The Head or the Heart?
Measuring the Impact of Media Quality

Gillian M. Wilson
Department of Computer Science
University College London
Gower Street, London, WC1E 6BT
g.wilson@cs.ucl.ac.uk

M. Angela Sasse
Department of Computer Science
University College London
Gower Street, London, WC1E 6BT
a.sasse@cs.ucl.ac.uk

ABSTRACT
The number of multimedia applications is constantly increasing. Subjective methods are typically used to determine the level of media quality required in applications, yet recent findings have shown that these have limitations. This paper introduces an objective method for assessing media quality - measuring physiological indicators of stress. An experiment examining the impact of video frame rate is presented. With low frame rates, physiological measurements indicated that users were under strain, even though subjectively most reported no differences between low and high frame rates. We conclude that the evaluation of media quality should not be conducted using solely subjective methods.

Keywords
Evaluation methods, physiological measurements, user cost, video frame rate, videoconferencing.

INTRODUCTION
Videoconferencing has many potential applications in areas such as distance learning and health care. Computer workstations and high-bandwidth networks can deliver high quality audio and video, but increased quality inevitably means higher financial cost. Therefore, effective evaluation methods are vital to determine the quality users need to successfully perform tasks in videoconferences.

Currently, subjective methods are widely used to evaluate multimedia quality. Recent research has raised concerns about the validity of using these methods in isolation, due to problems with continuous assessment scales and post-hoc assessment [6]. Also, subjective assessment is cognitively mediated, e.g. a recent study found that users accepted significantly lower media quality when it had a notion of financial cost attached [2]. The quality accepted by users was below the threshold generally thought to be required for that task. Therefore, can users always accurately assess the impact media quality has on their performance and well being?

NOVEL ASSESSMENT METHOD
This research is employing objective methods to measure the impact of media quality on users. Physiological signals, widely accepted as indicating stress, are being taken to measure increased *user cost*. The responses being measured are Galvanic Skin Resistance (GSR), Heart Rate (HR) and Blood Volume Pulse (BVP). Under stress GSR and HR increase, whereas BVP decreases.

Previous research using subjective assessment methods [1] found that users did not report the difference between 12 and 25 frames per second (fps) when involved in an engaging task. If users do not notice the difference in frame rate, can it be assumed that it has no effect on them?

EXPERIMENT
Twenty-four volunteers participated in an experiment comparing two frame rates, 5fps and 25fps (increasing the quality difference used in [1]). They watched two recorded interviews conducted using IP videoconferencing tools on a high-quality computer screen. The interviews were between a university admissions tutor and school pupils applying to UCL. The tutor and students were playing themselves in scripted interviews, which had been designed with help of an admissions tutor to reflect common questions and interactions. The interviews lasted fifteen minutes each. Participants saw two interviews at 5-25-5fps or 25-5-25fps and were asked to make a judgement on the suitability of the candidates. The frame rate changed twice to counteract any expectancy effect. Audio quality was good and did not change.

Participants were asked to rate the video quality continuously using a software slider [3]. After both interviews, a questionnaire was given. This addressed how participants felt during the experiment and their opinions on the quality. Physiological measurements were taken throughout the experiment. We posited the following hypotheses:

1. There will be different physiological responses to the two frame rates: 5 fps will cause more stress.
2. Participants will not register the frame rate change subjectively.

RESULTS

Table 1 shows the mean physiological responses for each participant during 5fps and 25fps periods. T-tests were performed on the physiological data. These showed statistically significant increases in stress at 5 fps for all signals (GSR, p=0.002; HR, p=0.003; BVP, p=0.04). There was no significant correlation between subjective and physiological results. Additionally, questionnaire analyses showed that only 16% of participants noticed the frame rate change. These results indicate that participants did not report the change during or after the interviews, however they did respond to it physiologically.

