Behaviour & Information Technology, Vol. 25, No. 3, May–June 2006, 233–237

Taylor & Francis
Taylor & Francis Group

A study of the suitability of videophones for psychometric assessment

GEORGE DEMIRIS*†, DEBRA PARKER OLIVER‡ and KAREN COURTNEY†

†Health Management and Informatics, University of Missouri-Columbia, USA
‡School of Social Work, University of Missouri-Columbia, USA

In order to determine whether videophones are appropriate communication tools for psychometric assessments, we need to determine whether the quality of videophones is adequate to enable this type of assessment or whether it places a burden on the communication. The purpose of this study is to measure the subjective quality of video and audio features of commercially available videophones in the context of a psychometric assessment session. We recruited 52 subjects who used the videophone to participate in a psychometric assessment using the Perceived Stress Scale. After each session, participants filled out the ITU-T P.920 that assesses the context-specific quality of the video-call. Findings indicate that the overall audio and image quality of the video-call was satisfactory and participants perceived the videophones as useful in the context of psychometric assessment. These findings strengthen the call for use of video mediated communication in home and hospice settings and disease management.

1. Introduction

Telemedicine is broadly defined as the use of telecommunication technologies to enable participants separated by geographic distance to interact with the goal to improve healthcare delivery and/or support medical education. Advances in telecommunication have created opportunities for the use of commercially available videoconferencing products such as videophones at the patient's home. Such systems have been developed both in home and hospice care (Dansky and Bowles 2002, Demiris *et al*. 2004, Young *et al*. 2004) and can operate over regular phone lines without the requirement of a costly modification of the residential infrastructure. The aim of such systems is the utilisation of the technology to conduct 'virtual visits' (Demiris *et al*. 2003), namely interactions between patients at home and healthcare providers at a medical site. There are obviously issues that cannot be easily addressed in a virtual visit that often take place in actual face-to-face visits, such as the inspection of the housing conditions, assistance with housework, etc. However, it is assumed that during a video-call a healthcare provider can interact with the patient and use standardised tools to interview the patient and assess physiological and/or psychological symptoms and overall wellbeing (Demiris *et al*. 2003).

The use of psychometric tools during a virtual visit allows for the healthcare provider to ask the same questions that he/she would have asked in a face-to-face interaction. The video-call could have added value when compared to a regular phone as only about 7% of the emotional meaning of a message is communicated through explicit verbal channels, while about 38% is communicated by paralanguage, and about 55% comes through nonverbal communication, which includes such things as gesture, posture, facial expression, etc. (Mehrabian 1972). Herein lies the area where the video component may enhance the communication, as visual information supports different non-verbal aspects of remote interpersonal communication. Specifically, it supports the transmission of several cues (Whittaker 1995), such as: (a) cognitive cues that indicate remote participants' understanding, such as head nodes and visual attention (Clark and Schaefer 1989, Clark and Brennan 1991); (b) turn-taking cues afforded by head turning, posture and eye gaze, which facilitate management

*Corresponding author. Email: DemirisG@health.missouri.edu

Behaviour & Information Technology
ISSN 0144-929X print/ISSN 1362-3001 online © 2006 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01449290500195585

of the interaction, such as smooth transitions when there are changes of speakers (Argylle *et al*. 1968, Duncan 1972); and (c) social or affective cues that reveal participants' emotional state or interpersonal attitudes which are manifested in facial expression, posture and eye gaze (Mehrabian 1971, Ekman and Friesen 1975, Reid 1977).

