How Bad is Good Enough? Exploring Mobile Video Quality Trade-offs for Bandwidth-Constrained Consumers

Anne Oeldorf-Hirsch
Pennsylvania State University
anneohirsch@gmail.com

Jonathan Donner
Microsoft Research India
jdonner@microsoft.com

Edward Cutrell
Microsoft Research India
cutrell@microsoft.com

ABSTRACT
In developing countries, many would-be mobile internet users perceive downloadable video content as too expensive. Aggressively degrading this video could reduce its file size and therefore its cost. The studies presented here explore extreme cases of this quality/cost trade-off for mobile phone users in urban India. A series of online studies tested the effects of manipulating a video's content, bit rate, frame rate, and audio quality on quality ratings and enjoyment. Results show that video quality and thus file size can be greatly reduced with relatively little decrease in these outcomes. A field experiment with low-income users in urban India explored consumers' choices when presented with a trade-off between video quantity and quality and found that nearly one-third selected a lower quality video for the benefit of more video content. Results suggest that offering lower-quality videos to bandwidth-constrained users could provide monetary savings with only minimal reduction in consumer satisfaction.

Author Keywords
Mobile, video quality, video degradation, emerging markets

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

General Terms
Experimentation.

INTRODUCTION
Thanks to the mobile (cellular) telephone, the world is in the midst of a historic transformation in connectivity. It is remarkable enough that in the past decade, billions of people, particularly in the developing world, secured the ability to make phone calls and exchange basic text messages (SMS) with each other [10]. But there is a second wave underway, as phones converge with computing, smartphone and data-enabled feature phones become less expensive, and access to the internet follows the same trajectory as access to the phone network. By 2015, there may by over 700 million people who access the internet exclusively via their mobile phone [5].

But the drive to widespread global wireless data use may not be as rapid or successful as the one we saw with basic telephony. In this paper, we explore how the inherent technical and pricing constraints around metered and constrained bandwidth may be reducing the appeal of wireless data use. In particular, we focus on the reluctance to consume (pay for) mobile video, which, regardless of whether it is downloaded or streamed, dwarfs the file sizes of email, text, music, or images.

The mobile boom in the developing world was propelled partly by pay-as-you-go or "prepay" pricing [2]. Operators discovered they could sell airtime in sachets or top-ups, via scratch cards or point-of-sale terminals in the tiniest of neighborhood corner stores. Operators avoided costly credit checks, electronic transfers, complex billing, and identity tracking. Consumers got to pay as much (or as little) as they could afford each day/week/month. The result has been an arms-length, low-balance, low-spend, high-churn, frequent-transfer relationship between operators and their customers. Industry estimates suggest 54% of users in developed economies are prepay, 83% in emerging economies [1]; India is overwhelmingly prepay, with over 90% of subscribers "paying as they go" there [20]. Airtime itself has become a currency, to be conserved through careful practices such as missed calls [6], or traded in a web of obligations and social norms [8,17].

If mobile data is to "boom" across the developing world, its users will pay for it the same way they have paid for airtime (minutes). Indeed the metered mindset is not limited to developing countries, as mobile operators worldwide replace unlimited data plans with "capped" monthly plans – partly in direct response to the proliferation of data-hungry smartphones on their networks. The economics of wireless data has birthed a growing cadre of consumption-sensitive users.

Consumption sensitivity is now and will continue to be particularly acute around mobile video. Cisco estimates that video was half of all mobile data traffic in 2011, and will be two-thirds of all mobile data traffic by 2015 [5]. And yet, many users won't download a movie to their handset, lest it "eat their cap" or "drain their airtime" [4]. But that does not mean that videos aren't *on* phones. Prosperous users might have access to a PC and a cheaper broadband connection at home and they simply shift content to their phones for later viewing via Wi-Fi or direct

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NordiCHI '12, October 14-17, 2012 Copenhagen, Denmark
Copyright © 2012 ACM 978-1-4503-1482-4/12/10... $15.00

connections. But for those users without easy access to PCs and cheap broadband, the current solution relies on sneaker-nets, download shops, SD card swapping, and other forms of side loading video content, often with little regard for copyright laws [18,21]. It is not uncommon to find individuals with multiple gigabytes of video content on a simple feature phone; but this content has been acquired haphazardly, in batches, with little ability for systematic search, archiving, socially-informed sharing, or other tools commonly seen in modern smartphones.

Our guiding research problem is that many would-be users perceive video content to be too expensive to download over their mobile wireless connection. However, video-viewing on mobile devices is greatly desired by these users. Those who do manage to pay for mobile internet in India cite video song and movie downloads as a key use of the service [16]. Assuming that the technical and pricing components were fixed, we elected to explore the nature of the content itself. By degrading the quality of a video, we could reduce its file size and therefore its cost to users. A first step in pursing this strategy is to ascertain "how bad is good enough?": how much could we degrade a video before it became undesirable? Making consumption-sensitivity more central to the design equation will be important for making digital experiences inclusive and affordable for the majority of the world's population.

This paper provides three primary contributions. First, we provide data from a series of experiments using online simulations to understand how viewers of mobile phone videos respond to degradation of video associated with reduced file sizes. Second, based on the data from the simulation study, we asked low-income users in urban India to make a trade-off between quality and quantity of video. One third of the people we spoke to in the field indicated that they would prefer the visibly degraded video if it meant receiving corresponding more content. Finally, we believe that this study helps to bring the notion of "consumption sensitivity" and issues related to metered access of bandwidth to the design community. For millions of people who spend a disproportionate amount of their income on airtime, these considerations are critical.

