone use. ‘Events’ demand videoing, but digital cameras and camera phones are used whenever, for whatever.

This is getting us to the distinct patterns of behavior of the other group in our study, the teenagers: for them, spontaneity in capture was important. In our focus group, this was the main reason why they recorded. But they did not use video cameras; they recorded with their mobile video phones. And the reason for this is only partly for the obvious fact that this is what they had in hand. This apparent platitude hides an important supplementary issue, brought out in the focus group. They could not see the merits of buying a digital camera or video for themselves. And this was not because they had a video function on their mobile phone and thus thought this unnecessary duplication. It was because they recognized that what one did with a video camera was different to what one did with a camera phone. The latter was to play with, something that let them do things on the spur of the moment; the other, something you did when you were being ‘serious’.

It should be added that this did not reflect a lack of interest in videoing on the part of our subjects. Indeed, quite the reverse. It was rather that they recognised a relationship between the type of device in hand and the event in question. Teenagers reported that they spontaneously used

their cameras at any and all times, and not, so to speak, at ‘special times’.

#### *Edit on Device*

At the ‘at-capture’ stage of the lifecycle, one other property of user behavior, also reflecting differences in the two user groups and their chosen devices is worth noting. Editing video on devices at the point of capture was rare and in any case was mainly possible only for the participants who used digital cameras and mobile phones. For those using their digital camera, recently recorded clips would be replayed and possibly deleted if unwanted; for those using mobile phones, triaging at point of capture occurred, which basically meant keeping or deleting. One teenager reporting carrying out other kinds of minor editing tasks (adding titles and fades to a recorded video on his phone before he shared it), but he commented that he had only ever done that once, and didn’t normally see the point of it.

#### *Share on / from device*

A further difference was also noted in what teenagers and other users of mobiles did once they had captured some video. If it is the case that the planful had in mind replaying and editing video in the future, for posterity as we say, for the younger age group with mobile phones, it was to use these video capable devices in the ‘there and then’. And this led in turn another different practice: one of sharing video at the point of capture. Participants reported that they would often immediately share their footage with those they were with. And they described how with the connectivity of the mobile phone this occasionally meant sharing from phone to phone either through Bluetooth connections or using email over longer distances.

In sum, there are two broad schemas of video capture, one entailing spontaneous and ad-hoc capture, some editing at the time of capture, and showing the captured video at the time of taking. This is a “capture-and-share-straightaway” experience, if you like. The other experience is less frequent but much more planned, with greater attention being given to the ‘quality’ of the device that ensures that the work it supported is done. There is little use or amendment of content at point of capture, though what happens later on is important and something we shall come to shortly.

#### **Post-Capture**

If we have begun to discern basic differences in the use of devices at point of capture, then this in turn is reflected in the diverse patterns of videowork in subsequent steps in the lifecycle.

#### *Leave on Tape*

At first glance, these diverse practices seem complex, bound to the foibles of particular video camera type. For example, all of our participants that used a camcorder (either digital or analogue) were concerned about leaving footage on tape (either DV tape, Hi 8 or VHS-C, which were the predominant formats used) all of which cost money. They thus tried to use up the tapes with ‘more than one event’.

“Occasionally if I’ve not got enough tape I find a tape that’s got a bit on the end and then I just add a bit on the end of it.” (Participant 15)

Digital users left content on tape either as a means of archiving the footage or as a means of keeping a physical copy. Given the cost of tapes all of our respondents who used camcorders suggested that they had captured multiple events on individual tapes to prevent wastage.

Clearly for those with analogue camcorders the process of leaving the footage on tape was also because they had no means by which they could edit the footage. However, an interesting by-product of this was that the footage was immediately available to share although this was only possible in a local context (camcorder and viewing device needing to be physically linked, which we will say more about later).

The participants who had used analogue camcorders also reported a special form of download practice of tape-to-tape transfer (which we have shown on our videowork diagram, Figure 2, as a link between the ‘leave on tape’ and ‘download’ boxes). It is principally a means by which the recording tapes can be freed up to be used again. To do this, participants reported that they would connect their camcorder to a TV and play back footage through the TV whilst recording to VHS..

#### *Download*

For those with video footage in a digital medium, the download stage was the first key element of the post-capture stage of videowork. All of our users of lightweight capture devices (mobile phones and digital still cameras) downloaded all of their footage to their computer at some stage. Likewise for all our users of digital camcorders there was at some point transfer to another digital medium.