While this seems to be an efficient use of the technology, which could reduce travel costs and provide a cost-saving method of performing routine interviews, it has not been investigated whether video-based interaction does enhance or disrupt the participants' ability to clearly communicate problems and concerns and respond to questions as part of a psychometric assessment. As Zuiderent *et al*. state, 'what often remains obscure in telemedicine experiments is how the technological mediation that takes place actually alters practices and relationships' (Zuiderent *et al*. 2003). Issues of image and audio quality, resolution and synchrony need to be addressed in this context in order to determine the extent of audio and video degradation that would be tolerated without compromising the utility of the overall interaction. Bruce argues that for most uses which are made of facial information, dramatic reductions in spatial and pixel resolution of images of videophones can be tolerated, while temporal information is likely to be much more crucial for communication efficiency (Bruce 1996). In order to determine whether videophones are appropriate communication tools for psychometric assessments, we need to determine whether the quality of videophones is adequate to enable this type of assessment or whether it places a burden on the communication.

While there are international standards for network and vendor interoperability, there are no widely accepted standards for performance evaluation of videophone applications. There are both objective and subjective assessment tools that can be used for the evaluation of videophones. Objective assessment tools would include measurements of frame rate, linearity, latency, lip sync, video resolution, and image color clarity. However, these measurements do not refer to specific tasks or the purpose of usage. In other words, the image color clarity can be sufficient for a given task or type of interaction and insufficient or burdensome for another purpose. Thus, an assessment of the videophone quality needs to determine whether a given level of quality is suitable to support the user in carrying out a particular task in a specific context (such as administering psychometric instruments via videoconferencing) rather than whether the user notices a degradation in passive viewing. In other words, the assessment needs to take into account users' tasks and the context of the videophone use. This can be addressed by the subjective assessment of the videophone quality. In this context, the quality is assessed as perceived by the actual user performing tasks in the context of the intended use.

The purpose of this study is to measure the subjective quality of video and audio features of commercially available videophones in the context of a psychometric assessment session. We aimed to assess the suitability of videophones as a communication tool for psychometric assessments by measuring the perceived quality of the communication in the context of an interview session using a behavioural instrument.

2. Methods

We created a flier explaining the purposes of the study and asked individuals interested in participating to contact a member of the research team to schedule participation in one of the test sessions. Fliers were posted and distributed within academic units of the School of Medicine at the University of Missouri. Test sessions were scheduled with two subjects at a time. We placed two videophones in two partition-based offices (cubicles) at a distance from each other. The videophone model used for this study was the Vizufon GVP-1000F (C&S Technology Inc., Korea), which operates over regular phone lines and complies to the International Videophone Standard ITU-T H. 324. The display is an Active Matrix TFT LCD screen (4'' diagonal with a resolution of 480 x 234 pixels). Figure 1 displays such a videophone.

Each subject who volunteered to participate, met with a member of the research team who explained the purpose of the study and provided details on the specifics of the test session. Then, the subject was asked to sit in the office and use the videophone to interact with the other subject (at the remote cubicle). A member of the research team established a connection with the remote party and initiated the video-call to ensure that the connection was established.

The environment in which a videoconferencing system is used can vary greatly in the amount of background noise and issues impacting visual display. Noise in the background can impact comprehension. There are also distracting features that can impact the visual display such as motion in the background (which might require a higher frame rate (Apteker *et al*. 1995)), lighting conditions or placement of the speaker in front of a window. In order to standardise across all sessions and maximise ideal conditions for the video-calls, we installed the videophones in two areas maintaining the same lighting conditions and minimising any kind of background noise for all test sessions. In addition, a member of the research team made sure that both parties were seated so that their face could be seen on the remote screen. We controlled for such environmental factors; however, we did not instruct users on behaviour or movement that would improve the overall quality of the video-call (e.g., we did not instruct users to avoid rapid hand movements which causes a 'pixelation' of the video)

Figure 1. A display of the videophone used for the study

to ensure that the users would interact as freely and intuitively as possible.