Prior Research on Video Quality

Previous studies have tested the effects of video quality on enjoyment, commonly varying factors such as resolution, video bit rate, audio bit rate [13], and frame rate [7], and have offered nuanced findings which suggest that more is not always better. As expected, satisfaction with video clips decreases as frame rate decreases, but this effect is not linear, and varies by the amount of movement present in the content [7]. However, decreases in frame rate are still largely acceptable as long as the quality of those frames is high [15]. Likewise, the improvement in satisfaction that comes from increasing video size also has a limit, increasing only to a point at which no additional benefits are gained by the user [12]. Audio quality, while perceived, may not contribute strongly to the overall quality of experience (QoE) [3]. Even higher audio quality is not uniformly better: lower quality audio was found to be more acceptable when video quality was also lower [12], and even preferred when frame rates are too low [14]. Notably, these effects are highly dependent on the content of the video (e.g., sports, news, or music).

These findings introduce a number of trade-offs in video watching. Available data bandwidth being equal, reduction in frame rate brings an improvement in frame quality, just as increasing a video's resolution causes the sharpness of the video to suffer. As video viewing moves to contexts susceptible to these trade-offs (e.g., small mobile devices, bandwidth limits for streamed data), models of QoE have accounted for the combined effect of bit rate, frame rate, screen size, and content [11]. Recent work has revealed specific thresholds at which mobile video quality becomes unacceptable [19], suggesting that for smartphone screen sizes video bit rate should be at minimum 120kbps to 500kpbs, depending on content type.

Present Research

The results of these previous studies have paved the way for testing just how far some of these quality indicators can be reduced in favor of other factors. However, these studies have not tested the threshold of quality for enjoyment by bandwidth-constrained users, such as those in the developing world who are using technology that pre-dates smartphones. While, for instance, [19] proposes thresholds for mobile video quality, these were tested on a minimum smartphone screen resolution of 320x240 with a minimum 12.5fps frame rate. The present research aims to test video quality below the previously established threshold in order to find just how much video can be degraded before it is rejected by price-sensitive users. Rather than establishing the best mobile video experience, this research aims to find the cheapest acceptable option for those consumers who are price-sensitive. The following studies seek to answer the following questions: How low can video be degraded without loss in perceived quality and enjoyment? What trade-offs in quality will individuals make to save money/time/bandwidth or get more content? More specifically, given the choice of paying the same price for a given amount of high quality content or a greater amount of low quality content, which would price-conscious users choose?

To explore these conditions, two studies were conducted with mobile phone users in India, as described in the following sections. First, information about the current mobile media market was gathered by visiting local mobile shops. Then, an online study had participants provide quantitative ratings of several clips of varied quality to assess which factors primarily influence the viewing experience. Following this, a field study assessed what quality/quantity trade-offs low income individuals might make when presented with higher and lower quality videos.

Preliminary Interviews and Media Collection

To better understand the content available in the current mobile phone market in urban India, we made visits to three mobile download shops in the city to gather content and interview the shop owners about their content and their clients. These mobile download shops function as either mobile phone repair shops or other types of shops, selling songs, videos, and sometimes other content to customers on memory cards (SD cards) for use in their phones. All three sold content to customers who could either request a given amount of data (e.g., 2GB), or specific songs, video songs (music videos), or even full-length movies. The first shop was a small mobile repair and accessory store on a busy market street in one of the city's oldest neighborhoods, owned and operated by just one man, who was interviewed. The second shop was primarily a stationery and office supply store located in the same neighborhood, in a dense shopping area. A man and two women were working in the store, and information about the store was provided mainly by the male employee. The third shop was a mobile accessory and repair shop in a partially-enclosed shopping center in another historically popular shopping neighborhood, and was run by one man. His brother, who previously worked in that store and now ran his own elsewhere in the shopping center, was also present and answered many questions about the shop and the current mobile video market. In total, four individuals were interviewed.

Interviews

Shop owners varied in their experiences with and opinions about the current market, but three key themes emerged regarding their clientele: 1) users shy away from video purchases; 2) people are moving to content downloads; and 3) consumers' understanding of video quality is vague. First, customers tended to buy more music files than video files; songs account for as much as 80% of the media these shops sell. The primary reasons for this are that videos take up much more room on their phones than music files, and that they consume too much battery life because of the screen display. In turn, many consumers end up only listening to the audio from these videos (with the screen off). Second, all shops were experiencing a change in business with customers increasingly downloading content on demand rather than waiting for it at mobile shops. One shop owner stated that customers now are primarily kids and "locals," not those of college age who seem to have the financial and technical ability to turn to downloadable content. Another shop owner stated that only those that do not know how to download content straight to their phones still come to his shop for content, and that today only 10% of his business is from content sales; most earnings now come from repairs and accessories. Finally, when asked about what customers request in terms of quality and what the shops can provide, answers were inconsistent. Shop owners were in agreement that customers never ask about the quality of the video and certainly never request videos of any particular level of quality. However, two shops insisted that customers cared about and expected quality, and assume the videos they buy are of good quality. Only one shop owner edited the videos in any way, but only by converting them into different formats, not changing video qualities such as the resolution, or making edits to decrease file size.

Media content

We purchased six GB of video content for 350 Rupees total (about $7). This included 408 videos, mostly video songs, but also some comedy clips, ranging in length from 49 seconds to 13 minutes, plus one 3-hour movie file. Most videos (66%) were in 3gp format, 29% were in mp4 format, and the rest were in avi format. The video bit rate ranged widely from 40 kilobytes per second (kbps) to over 2,000 kbps, but the median bit rate for a standard feature phone resolution of 176 x 144 pixels was 120 kbps, and the median bit rate for a standard smartphone resolution of 320 x 240 pixels was 464 kbps. The median frame rate was 15 frames per second (fps), although 12 fps was the most common. Analysis of this content informed the encoding details of videos for the following study.

EXPLORING USER RATINGS OF DEGRADED VIDEO

First, an online assessment with a set of manipulated videos was conducted to test how differences in various mobile video quality metrics are perceived. Participants were asked to view and rate video clips that had been edited by manipulating four video characteristics: content, video frame rate, video bit rate, and audio bit rate. The goal of this first experiment was to test whether such quality differences are even perceptible to users.

