

# Investigating the Effects of using Biofeedback as Visual Stress Indicator during Video-mediated Collaboration

Chiew Seng Sean Tan    Johannes Schöning    Kris Luyten    Karin Coninx

Hasselt University  
tUL - iMinds

Expertise Centre for Digital Media  
3590 Diepenbeek, Belgium  
firstname.lastname@uhasselt.be

## ABSTRACT

During remote video-mediated assistance, instructors often guide workers through problems and instruct them to perform unfamiliar or complex operations. However, the workers' performance might deteriorate due to stress. We argue that informing biofeedback to the instructor, can improve communication and lead to lower stress. This paper presents a thorough investigation on mental workload and stress perceived by twenty participants, paired up in an instructor-worker scenario, performing remote video-mediated tasks. The interface conditions differ in task, facial and biofeedback communication. Two self-report measures are used to assess mental workload and stress. Results show that pairs reported lower mental workload and stress when instructors are using the biofeedback as compared to using interfaces with facial view. Significant correlations were found on task performance with reducing stress (i.e. increased task engagement and decreased worry) for instructors and declining mental workload (i.e. increased performance) for workers. Our findings provide insights to advance video-mediated interfaces for remote collaborative work.

## Author Keywords

CSCW, Biofeedback, Stress, Video-mediated Collaboration.

## ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): User Interfaces, Interaction Styles.

## General Terms

Performance; Experimentation; Design; Human Factor.

## INTRODUCTION & MOTIVATION

A video-mediated interface is an essential tool to help users coordinate their activities when performing collaborative



**Figure 1.** Left: The interface condition was shown to the instructor on a laptop. The instructor conceptualised the build plan of the model and communicated each step of the building process as instructions to the worker. Right: A camera was fixated at the physical workspace and a second camera captures a facial view of the worker. Physiological sensors are placed on the fingers and strapped around the upper body of the worker.

tasks. Particularly during collaborative construction tasks, which require extensive coordination between two or more individuals working together to manipulate three-dimensional physical objects in the real world: e.g. a remote expert (instructor) might guide a field technician through a complicated chemical sampling procedure or help a domestic user troubleshoot an electrical home appliance.

The role of an instructor in video-mediated assistance [16] is to perform two primary tasks: (1) guiding workers to diagnose a problem, and (2) instructing them to perform unfamiliar operations. Issuing instructions that are not understandable by the workers or overloading them with instructions can result in severe performance deterioration during remote collaboration. One factor that may affect such performance might be stress, which can be defined as a process that disrupts, or threatens to disrupt, a person's equilibrium. Fussell, Setlock and Kraut [7] identified two mechanisms that might be useful for the instructors to coordinate their assistance. The first mechanism provides situation awareness about the workers' activities, status of the task, and the environment. The second mechanism relates to conversational grounding and refers to the common ground in which communicators work together to ensure messages are understood as intended.

We argue that instructors should be able to optimise these two mechanisms through continuous access of relevant visual information that provides assessment of stress. This will prove

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. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).  
CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada.  
Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-2473-1/14/04\$15.00.  
<http://dx.doi.org/10.1145/2556288.2557038>

to be valuable in task intensive scenarios, particularly tasks for which failure could lead to physical injury or worse in which situation awareness and conversational grounding will be ameliorated as instructors access work-related stress information to monitor the workers' comprehension and task progression. For example, scenarios could be a nuclear power plant surveillance or information-sensitive consultation between a doctor and a patient. Rich media, which is capable of providing an embodiment for the stress indication, can be used to disseminate workers' stress symptoms to be remotely interpreted by instructors.