Once the video connection was established and the members of the research team controlled for background noise and camera positioning, the subjects were asked to continue with the video-call. For each test session, one of the two subjects was randomly chosen to be the 'interviewer', namely the person asking the questions of the psychometric assessment tool and the other subject was the 'interviewee' responding to the questions. For the purposes of the psychometric assessment, we used the Perceived Stress Scale (PSS), one of the most widely used psychological instruments for measuring the perception of stress (Cohen and Williamson 1998). It is a measure of the degree to which situations in one's life are appraised as stressful. The PSS was designed for use with subjects having at least a junior high school education. The items are easy to understand. The questions are of a general nature and hence are relatively free of content specific to any subpopulation group (Cohen *et al.* 1983). The instrument includes ten items and subjects are asked to rate the frequency with which they experienced stress or other feelings on a five-point Likert scale. The interviewer recorded the responses of the interviewee on a sheet. However, since our focus was the appropriateness of the videophone for this type of assessment and not the stress level of the subjects, these datasets were discarded. Once all questions of the PSS were covered, the subjects terminated the video-call (by hanging up).

After the session was over, each subject was asked to fill out the quality scale ITU-T P.920 that captures the perceived audio and video quality. The International Telecommunication Union (ITU) is an international body that creates and maintains quality assessment methods for telecommunications. For the purposes of this study we used the quality scale ITU-T P.920 (interactive test method for audiovisual communication), which includes items that assess audiovisual quality as perceived by the user (ITU-T 2004).

In addition, we recorded the age and gender of each subject. After subjects filled out the form, some provided comments and other feedback that were recorded by a member of the research team. The responses to the ITU-T-P.920 were entered into an Statistical Package for the Social Sciences (SPSS) program for analysis. Additional comments and remarks were reviewed and coded by members of the research team.

3. Results

Fifty-two subjects participated in 26 sessions. Twenty-eight were male and 24 female. The average age was 35.8 years (age range 22 – 58 years old). The overall quality of the video-call was rated in average as high (mean 4.19 on a five-point Likert Scale, SD 0.59). Thirteen respondents (25%) stated that they or their conversation partner had difficulty in talking or hearing during the conversation. Table 1 summarises participants' responses to the ITU-T P.920 survey.

Table 1. Findings and subjects' responses.

Number of total participants	N = 52
Age	35.84 (SD 10.74, Mean 34)
Gender	Male: 28 (54%); Female: 24 (46%)
Quality of speech-audio connection *5-Excellent* *4-Good* *3-Fair* *2-Poor* *1-Bad*	Average 4.44 (SD 0.69)
Effort required to understand the meaning of sentences: *5-Complete relaxation possible; no effort required* *4-Attention necessary; no appreciable effort required* *3-Moderate effort required* *2-Considerable effort required* *1-No meaning understood with any feasible effort*	Average 4.07 (SD 0.62)
Did you or your partners have any difficulty in talking or hearing over the conversation:	Yes: 13 (25%); No: 39 (75%)
Image quality: *5-Excellent* *4-Good* *3-Fair* *2-Poor* *1-Bad*	Average 3.96 (SD 0.68)
Image impairment: *5-Not Noticeable* *4-Noticeable, but not annoying* *3-Slightly annoying* *2-Annoying* *1-Very Annoying*	Average 3.98 (SD 0.69)
Overall quality of video-call: *5-Excellent* *4-Good* *3-Fair* *2-Poor* *1-Bad*	Average 4.19 (SD 0.59)

There was, as anticipated, a moderate correlation between perceived audio quality and overall quality of the video-call (Spearman's rho = 0.4; $p < 0.05$) and between image quality and overall quality of the video-call (Spearman's rho = 0.54; $p < 0.05$). The correlation between the ease of understanding the meaning of sentences and the overall perceived quality of the video-call was not significant, while the correlation between the image clarity and overall perceived quality of the video-call was moderate (Spearman's rho = 0.47; $p < 0.05$). These findings could indicate that the visual dimension played a great role (maybe even greater than the audio dimension) in the way subjects perceived the quality of video-calls. There was no significant correlation between gender and overall perceived quality of the video-call (nor between age and overall perceived quality of the video-call).