Stimulus

The videos used for the study were obtained from an Internet video site and were re-encoded with the desired properties and into embeddable FLV format using the AVS Video Editor program. Because perceptions of quality are known to vary across content types [4], three different videos were used: a news clip, a sports game clip, and a music video clip. Videos were chosen from content relevant to the Indian audience. The news clip came from a story aired by Al Jazeera about the Akash computer tablet developed in India, the sports clip showed a winning play from a local cricket match, and the music video clip was selected from a music video from a Hindi film that was released at the time.

Each of these videos was reduced to a 20-second clip and then edited to vary in quality by manipulating video bit rate, video frame rate, and audio quality. Video bit rate (kbps) is the amount of data contained within each second of video, which determines how pixelated the video looks, and is also the primary determinant of file size. Based on videos collected from the mobile shop, the average video bit rate was about 120 kbps. To test the effects of roughly doubling and halving this bit rate, videos were encoded at a low

video bit rate of 50 kbps, at a medium bit rate of 120 kbps, or a high bit rate of 250 kbps. These upper and lower values were selected to match the proportions of the frame rate values used. Video frame rate (fps) determines the visual smoothness of movement in a video. Videos were encoded at a low frame rate of 5 fps, a high frame rate of 25 fps, or at a midpoint of 12 fps. Thus a video with the highest video bit rate and the lowest frame rate would appear jagged in movement, but with sharp frames, whereas a high frame rate but low bit rate would show smoother movement, but with more blurred or pixelated frames.

Finally, because audio bit rate also affects the file size, and may have an effect on perceptions of quality and enjoyment, each video was encoded at low and high quality audio. A number of factors determine sound quality, including mono versus stereo, bit depth, bit rate, and sampling rate, but a test of various encoding methods for these video clips indicated that varying sampling rate provided the greatest difference in file size. Thus, videos were encoded in 16-bit stereo with a bit rate of 64 kbps at either 11,025 Hz or 44,100 Hz.

The four variables were fully crossed such that a video was created for each possible combination of factors (e.g., a news video at 5 fps, at 120 kbps, and at 44,100 Hz audio quality). This created 54 different videos, and thus 54 different conditions, resulting from the factorial combination of three types of content (news, music, or sports), three video bit rates (50 kbps, 120 kbps, or 250 kbps), three frame rates (5 fps, 12 fps, or 25 fps), and two levels of audio quality (11,025 Hz or 44,100 Hz). The file sizes ranged from .25MB to 1.02MB, with the lowest bit rate videos being on average 45% of the size of the highest bit rate videos.

Procedure

Amazon's Mechanical Turk [22] was used for the experiment. This service allows for data collection from large samples with quick turnaround and for little cost. The sample parameters were set to include only those individuals residing in India and the payment amount was $.05 per task (about 2.5 Indian Rupees at the time of data collection). The desired number of participants in each condition was set to 25.

Eligible participants on Mechanical Turk could to see the listing for this study in their list of available "Human Intelligence Tasks" or HITs. A HIT is a task, such as an online survey, that a Turk Requester can make available for Turk Workers to complete for payment. Workers looking for tasks to complete can browse or search all available HITs and choose to accept any HIT for which they are qualified. Upon completing a HIT, the Requester of that HIT approves (and pays) or rejects the Worker's submission. This study was set up such that each video created for the study produced a separate HIT, so once a participant chose to accept the HIT, the website randomly loaded one of the conditions created for the study. Upon completing a HIT, the participant could choose to load another HIT (in this case another condition containing another version of the video), and could potentially complete as many as all available conditions in the study, but they could not receive the same video more than once.[1] Order of videos was randomized. In an attempt to limit the number of participants who would watch and rate all 54 videos—the Turk interface does not allow for limiting users to only some conditions within one HIT—three separate listings were created: one for each type of video content (news, sports, and music).

To complete the HIT, participants were asked to view a 20-second video clip shown on the website as though it were being displayed on a mobile phone (Figure 1). After watching the video they rated it on overall quality, video quality, and audio quality on 5-point scales (Bad, Poor, Fair, Good, or Excellent) recommended for use in video ratings, and found to be comparable to other common ratings used [9,23]. They also rated their enjoyment of the video on a 5-point scale ranging from "Did not enjoy at all" to "Enjoyed very much."

Figure 1. Simulated feature phone and smartphone screens presenting video stimuli.

Phase 1: Feature phones

The first experiment tested how the videos may be perceived on a feature phone (Figure 1, left). Feature phones are differentiated from smartphones by their lack of a third-party operating system (such as Apple's iOS or Google's Android), a system for installing applications, and a reliance on mobile data or Wi-Fi, among other advanced features, but still offer some features, such as a color screen, games, and email, over the most basic mobile phones. The feature phone is still the most common phone format used in India. The videos for this study were encoded at a resolution of 176 x 144 pixels and were displayed on an image of a common style of feature phone with a 128 x 160 pixel screen. Of videos gathered from the mobile shops, those in 3gp format (intended for feature phones) most commonly had a 176 x 144 resolution, and of

[1] Due to this design limitation, analyses were run with number of videos viewed included as a control variable.

current feature phones on the market in India, 128 x 160 was the most common screen size. The phone on which the video appeared to be displayed was a popular style of phone from which the name brand was removed.

Data was collected from 25 participants for each of the 54 video conditions, resulting in 1350 data points. Ninety-two individuals participated, completing anywhere from one to all 54 conditions (i.e., watching up to all 54 of the videos). No personal information (e.g., age or gender) was collected about participants.