	GSR		HR		BVP	
Subject	5fps	25fps	5fps	25fps	5fps	25fps
1	11.81	11.4	94.09	93.25	25.39	25.46
2	12.05	12.49	74.88	74.15	25.36	25.5
3	8.38	7.78	106.36	105.58	25.66	25.76
4	6.39	6.32	77.88	73.9	25.22	25.278
5	3.78	3.39	73.68	72.17	25.7	25.71
6	2.39	1.25	75.32	75.73	23.7	23.7
7	18.02	18.3	79.58	78.52	24.1	24.2
8	11.89	12.28	65.26	62.21	23.91	23.95
9	12.15	12.41	69.49	63.67	24.17	24.18
10	5.22	5.54	70.89	71.49	23.64	23.66
11	18.84	18.1	85.67	81.97	23.53	23.54
12	11.37	10.62	74.29	72.88	25.29	25.27
13	11.33	10.66	70.74	71.87	23.9	23.89
14	3.2	2.83	95.93	96.65	23.76	23.78
15	13.85	13.2	66.15	67.81	23.81	23.86
16	5.14	5.3	82.64	81.45	23.91	23.87
17	8.56	8.36	74.79	74.57	24.06	23.99
18	5.08	4.87	102.66	99.31	23.94	23.93
19	2.8	2.51	57.1	56.25	25.94	25.95
20	1.59	1.27	85.91	85.61	25.67	25.67
21	12.78	11.97	76.45	74.69	25.06	25.67
22	8.46	7.62	72.89	73.16	25.97	26.02
23	3.17	3.12	66.32	65.77	25.17	25.17
24	2.8	2.59	76.25	75.94	25.84	25.84

Table 1. Mean GSR, HR and BVP over 5fps and 25fps

DISCUSSION AND FUTURE WORK

Both hypotheses were supported. This is important as it indicates when people are engaged in a task, they report no difference between two frame rates, yet the difference is registered physiologically. The implication of these results is that the level of media quality required for a task should not be determined by subjective assessment alone. To identify acceptable levels of quality and those that result in unacceptable user cost, objective data needs to be collected. Traditional HCI approaches to evaluation recommend that task performance, user satisfaction and user cost should be considered; we feel that in the evaluation of media quality, this is particularly appropriate.

Our continuing work in this area will produce three substantive contributions. Firstly, the optimum and minimum levels of multimedia quality for users performing specific tasks will be determined. Secondly, the findings of this research will be incorporated into the ETNA project [4], which will produce a taxonomy of real-time multimedia tasks and applications and their corresponding audio/video quality requirements. Thirdly, a utility curve will be created. This will allow a multimedia application to receive feedback on the users' physiological state from a wearable computer, like those being developed at MIT [5]. The wearable computer will report the state of the user to the application and modify any parameters causing stress to maximize user satisfaction.

A final aim of the research described here is to bring to attention the drawbacks of utilizing solely subjective methods in the assessment of multimedia quality. We believe that physiological measurements can be used as part of a 3-tier approach to give a more reliable indicator of how users are affected by quality. The ability to detect discomfort and stress unconsciously has wider implications in areas like teaching, stress control and product assessment.

REFERENCES

1. Anderson, A.H., Smallwood, L., MacDonald, R., Mullin, J. and Fleming, A. (in press 1999/2000) Video data and video links in mediated communication: What do users value? To appear in *International Journal of Human Computer Studies*.
2. Bouch, A. and Sasse, M.A. (1999) It ain't what you charge it's the way that you do it: A user perspective of network QoS and pricing. *Presented at IFIP/IEEE International Synopsium on Integrated Network Management*, Boston, USA, 24-28 May 1999.
3. Bouch, A., Watson, A. and Sasse, M.A. (1998) QUASS – A tool for measuring the subjective quality of real-time multimedia audio and video. *Proceedings of HCI '98*, 1-4 September 1998, Sheffield, UK.
4. http:///wwwmice.cs.ucl.ac.uk/multimedia/projects/etna/
5. Picard, R.W. (1997) *Affective Computing*. M.I.T. Press, Cambridge, MA.
6. Watson, A. and Sasse, M.A. (1998) Measuring perceived quality of speech and video in multimedia conferencing applications. *Proceedings of ACM Multimedia '98*, 12-16 September '98, Bristol, England, pp55-60.

ACKNOWLEDGEMENTS

We gratefully acknowledge the contribution of Anna Bouch from UCL Computer Science. Gillian Wilson is funded through an EPSRC CASE studentship with BT Labs.