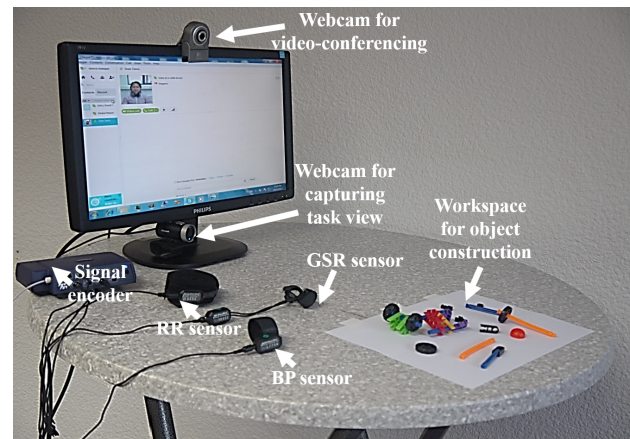
Recent advances in biosensor technology can allow instructors to remotely access the workers' physiological responses (e.g. heart rate and skin conductance). Physiological sensing and biofeedback presents a visual display of physiological responses and are commonly used by researchers to study stress symptoms in work-related scenarios. For example, vehicle driver struggling with difficult traffic [11] and office worker overloaded with assignments [13]. Electrocardiograms and galvanic skin response are commonly used to analyse the stress level of the driver for different warning notifications [11]. Skin conductance sensors are used to measure the electrodermal activity and are associated with stress levels to study the disruptive response of stress in call centers [13] and during social interaction (e.g. emails, phone calls, SMS) [1]. For a comprehensive review on physiological measures in lab-based and real-life stress scenarios, we refer readers to the survey by Biondi and Picardi [3].

The contribution of this paper is to provide a thorough investigation on the mental workload and stress perceived by both instructors and workers performing collaborative physical tasks in a video-mediated setup. We developed a biofeedback implementation that is capable of presenting a visual stress indication of a worker to an instructor, and demonstrated its technical feasibility for video-mediated collaboration. We then study the impact of using this biofeedback, with facial and/or task views, in a user study to construct complex devices. We conclude with a discussion and implications of using biofeedback to support remote collaborative work.

## RELATED WORK

Various authors have proposed displaying visual information to coordinate remote assistance and improve collaborative tasks. More specifically, recent systems incorporating situation awareness and conversational grounding in the form of visual information have been explored [17, 24] to engage remote (instructor) collaboration in a video-mediated setup. Research utilising physical environment as a mean for task engagement includes Lighthouse [21] and BISi [20], which explored physical setup and proxemics to establish a task-centred point of view for situation awareness.

Heath and Luff [12] pointed out that people having a face-to-face interaction may respond automatically to cues arising from gaze direction or physical gestures that are registered in peripheral vision, whereas users communicating via video-conferencing may completely fail to pick up these signals. Videoconferencing systems such as Multiview [19] and ClearBoard [14] were developed to address these shortfalls.



**Figure 2. Video-mediated setup at the worker's location: Devices include webcams for video-conferencing and capturing the task view, signal encoder, and biosensors for Respiratory Rate (RR), Galvanic Skin Response (GSR) and Blood Pressure (BP).**

These systems reproduce directional nonverbal cues and provide conversational grounding for people at different sites.

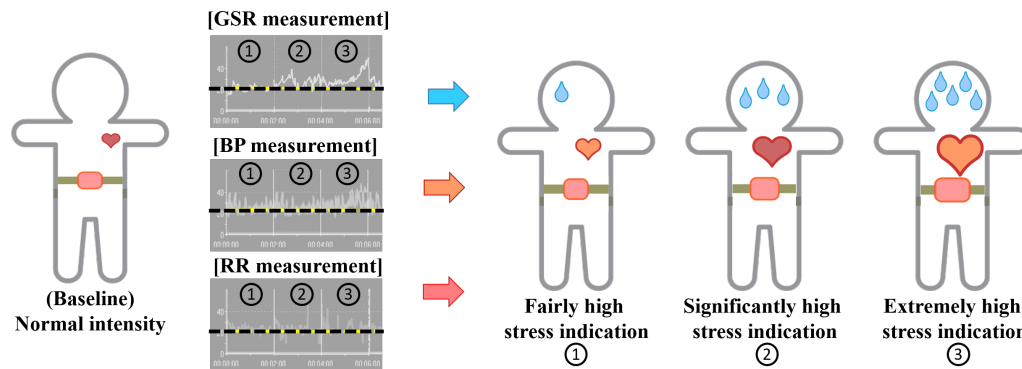
Systems that combine situation awareness and conversational grounding within a single interface have also been investigated to engage remote parties. For instance, Fussell, Setlock and Kraut [7] captured both the task performance and facial view of remote users to enhance collaborative physical tasks through shared visual space. People in Books [6] provides a feeling of emotion togetherness for children to interact with story characters superimposed with the faces of their remote parents. These systems capture both task and facial information about the users on the video-mediated interface to support situational awareness and conversational grounding.