Of primary interest was the effect of video bit rate, the main determinant of file size, on perceptions of overall quality and ratings of enjoyment. A mixed-model ANOVA was run to assess this and following effects. Video bit rate had a significant effect on perception of overall quality, $F(2,51) = 3.76$, $p < .001$, such that the low bit rate (50 kbps) was rated of significantly lower quality ($M = 3.56$, $SD = .94$) than the medium (120 kbps) and high (250 kbps) bit rate conditions. However, quality ratings for the medium ($M = 3.79$, $SD = .86$) and high ($M = 3.77$, $SD = .90$) bit rates did not differ significantly, indicating that improvements in video quality do not impact quality perceptions beyond a bit rate of 120 kbps. See Figure 2.

Video bit rate had the same effect on ratings of enjoyment, with the low bit rate video being rated as significantly less enjoyable ($M = 3.53$, $SD = 1.10$) than the other bitrates, $F(2,51) = 7.63$, $p < .001$. Again, the medium ($M = 3.77$, $SD = 1.01$) and high ($M = 3.75$, $SD = 1.06$) bit rate videos did not differ significantly on enjoyment. As shown in Figure 3, enjoyment of the videos, overall, drops off at a bit rate of 120 kbps. These effects are moderated by content, such that enjoyment still increases somewhat for the music video as the bit rate increases, and actually decreases for the sports video as bit rate increases. These effects are likely due to the amount of movement in the video which was highest in the music video and lower in the news and sports videos.

Frame rate showed a similar pattern, with the low frame rate (5fps) videos being rated of significantly lower quality ($M = 3.42$, $SD = 1.03$) than the medium (12 fps) and high (25 fps) frame rate videos, $F(2,51) = 36.29$, $p < .001$. Also, the low frame rate videos were significantly less enjoyable than the medium or high frame rate videos. However, the medium and high frame rate videos did not differ in either case, indicating that videos with frame rates higher than 12 fps did not additionally improve perceptions of the video's quality or enjoyment of the video.

Audio quality had no significant effects on ratings of overall quality or enjoyment, but the high quality videos were perceived to be of significantly higher audio quality than those of low quality, $F(1,52) = 14.17$, $p < .001$, indicating that participants did notice the difference, even though this variable alone impacted neither their perceptions of the video's overall quality nor their enjoyment of the video. These results match previous findings that audio effects do not significantly impact overall video quality perceptions [3].

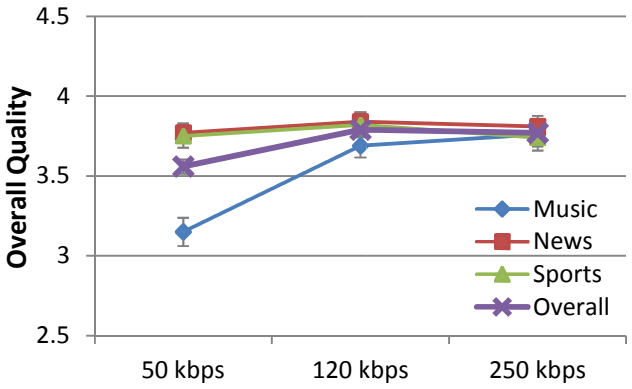

Figure 2. Overall quality ratings for video content by bit rate for feature phones.

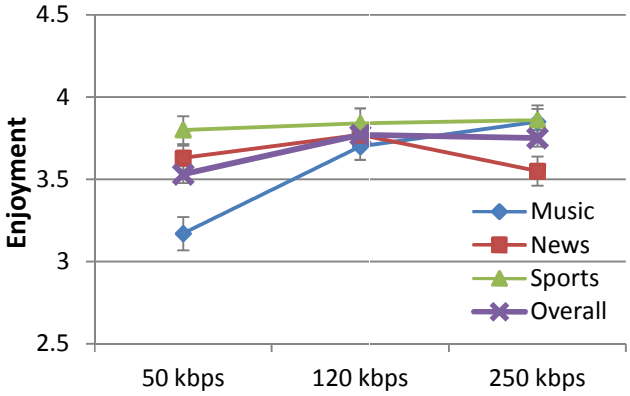

Figure 3. Enjoyment ratings for video content by bit rate for feature phones.

Figure 4 depicts how the two main factors that determine a video's file size (video bit rate and audio quality) impact ratings of video enjoyment. The peak in the figure displays which bit rate and audio quality combination results in the highest enjoyment. According to this graph, the most enjoyable video is one with high quality audio, yet only medium (120 kbps) quality video. No benefit was found for providing a bit rate higher than the median of 120 kbps on a feature phone.

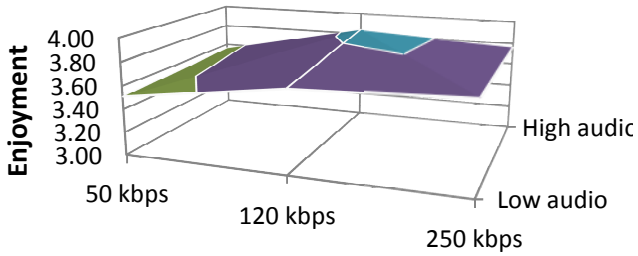

Figure 4. Enjoyment ratings by video bit rate and audio quality.

Phase 2: Feature phone follow-up

To assess in greater detail where the change in quality perceptions lies between 50 kbps and 120 kbps, a follow-up study was done with the same method, but this time testing a closer range of video bit rates: 85 kbps, 120kbps, and 185 kbps. Feature phone videos in the mobile shop video collection in 3gp format had a median video bit rate of about 120 kbps, and thus 50 kbps and 250 kbps represented much lower and much higher quality than the standard. The video bit rates of 85 kbps and 185 kbps were chosen as midpoints between 50 kbps and 120 kbps, and between 120 kbps and 250 kbps, respectively. The experiment again tested the same three types of video content, three frame rates, two audio quality levels, and these three video bit rates of 85 kbps, 120 kbps, and 185 kbps. This again resulted in 54 videos and thus 54 conditions.