For four of our families who principally used a digital camera to record video, the downloading of footage was performed in parallel to the downloading of digital pictures, and in essence the video clips were treated as digital photos in terms of archive and storage and reasons for download. (These issues have been detailed elsewhere in the literature [13]). One participant (who had rejected his camcorder in favour of his digital camera) was comparatively pleased with the ability to download from his digital camera at better speeds than he could manage with his camcorder, because of the need for real-time downloading with the camcorder.

“Transferring images onto the PC from the digital camera is a lot more straightforward...it takes a hell of a lot of time if you want to convert your miniDV cassettes or some of it or part of it into the PC format.” (Participant 4)

This issue of the time cost of downloading was something raised by several of the digital camcorder users. They reported being frustrated by the length of time it took, but this didn’t stop them from wanting to download and edit their footage.

For the mobile phone users, download was clearly an issue of memory management. Our teenage participants informed us that they had (on average) eight video clips (never more than a couple of minutes each, maximum) stored on their phone for posterity. As they captured clips, ones they liked would be downloaded at home to the computer, but they often left a copy of the best ones on the phone for easy access and sharing. For those participants that used digital camcorders, the download process was dictated by two main strategies of use. For three of our participant families there was a more regular cycle of capturing an event and then downloading the footage after the event to edit it for a specific event-focused video (such as a holiday compilation). For our other two digital camcorder users events were captured on a more ongoing basis, being left on tape until a sufficient point was arrived when someone wanted to edit together a large number of clips that had built up on the tape(s) over a period of time. Examples of this kind of download include the production of ‘baby’s first year’ type of films.

“So this is the second DVD I did of my little girl... we had like the baby footage until one year old and then this was like one year old until two-ish.” (Participant 5)

#### *File*

As video footage was loaded onto the computer it was immediately filed in some way for later retrieval. As highlighted above, for the lightweight capture technologies there was a sense in which the videos were treated similarly to photos. Our participants told us that they tended to keep their videos in either a ‘my videos’ folder or in amongst their digital pictures in their ‘my pictures’ folder. They described typical folder naming conventions for the storage of event based collections of clips. Our participants who used more heavyweight capture and transferred larger files from digital camcorders showed us how they stored that footage on their computers. Again there was a tendency to store the footage in an unprocessed format in a standard folder structure, again adopting regular naming conventions. What was of particular interest however, was the fact that only a small proportion of their video footage was actually stored on the computer at any given time. And it was usually sections of video that they were currently working with or intending to work on. Reasons given for not storing more of their collections on the computer included not wanting to take up too much space on the hard drive, not trusting the computer enough to store it safely and not wanting to actually view the footage on the computer so needing it stored in other formats.

#### *Editing on Computer*

Once the footage was downloaded to the computer and put into (potentially) temporary filing, some editing of the footage was normally undertaken. This was particularly the case with those downloading from camcorders but was also true for the users of digital cameras. Users of mobile phone video didn’t however engage in editing of their clips—they didn’t see the point of manipulating what were in essence short snippets of action. It would appear that much of the

meta-information that they felt was needed to explain what the footage was, was often represented in the clip title and folder naming conventions.

For those participants that did engage in editing of clips and movies, we observed a variety of practices. Our participants told us that they would crop frames, order clips (from multi-clip projects), add titles, captions and transitions. They described various reasons for wanting to perform these kinds of editing task. For one participant the process of editing and adding extra information to a video was extremely important, as he was sending his movies of family events to relatives back in Mexico and they required Spanish annotation to be comprehensible to his parents. For others, the desire to edit was to be able to have an easily digestible and viewable collection of bits of footage (e.g. the 'baby's first year' style project) and yet for others it was seen as something fun to do. For the adults, producing movies of holidays or pertinent family related scenes were always with the *intention to share*. Furthermore, with this there was a sense in which producing a video was a creative process that reflected on them when viewed by others. To this end they were happy to spend large amounts of time editing their footage. As one participant suggested, he was happy to spend an entire evening/night editing footage for a specific video.

Int: "So you're happy to spend a whole evening doing it [editing a video]?"

Participant 1: "One evening, yes one evening to edit one video, that's not so much time for me."