Three participants commented that they enjoyed the interview having the visual feedback and perceived the interview as more personal. Two of them commented that they are usually uncomfortable when interviewers ask personal questions over the phone where they cannot see each other, while the video feature created an intimacy that allowed them to comment on their stress levels. Twenty-one (40%) participants stated that they would be interested in purchasing a videophone if the cost is reasonable. One of the participants with a hearing impairment commented that he was experienced in the use of PC-based videoconferencing and thus, the slight delay of the video images did not surprise or disturb him, as he knew not to attempt to lip-read. Another participant commented on the ease of use and her surprise that the videophone looked like a regular phone with an added screen instead of a complex device that would require extensive training as she had anticipated. Four participants (7.6%) found the visual feedback useful in assessing the conversational partner's reaction to the questions. Four participants who had assumed the role of the interviewers commented that they could see on the screen that their conversation partner was thinking of a response to the questions and that helped them to interpret the silence after each question, an observation that they would not have been able to make over a regular phone.

4. Discussion

The study findings indicate that participants perceived the quality of the video-call to be satisfactory for the administration of the perceived stress scale (PSS). All of the participants found that there was no considerable effort required to understand the questions and/or answers of the interview. None of the participants felt that there was any point during the interview where they did not understand the meaning of sentences. The image quality was always rated as fair, good or excellent and none of the participants reported annoying image impairment.

These findings give strength to the argument that videophones can be used for regular patient-provider interaction and the administration of standardised behavioral instruments. One of the users characteristics that we did not assess during this study but could be of relevance to the overall user experience is their previous experience with technology. Users of multimedia communication systems often learn to develop skills to address the constraints of an application and take full advantage of the medium. One could expect that users who have previous experience with videophones or low-cost commercially available videoconferencing solutions would show different levels of tolerance to technical problems and rate the overall quality of the video-call differently than a user who tries out this mode of communication for the first time. Studies have shown that groups of users who have gained considerable experience of

video communication over a longer period of time provide higher quality ratings and have more positive attitude than users who are novice and are testing video-applications in a laboratory for a very short time (Tang and Isaacs 1993, Doherty-Sneddon *et al*. 1997, Rudman *et al*. 1997).

The subjective assessment of the video-call quality does not always reflect all potential uses and limitations of the technology. The focus is on the user's perception and not the technical features. In a study of videophone systems, for example, investigators determined that any increase in visual representation of the speaker increased the viewer's tolerance to audio noise (Ostberg *et al*. 1989). A similar finding was reported in experimental studies (Negroponte 1995) with High Definition Television (HDTV) where investigators improved the perceived quality of video by increasing the audio quality only. The fact that users perceive levels of quality depending on the task and the context adds value to subjective evaluation when the aim is to determine the appropriateness of the application for a specific domain area and a given purpose. In other words, our evaluation might not inform the vendors and developers about overall performance of videophone technology when compared to other applications or other videophones. It does, on the other hand, provide insight into the suitability of the technology for the administration of psychometric tools, which is an essential part of home and hospice care monitoring. Even if laboratory testing indicates satisfactory frame rate, latency or image clarity, a user's perception of the overall system quality and its utility in a given context will determine the system utilisation rates and ultimately, its success. Thus, determining the appropriateness of this communication mode helps us understand the process and impact of 'virtual' visits in healthcare. As technology advances, there will continue to be advanced systems that utilise videoconferencing to allow for remote monitoring and 'virtual' interactions. In addition to the cost-effectiveness of such applications we need to evaluate the actual process of video-mediated patient assessment. Patients' and healthcare providers' level of acceptance of this communication mode will greatly determine its success and diffusion.

Acknowledgments

This study is supported in part by the John A Hartford Social Work Faculty Scholars Program administered by the Gerontological Society of America and the Catherine Pouget Research Award of the MAPI Research Institute.

References

Apteker, R.T., Fischer, J.A., Kisimov, V.S. and Neishols, H., 1995, Video acceptability and frame rate. IEEE Multimedia, **2,** pp. 32–40.

Argyle, M., Lalljee, M. and Cook, M., 1968, The effects of visibility on interaction in the dyad. *Human Relations*, **28,** pp. 289–304.