## USER STUDY

We argue that informing instructors with biofeedback, which communicates the remote workers' physiological responses, can improve communication and lower stress for both instructors and workers. Our study analyses the effects of different interface conditions for collaborative construction of complex devices – K'Nex systems. The conditions differ in video composition of task, facial and biofeedback communication. For our analysis, we focused on effectiveness of communicating instructions and profiles of the workload and stress levels as indexed by the self-reporting of the instructors and workers.

### Study Design

Physiological measures for stress detection pose several challenges as identified by Plarre et al. [23]. They are, however, recommended by Shostak and Peterson [25] to provide objective measures and ground truth on human cognitive performance. We list three practical limitations of current physiological sensing and use the criteria for cognitive work analysis [27] to guide our user study and our implementation for biofeedback. (1) *Subjective ground truth*: quantification to the actual stress response can provide an understanding of the stress symptoms during collaborative task. However, interpretation of stress is highly subjective: it will differ between individuals and situations and even at different times.



**Figure 3. Animations for stress indication.** Galvanic Skin Response (GSR) measure is represented at the top followed by Blood Pressure (BP) and Respiratory Rate (RR) measures. The animation varies according to the physiological measurements with baselines indicated by the dotted lines.

We used both self-reports and physiology-based assessment to access and measure each participant directly so that the stress effects can be determined comprehensively. (2) *Large between-person differences*: the diverse range of physiological response to stress in each individual can make it difficult to build a stress classifier [23]. Instead we adopt a third-party stress evaluation approach, which has been used extensively for studying clinical stress, to calibrate each participant. The calibrated stress response is then animated by a visual stress indicator. (3) *Undesired behavioural expressions*: stress can induce nonverbal communication which influence social interaction, such as attracting attention and showing empathy. For that reason, we use a strict lab protocol to ensure that the task-induced stress we want to measure is caused by executing the task itself, i.e. dialog communication, time pressure, and demand overload. In addition, the lab protocol ascertain that physiological measurements resumed to baseline levels prior to the start of each task.

### Hypotheses

Drawing from psychophysiology and biosensor technology that physiological responses of geographically distributed workers can be accessed and displayed as biofeedback, we argue that informing instructors with such biofeedback in the video-mediated interface can improve communication and lower their stress. Our primary hypothesis is to determine whether pairs will experience less mental workload and stress to complete their collaborative tasks when instructors are informed with the biofeedback. We explore different configurations, which include (1) facial view of worker, (2) biofeedback communicating the stress indication, as well as (3) task-centred view on the physical workspace. For our study we define the following three hypotheses, covering the different configurations of a video-mediated interface to complete remote collaborative tasks, we are interested in:

- H1** Instructors with a combined worker facial and worker biofeedback view will result in a lower workload and less stress for both instructors and workers than using only worker biofeedback view.
- H2** Instructors with an integrated worker biofeedback view will result in a lower workload and less stress for both instructors and workers than using only a worker facial view.

**H3** Instructors with an integrated worker biofeedback view will result in a lower workload and less stress for both instructors and workers than using only the worker physical workspace view.

### Method

#### Participants

Twenty participants (13 males, 7 females) between the age of 19 and 35, median age of 26 years old, participated in the study. All had prior experience using video-conferencing tools and used them at least 1 to 2 hours a week for work-related purposes. Participants are paired up to complete the tasks in a video-mediated setup as shown in Figure 1.