Whilst engaging in lengthy bouts of editing was not *per se* seen as problematic for the users, actually finding the time to do it was, which explains to some extent why file structures on the computer were used as temporary storage. The time costs of downloading, then editing and then rendering video often meant the tasks had to be split over separate nights of activity.

To perform the editing, a variety of software tools were used amongst our participants. Most of our participants were operating Windows systems on which MovieMaker is available at no cost, although many users were not aware of it. And in some cases MovieMaker had been rejected because it appeared to be too simple compared to reports of other software.

"I'd got the impression that it was a bit crude." (Participant 6)

However more professional editing environments such as Premier were often considered too complicated to use,

"It [MovieMaker] doesn't look as intimidating as Adobe Premier." (Participant 2)

Despite its low profile and perceived simplicity, MovieMaker was used by our most prolific editor for the majority of his editing tasks. However, his editing process could never be completed with MovieMaker because he wanted to produce a DVD version of his edited work and MovieMaker is not seamlessly integrated with a DVD authoring environment. This consideration was important to

all of our video editors and was explicitly behind the choice of three of our participants using Macs for video editing tasks, precisely because of the seamless integration between iMovie and iDVD. The production and authoring of a DVD copy of the edited movie was seen as the most important end goal of all of our participants who engaged in video editing practices. The tangible resulting product was seen as a desirable commodity that justified the effort of the editing process and served a variety of purposes of digital archiving, back-up and sharing which we shall discuss in the next sections.

#### *Digital archiving and back-up*

The final element of the post-capture stage described by our participants was the digital archiving and backing-up of the edited films. As discussed above the footage stored on computer was largely temporary (unless we're considering mobile phone footage), with the end archival point being the authoring of a DVD. These DVDs were then stored much as described for the filed tape recordings, usually for easy viewing. Again the creation of a DVD archive stands as a transitional point where the video footage is transformed from one state to another. This instantiation as a solid tangible object reportedly gave our participants a sense of increased security in the safety of the recorded footage.

"You know one day that hard disk can be broken, like an electrical storm or something and then you would lose everything. I don't want to save everything on the hard disk, I want everything on DVDs. I just don't trust it [the hard disk]." (Participant 1)

An interesting point to consider is the quantity of recorded video found in participants' homes. Some of the work presented in the early sections of this paper would suggest that similar to digital photography, the advent of digital video will lead to an explosion in the amount of video that we record. If one considers the video habits of our older respondents however, who used analogue camcorders in the late eighties and early nineties (all three could be considered early-adopters of technology), we get a different picture. There was essentially no set limit to the amount that these people could record beyond the number of tapes they were willing to purchase and store. And yet with these participants there was no endless catalogue of tapes or rooms filled with footage. For our participants, the amount of home video footage, largely centered on family life and recorded over the several years of children growing, on average amounted to probably no more than 10-20 hours of footage. As one of our participants suggested, the choice to capture video tended to decline as changes in the children became less visible.

"I stopped [capturing video] when they stopped changing, like a lot of the photos stop about then." (Participant 16)

Obviously as a counter point, the digital video clips that are recorded in a lightweight fashion (digital camera and mobile phone) will likely increase in number and are perhaps more inclined to be archived on a central computer. However, our sample of teenagers did all agree that if they

had the capacity to create DVDs of collections of their mobile phone generated clips then they probably would do.

### End Use

The final stage of the videowork process that our participants described was the stage of sharing their completed (raw or edited) footage with others (or on occasion watching it themselves). There were various methods by which video was shared, and again here DVDs were particularly useful, either for watching with family or sending to relatives and friends. This issue of sending DVDs is of some importance as it highlights the desire of participants to share video at a distance but acknowledges that the participants who edited larger films were all clear that these were sizeable chunks of data which could not easily be emailed. For at least one pair of participants (who didn't own a DVD burner), the solution was to share via the internet, by posting their video clips to a blog, although this was facilitated by their video being captured on a digital camera and therefore being of relatively small size. For the rest of our adult sample, there was little interest in the notion of being able to post video to the internet, but amongst our youngest participants and our focus group sample of teenagers this was something that they were all either keen on, or had even already tried.

Interestingly, within the families we spoke to, especially amongst those with children older than toddlers but younger than teenagers, watching old videos was relatively common, the request often coming from the children themselves.