The study was listed on Mechanical Turk as previously, with 25 participants rating each of the 54 videos, resulting in 1350 data points. This time, 125 individuals participated, again rating anywhere from one to all 54 videos. ANOVA was used to test for differences between bit rates on outcomes. In this case there were no significant differences between the bit rates in terms perceptions of overall quality, $F(2,51) = 1.22$, $p = .30$. Likewise, there were no significant differences in enjoyment, $F(2,51) = .03$, $p = .97$.

Combining results from all five bit rates tested revealed that only the lowest bit rate (50 kbps) differs significantly from all higher bit rates in terms of overall quality perceptions, $F(4, 85) = 7.82$, $p < .001$, and in terms of enjoyment, $F(4, 85) = 8.49$, $p < .001$. That is, improving the video quality by increasing the bit rate to 85 kbps led to significant increases in perceptions of quality and enjoyment, but additional increases in quality did not further increase these outcomes. In other words, the file size of a feature phone video was reduced by 30% in reducing the quality from 120 kbps to 85 kbps, with no significant loss in quality perception or enjoyment. Again, these results vary somewhat by type of content, as depicted in Figures 5 and 6.

Phase 3: Smartphones

A final Turk study was run to assess how these results may differ when displayed on a smartphone screen (Figure 1, right) with a higher resolution where higher video quality may be expected. The same 20-second news, music, and sports clips were used, but this time encoded at a 320 x 240 resolution to match that of the average smartphone on the market in India. Additionally, this was the most common resolution for mp4 videos (intended for smartphones rather than feature phones) in the videos gathered from mobile shops. The previous five video bit rates were tested, plus an additional bit rate of 450 kbps to reflect the median bit rate of 464 kbps of the mp4 videos gathered from the mobile shops. Videos were again tested at frame rates of 5 fps, 12 fps, and 25 fps, and at low and high audio quality. The combination of each of these factors resulted in 108 different videos, and thus 108 conditions. Files sizes ranged

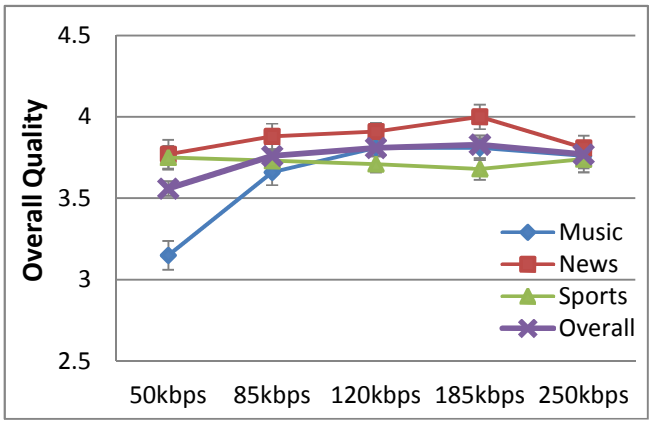

Figure 5. Overall quality ratings for video content for all bit rates tested for feature phones.

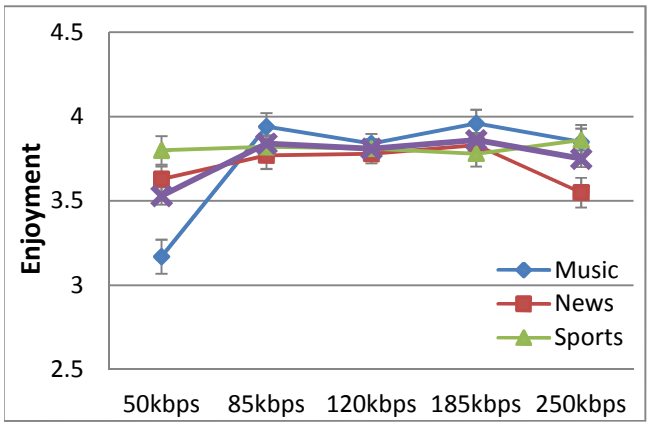

Figure 6. Enjoyment ratings for video content for all bit rates tested for feature phones.

from .25MB to 1.54MB, with the lowest bit rate video file size being on average 29% of the highest bit rate video.

Again, 25 participants completed each rating task, this time for 108 conditions, resulting in 2700 data points. 114 individuals participated, rating anywhere from 1 to all 108 videos. ANOVA was used to test for effects of video attributes on outcomes.

Video bit rate again had a significant effect on overall quality perceptions, $F(5,102) = 22.10$, $p < .001$, with 120 kbps remaining the breaking point. The lowest bit rates of 50 kbps and 85 kbps were rated as significantly lower quality than those of 185 kbps and higher, but 120 kbps differs only from the lowest and the highest bit rates, but not those of 85, 185, or 250 (see Figure 7).

Likewise, video bit rate had a significant effect on enjoyment, $F(5,102) = 12.54$, $p < .001$. Statistically, enjoyment did not increase significantly with bit rates higher than 120 kbps, but this varied by content. As depicted in Figure 8, enjoyment for all content types does not noticeably increase with videos with a bit rate higher than 185 kbps. Despite viewing the video on a larger screen with a higher resolution than on the feature phone screen, higher bit rates were not necessary for producing greater

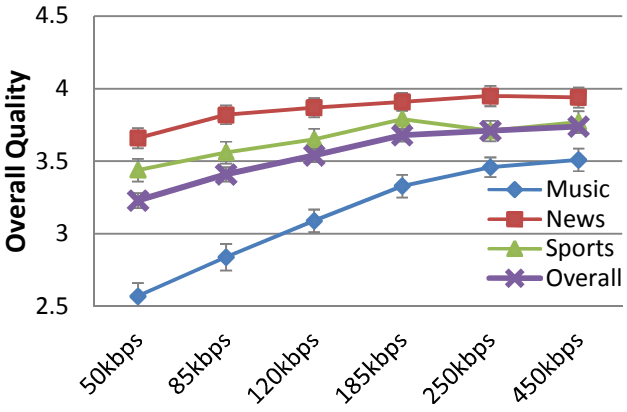

Figure 7. Overall quality ratings for video content by bit rate for smartphones.