Bruce, V., 1996, The role of the face in communication: implications for videophone design. *Interacting with Computers*, **8,** pp. 166–176.

Clark, H. and Brennan, S., 1991, Grounding in communication. In *Perspectives on Socially Shared Cognition*, L.B. Resnick, J. Levine and S. Teasley, (Eds), (Washington, DC: APA Press).

Clark, H. and Schaefer, E., 1989, Contributing to discourse. *Cognitive Science*, **12,** pp. 259–292.

Cohen, S. and Williamson, G., 1998, Perceived stress in a probability sample of the United States. In *The Social Psychology of Health*, S. Spacapan and S. Oskamp, S. (Eds) *The Social Psychology of Health* (Newbury Park, CA: Sage).

Cohen, S., Kamarck, T. and Mermelstein, R., 1983, A global measure of perceived stress. *Journal of Health and Social Behavior*, **24,** pp. 385–396.

Dansky, K.H. and Bowles, K.H., 2002, Lessons learned from a telehomecare project. *Caring*, **21,** pp. 18–22.

Demiris, G., Oliver, D.R., Porock, D. and Courtney, K., 2004, The Missouri Telehospice Project: Background and next steps. Home Health Care Technology Report, **1,** pp. 49–55.

Demiris, G. Speedie, S., Finkelstein, S. and Harris, I., 2003, Communication patterns and technical quality of virtual visits in home care. *Journal of Telemedicine and Telecare*, **9,** pp. 210–215.

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S. and Bruce, V., 1997, Face to face and video mediated conversation: a comparison of dialogue structure and task performance. *Journal of Experimental Psychology*, **3,** pp. 105–25.

Duncan, S., 1972, Some signals and rules for taking speaker turns in conversation. *Journal of Personal and Social Psychology*, **23,** pp. 283–292.

Ekman, P. and Friesen, W., 1975, *Unmasking the Face* (New Jersey: Prentice Hall).

ITU-T Recommendation P.920, Interactive test method for audiovisual communications. Available online at: http://www.itu.int/ (accessed 28 October 2004).

Mehrabian, A., 1971, *Silent Messages* (Belmont, NJ: Wadsworth Press).

Mehrabian, A., 1972, *Nonverbal Communication*, pp. 12–29 (Chicago: Aldine-Atherton).

Negroponte, N., 1995, *Being Digital* (London: Hodder & Stoughton).

Ostberg, O., Lindstrom, B. and Renhall, P.O., 1989, Contribution of display size to speech intelligibility in videophone systems. *International Journal of Human Computer Interaction*, **1,** pp. 149–159.

Reid, A., 1977, Comparing the telephone with face-to-face interaction. In *The Social Impact of the Telephone Pool*, I. Pool (Ed.) (Cambridge, MA: MIT Press).

Rudman, C., Hertz, R., Marschall, C. and Dykstra-Erickson, E., 1997, Channel overload as a driver for adoption of desktop video for distributed group work. In *Video-Mediated Communication*, K.E. Finn, A.J. Sellen and S.B. Wilbur (Eds), pp. 199–223 (Mahwah, NJ: Lawrence Earlbaum Associates).

Tang, J.C. and Isaacs, E.A., 1993, Why do users like video: study of multimedia supported collaboration. *Computer Supported Cooperative Work*, pp. 163–196.

Whittaker, S., 1995, Rethinking video as a technology for interpersonal communications: theory and design implications. *International Journal of Human-Computer Studies*, **42,** pp. 501–529.

Young, N.L., Barden, W., Lefort, S., Nijssen-Jordan, C., Daniels, C., Booth, M. and Dick, P.T., 2004, Telehomecare: a comparison of three Canadian models. *Telemedicine Journal and e-Health*, **10,** pp. 45–52.

Zuiderent, T., Winthereik, B.R., and Berg, M., 2003, Talking about distributed communication and medicine: on bringing together remote and local actors. *Human-Computer Interaction*, **18,** pp. 171–180.