“Like ‘L’ really likes watching them so she’ll say can I watch ‘P’, like they know all of the video tapes we’ve got, like ‘L’ was just saying can we watch that one of ‘P’ eating that orange.” (participant 15)

In these instances either a DVD would be put on, or for those who had not edited footage, the tapes themselves were watched. With the older analogue cameras this was sometimes achieved through inserting the recorded cassette into a VHS adaptor or more commonly with more recent models of camcorder, the video camera was connected directly to the TV for viewing. In these instances viewing of the video was performed in the main communal space of the home, where the family's main TV unit was, and the act of watching selected bits of tape was essentially collaborative. For some families, such an event might occur once every few months. What was interesting in particular was that amongst our participants who had older children or children that had left home, the videos often hadn't been looked at for a number of years. This intriguingly parallels the purported decline in recording new footage amongst families as the children grew older.

“This is probably the first time I’ve looked at them in 15 years...you look at them when they’re new when you’ve first seen them” (Participant 9)

When questioned about how they felt about this, parents tended to respond that they didn't really mind, but it was nice to know that the movies were there.

“I think I’m getting to the age now where memories are important aren’t they, you want to be able to remember what’s happened in the past and the joy you’ve had.” (Participant 9)

When we prompted our participants about the potential for having more casual methods of viewing video, such as a ‘video clip screensaver’ they were surprisingly resistant, arguing that a key issue of video was that it demands attention when viewing, pointing out that a reason for capturing video as opposed to pictures was that it caught some sense of motion and sound which was integral to its interpretation, and claiming that this could not be easily parsed in a casual manner like a photo.

This however, is quite atypical as an approach to video for our teenage sample, whose video viewing was much more casual, much more ad hoc, and often immediately after capture.

### DISCUSSION

Having articulated the key elements of the videowork process, we reflect now on some of the central issues the interviews raised. These highlight the ways in which existing approaches to designing technology support for home video users have failed to take into account what users really do, and want to do with video.

#### Two Kinds of Videowork

First, most studies categorically consider ‘video’ as a singular type of process or product. However, our interviews have revealed that even within the narrowly defined scope of “home video” there are at least two distinct types of video data: the kind captured spontaneously in an ad hoc way and the kind captured in a more planned, intentional way. These two forms of video generation are associated with differing capture devices, editing practices, and above all, end-uses. Thus we use the terms *lightweight* for the first, *heavyweight* for the second (see Table 1). We do so not in a literal sense (camcorders typically being physically larger and more cumbersome than mobile phones) but in the sense of the nature of the process - ease of use, ease of connection to other devices and so on - and in terms of purpose - why they were captured, for what ends, for whom.

Lightweight videowork is about creating visual traces of an engaged in event that are used within that same event itself to laugh, to rue, and to reflect on that event by those mostly who are participants within it. Sometimes these materials are kept for posterity, but mostly they are as ephemeral as the events in question. They do not capture things that matter; they simply augment the experience for those who so desire.

In contrast, heavyweight videowork is intentionally about capturing events that matter. Thus the ‘work’ here is greater, with a heavier burden on it. Users ‘work at’ making sure the video is of good quality, for example, capturing all that needs to be captured. Users work too at producing some kind of robust artifact that will stand as a testament of

the event in question. This kind of footage is too important to disappear into the digital ether of either the Web or even the PC; nor should it be vulnerable to the vagaries of electronic gadgetry that might break at any time. These materials need making real and making safe.

| <i>Lightweight (Spontaneous)</i>  | <i>Heavyweight (Intended)</i>  |
|---|--|
| Mobile Phones   | Camcorders   |
| Multi-function  | Dedicated to video   |
| Ad hoc, spontaneous capture   | Intended capture   |
| These devices are 'end use' devices, the majority of the video activity ends with the device, it can be created, stored and consumed within the device. | These devices are more focused on the capture element of the cycle, they are temporary holders of the video data.                  |
| Small clips are captured, but are ever increasing in number and size.   | Relatively limited use tied to the families' lifecycle. Although clips are longer, collection rarely becomes unmanageable in size. |
| Easy to upload (small file size – USB transfer)   | Upload barriers (large files – firewire transfer – real time)  |
| Less emphasis on tangibility of end result  | Importance of tangibility of end result  |
| Focus on sharing  | Focus on creativity  |
| Users do not want to edit   | Users want to edit   |
| Sharing practices<br>- In the moment<br>- Face-to-face<br>- Small clips<br>- Internet   | Sharing practices<br>- Giving DVDs<br>- Making gifts<br>- Watching with family<br>- Edited movies                                  |