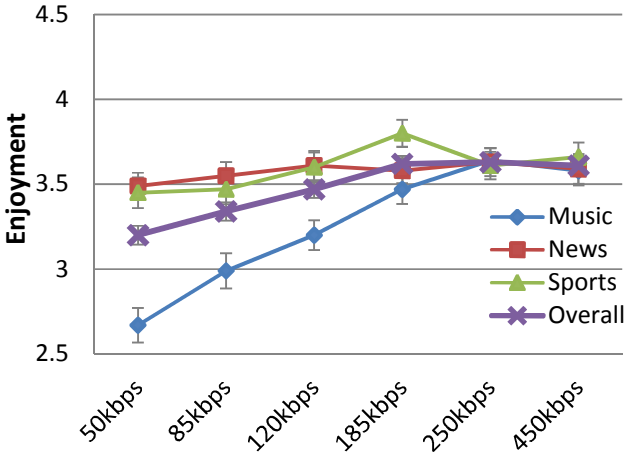

Figure 8. Enjoyment ratings for video content by video bit rate for feature phones.

enjoyment. Thus, video file size for smartphone screens can be reduced by 48% (from 1.38MB at 464 kbps to .72MB at s185 kbps) without a marked decrease in enjoyment.

Summary

These studies show that a 47% reduction in file size from 450 kbps to 185 kbps (for the smartphone) and 15% reduction in size from 120 kbps to 85 kbps (for the feature phone) is possible without presenting a perceptible difference in quality (see Figure 9), and thus without affecting enjoyment of the video. Based on the videos used here, for feature phones, a three-minute music video would "cost" 5.4mb at 120 kbps (20sec encodes at .6mb) or 4.6mb at 85 kbps. At the rough level of 1.2 Rupees for 1 MB of data [20], this would save the user about 1 Rupee per feature phone music video downloaded while maintaining a desirable user experience.

FIELD EXPERIMENT

Our studies on Turk begin to map the ways in which users' subjective enjoyment is related to various quality manipulations of video (which are directly related to

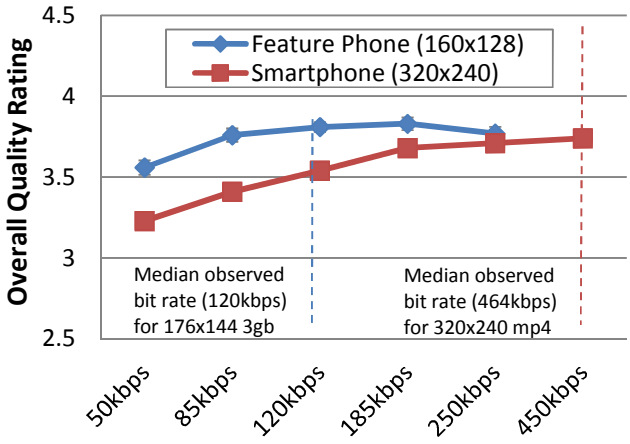

Figure 9. Overall quality rating by bit rate for feature phones and smartphones.

bandwidth consumed). While this is somewhat interesting in and of itself, our real interest is whether and how mobile users would view trade-offs between video quality and cost. To explore this, we conducted a field experiment to assess how low income mobile users would make such trade-offs. That is, if users had the choice between a given amount of high quality content or a larger amount of low quality content for the same price, which would they choose?

It is important to note that our target population spends USD2-3 a month on airtime, but this is also a significant expense for them. Identifying and consulting them proved challenging in various ways and required a new set of methods and considerable flexibility. As a result, our experiment has some flaws, but we believe that the benefit of performing the study with this population provides a novelty and realism that could not otherwise be obtained.

Method

Participants were recruited from low-income communities which are likely to be representative of "download shop" patrons. Data was collected from individuals living in a local slum neighborhood and from service staff at a local organization, totaling 38 participants. The sample was 76% male and ranged in age 17-39 years old (M = 24.74, SD = 5.90). Only two participants did not have their own mobile phone, and used a family member's instead. Ninety-two percent of participants watch videos on their phone and 89% share videos with others, most commonly through Bluetooth transfers.

The experiments took place in the two environments from which participants were recruited. Individuals from the slum were recruited though a local research contact from the slum. Community members who were interested were asked to show up at the contact's home on the day of the study. One researcher conducted the study in the contact family's home with the use of a translator. Because of the small space for conducting this research, the family of the house and many of the neighborhood participants were

present throughout the study, but were instructed not to view the videos as others were participating. The service staff participants were recruited through a supervisor at their place of employment. Those interested in participating met the researcher in a common area of their workplace at the instructed time. The same researcher conducted the study with translation to the local language by the supervisor. Due to the work area in which staff came and went as they completed their work, several employees were present during the experiments, but again were asked not to view the videos before their own participation or disrupt the current participant.

Each participant was randomly assigned to one of three conditions of video viewing: low/medium, low/high, or medium/high. They were shown two sets of four videos on a mobile feature phone and asked to rate each of them on the same scales used previously in the Turk study (overall quality, video quality, audio quality, and enjoyment). The sets of videos included the three 20-second video clips used in the Turk study (music, news, and sports clip), plus a 20-second clip of a popular video of a man dancing in various countries set to music. The two sets of videos that participants watched were identical in content, but differed in quality such that they first watched the four videos at low, medium, or high quality, and then the second set at one of the other levels of quality. Order of video quality and order of videos within video sets were randomized to control for order effects.