The conditions of our study corresponded to different interface configurations. The study involved pairs of participants building K'Nex models in an Instructor-Worker scenario: one participant acted as an expert user (instructor) who provides instructions for the other participant who acted as the field user (worker). The workers had to perform building tasks on physical objects, while the instructors' tasks were to diagnose and guide workers through the build plans. The workers did not have access to the build plans, which were conceptualised by the instructors. Thus, the workers need to be informed by the instructors for each step in the building process.

#### Apparatus

In each trial, the instructor and worker were connected through a Skype video-conferencing call for facial and audio communication. We used Procomp hardware and Biograph infinity from Thought Technology to measure the worker's Galvanic Skin Response (GSR), Blood Pressure (BP) and Respiratory Rate (RR). The GSR and BP sensors were attached to the last 3 fingers on the non-dominant hand to reduce noise contamination from object handling. The RR sensor was attached above the abdomen to measure the chest expansion for respiratory rate.

The measurement for each physiological modality was obtained from a physiological stress assessment and subsequently used to calibrate each worker. We then normalised the stress response from each physiological modality, and transformed it into a visual representation that can be interpreted as a stress indication. We apply a transformation



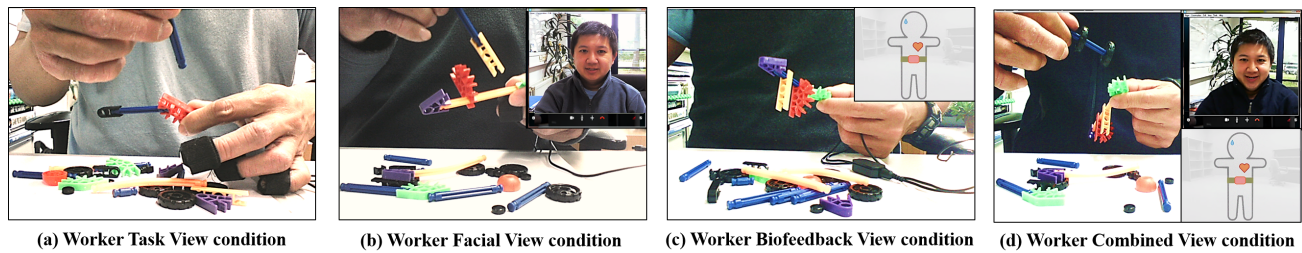


Figure 4. The four interface conditions in the experimental design for video-mediated collaboration.

to correct the physiological measures for mapping nonlinear relationships so that the sigmoidal nonlinearity can account for abrupt changes in an emotional response [22]. Basically, this is achieved by mapping the normalised stress response to regular intervals in a sigmoid distribution. For example, no perspiration indicates normal GSR reading and having 5 drops of perspiration indicate extremely high stress indication from GSR reading. Likewise, animations for BP and RR are represented with two different pulsating movements, corresponding to the actual measurements. They are normalised with settings at 60 to 180 heartbeats per minute and 12 to 24 breaths per minute for the BP and RR respectively. The individual component changes according to its assigned biofeedback (e.g. heart size increased for BP alone).

Animations of physiological cues are depicted on a human avatar so that instructors or non-expert users are able to develop their own interpretation of the stress indication. The animations provide concordance with stress response (e.g. increasing heart rate with heightened stress) and they collectively depict the stress relation in the physiological responses, which is illustrated in Figure 3.

#### Conditions

We designed the interface conditions to explore new use experiences for video-mediated collaboration, which could potentially lead to different ways of communicating the build plan in each condition. Figure 4 shows the different visual layouts we compared. The task objects and finger gestures are within the view of the camera as illustrated on the physical workspace in Figure 2. The first condition shows only the task workspace on the main display; facial and biofeedback information is combined with the main display in different configurations and presented in the other conditions. Other factors, i.e. screen size and resolution, remained unchanged.

**Worker Task View** In this condition, the primary camera is focused on the task object (K'Nex model), which is situated in the physical workspace. The workspace takes up 100% of the window display as shown in Figure 4a.