**Table 1. Comparison of lightweight and heavyweight videowork.**

While we noted at the outset a dearth of research on video use, there is, interestingly, relevant literature on the use of mobile camera-phones which confirms our findings about lightweight videowing. If we consider the work of Kindberg et al [21], for instance, we can see similarities between the social uses of mobile phone based photos and the social uses of mobile phone based videos. These kinds of materials are often used to 'enhance the moment' and are certainly rarely planned. Consequently the immediacy with which footage is generated and consumed means that editing of the footage is largely considered inappropriate and outside of the bounds of the normal cycle of use.

What we have also found in this study that digital cameras constitute an interesting sub-group of technologies that exist somewhere between lightweight and heavyweight capture. There does not appear to be much prior research on this. Our findings suggest that whilst in principal retaining many of the features of lightweight technologies, when it come to the context of use, digital cameras are often used in far less ad hoc ways; indeed the types of footage recorded bear similarity to the footage recorded with the heavyweight devices, demonstrated by a preponderance of family footage.

The reason why digital cameras might be replacing other heavyweight technologies might point toward some important limitations in existing video technologies. Most obviously cost might be an issue: digital cameras are often less expensive than video cameras. They are smaller, lighter and hence less burdensome too. Perhaps more interestingly in terms of design, they appear simpler and more robust, and thus to users may perceive them as a 'safer bet', not being likely to break down at that special moment. Part of the appeal here might also be to do with how their relative simplicity can make them easier for users to use. They may do less, but at least the user knows what they are doing. This in turn might reduce the intrusion that often occurs when people have to capture video. These benefits have to be weighed against the quality of footage, of course. Heavyweight use is about important events, it should not be forgotten.

Irrespective of the shifting of the devices used to support heavyweight work, the practice itself will, it seems to us, continue, since there is no reason to suspect that the importance of events and thus the need to document them will diminish. But this ought not to distract from what we think is the failing of current attempts to support this kind of videowork, particularly the editing stage, issues to which we now turn.

### **Understanding User Creativity**

Prior studies seem to have misconstrued the role of the user in the editing process, the ways in which they want to edit their footage, and their editing priorities. Whilst many existing studies have correctly presumed that the time required to create edited video is a barrier to editing, they have failed to consider that a critical part of the editing process is the generation of some tangible media object, such as a DVD. This is also a creative process in which the user is often happy to engage.

There are many reasons why the creator of an edited video might want to produce a DVD and it is certainly true with other media, such as audio, that having a tangible copy of an artifact is often seen as a desirable attribute, digital copies usually being regarded as inferior versions [22]. With home movies, the desire for a tangible product is perhaps less an issue to do with quality than it is to do with other key aspects of the editing process. Editing in order to produce a tangible object - a DVD - is a common end goal. Yet at the same time, the design of most software systems reveals that this end goal is rarely considered as an important part of the process users want to engage in.

As a further related point, looking at the nature of these edited DVDs, and the nature of much home footage, shows that it does not conform to the kinds of narrative structure called for by Adams et al [2]. This structure basically entails building a story-like construct, with a beginning, middle, and end. Rather, we found that edited DVDs were usually made up by gathering together disparate but interesting clips into an organized form. By organized we do not mean a narrative structure; we mean a method for accessing and navigating through them. Here we found that

people added menus and titles not to “story-tell”, but to identify the different points or places that a viewer could go to. It is in the authoring of a viewable, navigable record of these stories that people were seen to exercise their creative talents. Further, for those that engaged in this kind of editing of video, it was often a labour of love. It was work for others on a video compilation so that those others could relive events as they saw fit.

## CONCLUSION

Clearly videowork is a multifaceted process and this initial research can only scratch the surface layers of complexity that our interviews and analysis revealed. The remit of the study was, however, to redress some of the imbalances in the existing literature, shifting the focus on video from a techno-centric to a more user-centric approach. In doing so, we have not only described the cycle of that work, but identified two distinct kinds of videowork that get undertaken. This we hope opens up the way for thinking about various ways in which the design of the different kinds of capture devices, software tools, and methods of sharing may be better tailored to the needs and desires of its users.