All videos were encoded at the feature phone resolution of 176 x 144 and varied between quality levels primarily on video bit rate. The videos in the high quality condition were encoded at bit rate of 250 kbps, 25 fps frame rate, and high quality audio. The medium quality videos were encoded at 120 kbps, 12 fps, and low quality audio. The low quality videos were encoded at 50 kbps, 12 fps, and low quality audio. These values were informed by results of the Turk studies.

Each video was started the researcher and the phone was handed to the participant to view as he or she would typically view a video. After viewing each video, the researcher verbally asked the participant each of the rating questions and gave the five possible choices (Bad, poor, fair, good, or excellent). These questions were translated verbally to the local language by the translator and answers from the participant translated verbally back to the researcher and recorded on paper.

After completing the video task and answering additional questions about their mobile phone use, participants were thanked for their participation and offered a 4GB SD card loaded with video content as compensation. Each card had 14.36MB of the "dancing man" video on it, varying inversely by length and quality. The cards contained eight minutes of the high quality video, 11 minutes of the medium quality video, or 15 minutes of the low quality video. Participants were asked to make the choice between the two quality levels of video offered based on their condition. For instance, participants in the medium/high condition, who watched one set of videos at medium quality and one set of videos at high quality, were asked to choose between the medium and high quality SD card options. The word "medium" was not used in the study; rather, participants were always asked to choose between the "lower" or the "higher" quality of video *that they saw*.

Due to logistical issues that limited the sample size, the low/medium condition was dropped early in the study to focus on obtaining adequate participants for the medium/high and low/high conditions. These two conditions were of greater interest as they contained the high quality condition, testing what choices participants make when they have seen the best possible option compared to a video that is somewhat worse in quality or one that is considerably worse in quality, rather than testing how participants decide between two sub-par conditions.

Results

Five of the participants were in the low/medium condition so were dropped from further analysis. Of the remaining 33, 16 were in the low/high condition and 17 were in the medium/high condition. A contingency analysis indicates that the lower quality video was selected by 38% of those who compared the high quality to the low quality videos, and by 24% of those who compared the high quality to the medium quality videos. The proportion of participants who selected the lower versus higher quality video did not differ significantly by which comparison they were making, $\chi^2(1, N = 33) = 0.76$, $p = .38$. Of all participants, nearly one-third of participants selected the lower quality video, even when given the option for higher quality, indicating that lower quality video can a desirable choice when it means several more minutes of content.

Summary

The results from the field experiment reveal an important idea regarding how people judge mobile video quality. Participants were not able to consistently judge the differences between the low, medium, and high quality videos they watched. This is evident in the data, which shows that there were no significant differences between the low, medium, and high quality videos in ratings of overall quality, video quality, audio quality, or enjoyment. Comments made by participants indicated that they saw more quality differences in comparing the various content. This is corroborated by the ratings, which show that regardless of quality, the dancing video was rated significantly higher in overall quality than the other three videos, $F(3,290) = 63.95$, $p < .001$, and the sports clip was rated as significantly more enjoyable than the other clips, $F(3,290) = 3.46$, $p < .05$. Thus, even quantifiable differences in quality are influenced by other factors such as personal preferences for given content.

As a result, participants had difficulty in deciding which SD card to take, likely because of the lack of perceptual

difference in the videos they had just watched. Because judgments about better and worse quality varied by content rather than by actual video quality, participants were likely considering their preferences for the *content* they had just seen rather than the objective *quality* of the video sets. Of those who selected the higher quality video card, a few stated that they wanted it for its higher quality, but most did not have a clear explanation. Those who selected the lower quality card provided similarly inconsistent answers, although a few participants stated they wanted more of the video. Some thought that the SD card itself had more space in one condition over another.

There were some logistical issues with this field study. Participants and the researcher did not speak the same language, requiring the use of a translator. Despite a written script, it is possible that the translator was inconsistent in translating, or that some communication lost its nuanced meaning in translation. Also, it was difficult to isolate the current participant from the rest of the group in the study environment, leading many participants to hear and see others taking part in the study before participating. Additionally, those who had already completed the study sometimes attempted to influence the current participant. This experience highlights the difficulty of running such field experiments, but also serves as a reminder of the type of environment in which mobile video use occurs. Just as participants could not feasibly complete the study in isolation, they do not view videos or make decisions about entertainment in isolation. This is evident by the popularity of sharing content via Bluetooth and also physically sharing one's screen in a group viewing session.

CONCLUSIONS

The present research reveals several important findings about the acceptance of video quality by price-sensitive mobile consumers in India. The results found from the Turk studies regarding quality and enjoyment ratings indicate that mobile users, at least in this market, may be just as happy with mobile video offered at a lower quality than presently available. Virtually no decline in quality rankings was found even when video file sizes were reduced below the median by 15% for feature phones and 40% below the median for smartphones. This reduction in file size directly represents considerable potential savings in bandwidth and money for price sensitive users, such as those who pay for data by the MB. Furthermore, 30% of those in the field experiment selected the SD card with the *lower* quality video even though a higher quality version was available. These participants chose more content at the cost of quality, just as mobile phone users may do given the same option for downloaded content.

The present research is only a first step in exploring the potential desirability of degraded mobile video content for bandwidth constrained/price-sensitive users, and leads the way for new paths of future research. Future work needs to assess actual price tradeoffs through other methods such as forced choice simulations or actual observations on a download site. Such a site could provide cheaper but lower quality video alternatives to the standard video and then track users who choose that option.

Evidence from the Turk studies suggests that downloads for low-resolution feature phone consumers can be smaller and that current video providers may be over-delivering. While the industry is making strides in optimizing download sizes for mobile web – such as the Opera Mini browser which allows for the compression of text and images on web pages – the comparatively large size of video files remains a major strain on available data. For a given provider, different kinds of lower quality video might be compelling. Such lower quality video options may be particularly useful, for example, for public health campaigns, mobile learning videos, or for producers of other such "free" content that seek a wide audience and take-up by consumption-sensitive users. That is, even such "free" content is not really free if it works against one's airtime balance, any more than it is "free" to call a long distance telephone number and listen to a recording of a public service announcement. An option of a smaller size of such videos might be valued by a significant proportion of resource-constrained users.