**Worker Facial View** In this condition, the primary camera is focused on the task object and a secondary camera is focused on the face of the worker. The workspace and facial view in the window display are as shown in Figure 4b.

**Worker Biofeedback View** In this condition, the primary camera is focused on the task object. The visual stress indicator for biofeedback is an animation, which highlights

the physiological responses of the worker. The workspace and biofeedback in the window display (see Figure 4c).

**Worker Combined View** In this condition, the primary camera is focused on the task object and a secondary camera is focused on the face of the worker. The visual stress indicator for biofeedback is an animation, which highlights the physiological responses of the worker. The physical workspace, facial view and biofeedback in the window display are as shown in Figure 4d.

#### Protocol

Figure 5 summarises the complete laboratory study protocol. Participants were welcomed and were given a brief description of the study, including assurances on the anonymity of their responses. They were led to separate rooms to complete the general survey and pre-task questionnaires. The conditions were randomly assigned to each pair in a within-subjects design, which enables instructors to be in a better position to appraise the use of biofeedback with stress indication for remote video-mediated task. In each randomly assigned pair, we select the participant with higher educational and research experiential background as instructor which implied its role of remote expert. E.g. a post-doc researcher is assigned the role of instructor, while a Master student filled the role of worker. Note that we also had participants with other backgrounds and participants in each pair do not know each other prior to the experiment. Order effects were also avoided by counterbalancing the experimental design. The experimental trial lasted approximately 90 minutes for each pair of participants. Prior to the start of the lab protocol, the visualisation of biofeedback with stress indication corresponding to each physiological modality, was functionally described to the participants. A physical illustration of Figure 3 is provided to the instructors as a look-up reference.

After the workers were fitted with the biosensors, the Biograph stress assessment script<sup>1</sup> was administered to identify hyper- or hypo-reactivity in the recorded physiological signals. The workers started with a 3-minute baseline-recording period, during which the participants were relaxed while focusing on a distant nature scenery outside the window. Following that, the participants were exposed to three stressors in the Biograph Stress Assessment script, which are namely colour words, mental arithmetic, and stress event recall. The

<sup>1</sup> <http://www.thoughttechnology.com/physuite.htm>

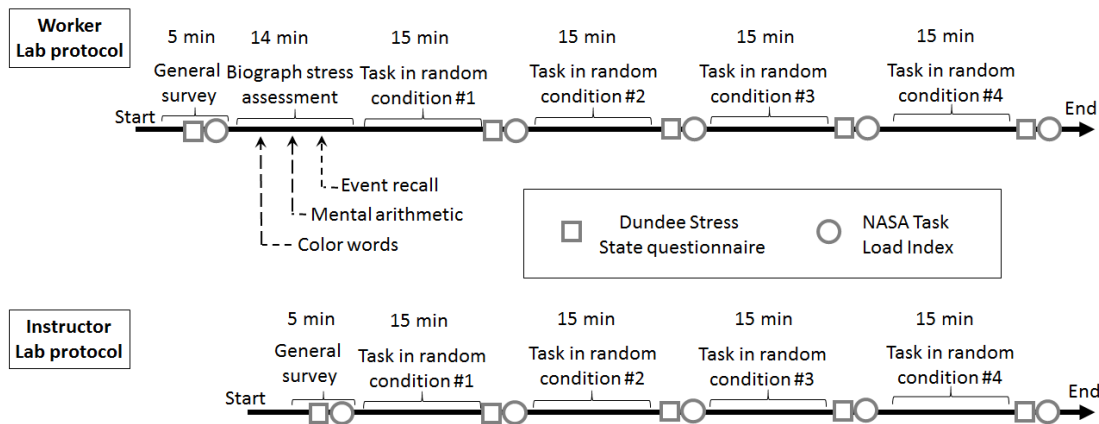


Figure 5. The Lab Study Protocol. Dundee Stress State Questionnaire and NASA Task Load Index were used to assess the stress and mental workload respectively. They were administered prior to the experimental trial and after each task of building K'Nex models.