## ACKNOWLEDGMENTS

We would like to thank our participants for giving us access to their time, homes, families and video archives and Amy Walters and the students of Rawlins Community College for their help with the focus group work.

## REFERENCES

1. Abowd, G. D., Gauger, M. and Lachenmann, A. The Family Video Archive: An annotation and browsing environment for home movies. *In Proc. MIR'03*, ACM Press (2003), 1-8.
2. Adams, B., Venkatesh, S. and Jain, R. IMCE: Integrated Media Creation Environment. *ACM Transactions on Multimedia Computing, Communications and Applications*. 1, 3 (2005), 211-247.
3. Beyer, H. & Holtzblatt, K. *Contextual Design: Defining Customer-Centered Systems*. Academic Press: Kaufmann Publishers, (1998).
4. Boreczky, J., Girgensohn, A., Golovchinsky, G. and Uchihashi, S. An Interactive Comic Book Presentation for Exploring Video. *In Proc. CHI 2000*, ACM Press (2000), 185-192.
5. Casares, J., Long, A. C., Myers, B., Bhatnagar, R., Stevens, S. M., Dabbish, L., Yocum, D. and Corbett, A. Simplifying Video Editing Using Metadata. *In Proc. DIS2002*, ACM Press (2002), 157-166.
6. Davis, M. Editing Out Video Editing. *IEEE Multimedia*, 10, 2 (2003), 2-12.
7. Davis, M. Garage Cinema and the Future of Media Technology. *Communications of the ACM*, 40, 2 (1997), 43-48.
8. Doherty, J., Wilcox, L. and Girgensohn, A. A Hitchcock Assisted Video Edited Night at the Opera. *In Proc. Multimedia'02*, ACM Press (2002), 660-661.
9. Gemmell, J., Bell, G. and Lueder, R. MyLifeBits: a personal database for everything. *Communications of the ACM*, 49, 1 (2006), 88-95.
10. Girgensohn, A., Bly, S., Shipman, F., Boreczky, J. and Wilcox, L. Home Video Editing Made Easy – Balancing Automation and User Control. *In Proc. INTERACT'01*, IOS Press, Tokyo, Japan, (2001), 464-471.
11. Girgensohn, A., Boreczky, J., Chiu, P., Doherty, J., Foote, J., Golovchinsky, G., Uchihashi, S. and Wilcox, L. A Semi-automatic Approach to Home Video Editing. *In Proc. UIST'00*, ACM Press (2000), 81-89.
12. Hua, X., Lu, L. and Zhang, H. AVE – Automated Home Video Editing. *In Proc. MM'03*, ACM Press (2003), 490-497.
13. Kirk, D. S., Sellen, A., Rother, C. and Wood, K. Understanding Photowork. *In Proc. CHI 2006*, ACM Press (2006), 761-770.
14. Pan, Z. and Ngo, C. Structuring Home Video by Snippet Detection and Pattern Parsing. *In Proc. MIR'04*, ACM Press (2004), 69-76.
15. Stevens, M. M., Abowd, G. D., Truong, K. N. and Vollmer, F. Getting into the Living Memory Box: Family archives & holistic design. *Personal and Ubiquitous Computing*, 7, (2003), 210-216.
16. Suchman, L. *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press: New York, (1987).
17. Van Dijk, J. From shoebox to performative agent: the computer as personal memory machine. *New Media & Society*, 7, 3 (2005), 311-332.
18. Yan, W. and Kankanhalli M. S. Detection and Removal of Lighting & Shaking Artifacts in Home Videos. *In Proc. Multimedia'02*, ACM Press (2002), 107-116.
19. Yip, S., Leu, E. and Howe, H. The Automatic Video Editor. *In Proc. MM'03*, ACM Press (2003), 596-597.
20. Zhu, C., Mei, T. and Hua, X. Natural Video Browsing. *In Proc. MM'05*, ACM Press (2005), 265-266.
21. Kindberg, T., Spasojevic, M., Fleck, R. and Sellen A. I Saw This and Thought of You: Some Social Uses of Camera Phones. *In Proc. CHI '05 Extended Abstracts*. ACM Press (2005), 1545-1548
22. Brown, B. and Sellen, A. Sharing and Listening to Music. In K. O'Hara and B. Brown (Eds.) *Consuming Music Together: Social and Collaborative Aspects of Music Consumption Technologies*. Springer: Netherlands, (2006) 37-56