The Turk study does have some limitations: notably, the setting is not necessarily realistic of typical mobile video viewing experience. While the environment of each participant completing the online study is unknown, sitting at a computer to view the video is presumably quite different from the experience of viewing such a video on their mobile phone, possibly with distractions in public, with others, or while in motion. Additionally, there was no control over how many times participants viewed the video, potentially creating unequal experiences for participants. Thus, findings from this exploratory work should be considered as guidelines for continued research, which should test these factors on a mobile device in a field setting.

Also, the difficulty revealed in the field experiment with terms like "quality" (especially when differences are not very perceptible) means that the framing of whether a video is "high" or "low" quality may be just as important to the overall rating of enjoyment as the objective measures of video quality such as bit rates and frame rates. These effects are also highly dependent on preference for the content, which could influence perceptions of quality. Exploring these framing effects would be another fruitful path for further research in mobile video.

This paper makes a more general contribution to the research design community by bringing "consumption sensitivity" and awareness of metered and pre-pay use to mind. With video becoming a primary form of information distribution, it is useful to consider that for bandwidth- and cost-constrained users, such as those in the developing world, big video files can be problematic to deal with.

These are critical issues for inclusion and affordability of mobile digital experiences in the years to come.

REFERENCES

1. Ahonen, T. *Tomi Ahonen Almanac 2011: Mobile Telecoms Industry Review*. 2011.
2. Beaubrun, R. and Pierre, S. Technological developments and socio-economic issues of wireless mobile communications. *Telematics and Informatics 18*, (2001), 143-158.
3. Belmudez, B., Moeller, S., Lewcio, B., Raake, A., and Mehmood, A. Audio and video channel impact on perceived audio-visual quality in different interactive contexts. *2009 IEEE International Workshop on Multimedia Signal Processing*, (2009), 1-5.
4. Chetty, M., Banks, R., Brush, A.J.B., Donner, J., and Grinter, R.E. "You're Capped!" Understanding the Effects of Bandwidth Caps on Broadband Use in the Home. *Proceedings of CHI 2012, May 5-10, 2010, Austin, TX, USA.*, (2012).
5. Cisco. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010 – 2015*. San Jose, CA, 2011.
6. Donner, J. The Rules of Beeping: Exchanging Messages Via Intentional "Missed Calls" on Mobile Phones. *Journal of Computer-Mediated Communication 13*, 1 (2008), 1-22.
7. Ghinea, G. and Thomas, J.P. QoS impact on user perception and understanding of multimedia video clips. *Proceedings of the sixth ACM international conference on Multimedia - MULTIMEDIA '98*, (1998), 49-54.
8. Horst, H.A. *The Cell Phone: An Anthropology of Communication*. Berg, Oxford, 2006.
9. Huynh-Thu, Q., Garcia, M.-N., Speranza, F., Corriveau, P., and Raake, A. Study of Rating Scales for Subjective Quality Assessment of High-Definition Video. *IEEE Transactions on Broadcasting 57*, 1 (2011), 1-14.
10. ITU. Global ICT developments, 2001-2011. *Global Statistics*, 2011. http://www.itu.int/ITU-D/ict/statistics/material/excel/2011/Global_ICT_Dev_01-11.xls.
11. Joskowicz, J. and Ardao, J.C.L. Combining the effects of frame rate, bit rate, display size and video content in a parametric video quality model. *Proceedings of the 6th Latin America Networking Conference on - LANC '11*, (2011), 4.
12. Knoche, H., McCarthy, J.D., and Sasse, M.A. Can small be beautiful?: assessing image resolution requirements for mobile TV. *Proceedings of the 13th annual ACM international conference on Multimedia - MULTIMEDIA '05*, ACM Press (2005), 829.
13. Knoche, H., McCarthy, J.D., and Sasse, M.A. How low can you go? The effect of low resolutions on shot types in mobile TV. *Multimedia Tools and Applications 36*, 1-2 (2007), 145-166.
14. Knoche, H., de Meer, H., and Kirsh, D. Compensating for low frame rates. *CHI '05 extended abstracts on Human factors in computing systems - CHI '05*, ACM Press (2005), 1553.
15. McCarthy, J.D., Sasse, M.A., and Miras, D. Sharp or smooth? *Proceedings of the 2004 conference on Human factors in computing systems - CHI '04*, ACM Press (2004), 535-542.
16. Rangaswamy, N. and Cutrell, E. Anthropology, development and ICTs. *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development - ICTD '12*, ACM Press (2012), 85.
17. Sambasivan, N. and Cutrell, E. Understanding negotiation in airtime sharing in low-income microenterprises. *Proceedings of CHI 2012, May 5-10, 2010, Austin, TX, USA.*, (2012).
18. Smyth, T.N., Kumar, S., Medhi, I., and Toyama, K. Where there's a will there's a way. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, ACM Press (2010), 753.
19. Song, W., Tjondronegoro, D., and Docherty, M. Saving bitrate vs. pleasing users. *Proceedings of the 19th ACM international conference on Multimedia - MM '11*, ACM Press (2011), 403.
20. TRAI. *The Indian Telecom Services Performance Indicators July - September 2011*. New Delhi, 2012.
21. Wyche, S.P., Smyth, T.N., Chetty, M., Aoki, P.M., and Grinter, R.E. Deliberate Interactions: Characterizing Technology Use in Nairobi, Kenya. *Proceedings of CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.*, ACM (2010).
22. Amazon Mechanical Turk. http://www.mturk.com.
23. *RECOMMENDATION ITU-R BT. 500-11 Methodology for the subjective assessment of the quality of television pictures*. 1996.