colour word task is similar to the Stroop test [26] as participants were asked to read out the colour instead of the perceived word. For example, participants have to read out *blue* for the word *red* which is written in blue ink. For the mental arithmetic task participants were asked to continuously subtract a fixed integer from a designated number. For the stress event recall, the participants were asked to articulate (3 minutes) a stressful event that they had experienced recently. In the occurrence that the participants were unable to articulate further details, the experimenters enquire the participants for other stressful events in the distant past. These stressors are representative of the different aspects of cognitive stressors that are known to clinically induce stress in most people [2]. Participants were allowed to relax between stressors and continue the session again when they feel ready. We also ensure that relaxation resumes the physiological readings to baseline levels respecting the different times that participants needed for this.

Subsequently, the workers and instructors were given four tasks of building K'Nex models. The workers and instructors were required to communicate in a comprehensible manner and diagnose the building process collaboratively. The instructors had to continuously formulate a build plan to assemble each model and communicate this plan to the workers. The workers in turn had to follow instructions from the instructors and translate the instructions to physical operations on the final output. The instructors were equipped with the instruction manual for K'Nex models in each task and were told to complete the task within a 12 minutes time frame. Reminders were given every minute for the last remaining 5 minutes. We recorded the number of instructions that are issued and also questions from the instructors that were directed at the workers. The sessions were also video-recorded.

#### Measurement Instruments

In this paper we concentrated on the participants' perception of their cognitive experience and this was assessed with ten self-reports for each instructor and worker throughout the experimental trial (see Figure 5). Current sensors are inept to measure for such cognitive state. To gain understanding of difficulty that users might have experienced, we use (mental

workload and stress) self-reports for expressing their cognitive state. We further cross-check self-reports with physiological sensor readings to identify the relation between cognitive state and stress. This helps us to determine the role of biofeedback during video-mediated collaboration. Initial self-reports to measure stress states and workload were administered to determine the participants' pre-test status. Subsequently, self-reporting were administered after each task of building K'Nex models. The purpose was to determine the extent to which stress and mental workload changes after completing the task in different conditions.

#### Dundee Stress State Questionnaire

We used the Dundee Stress State Questionnaire (DSSQ) as our self-report measure to assess participants' immediate stress moods, motivations, cognitions and coping strategies. The DSSQ was developed by Matthews et al. [18] to provide a comprehensive multi-dimensional assessment instrument for transitory states associated with stress, arousal, and fatigue. It consisted of a total of 77 Likert-items, which represented three primary categories: *Task Engagement* corresponds to energy, task motivation, and concentration, e.g. "How eager are you to do well at the task?", *Distress* measures the focus on arousal, tension, and hedonic tone, e.g. "I thought about the difficulty of the problems.", and *Worry* quantifies the awareness, inference, self-focus, and concentration, e.g. "I am worried about what other people think of me.", on a five-point Likert scale.

#### NASA Task Load Index Questionnaire

The self-reporting instrument for mental workload was the NASA Task Load Index (NASA TLX), which is based on the assumption that subjective workload is a hypothetical construct representing the cost incurred by a human operator to achieve a particular level of performance. NASA TLX was developed by Hart and Staveland [10] to assess mental workload using six dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. The items in each dimension present a statement, such as "How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, etc)?" on a Likert-type scale consisting of 20 five point steps from 0 to 100.

## RESULTS AND ANALYSIS

Evaluations were conducted on the self-reported measures and recorded sessions of participants in the four conditions (as shown in Figure 4). We presented the results in three parts. First, we examined the task performance by analysing the effectiveness of task-oriented conversation. We then computed the DSSQ scores and determined their correlation with the task performance. Similarly for the third part, we analysed the NASA TLX ratings and delineated the correlation with task performance. The results of perceived stress and workload were examined separately for instructors and workers.

### Task performance

We initially considered a performance metric using completion time, but there does not seem to be an effect across our tasks with similar complexity. We used the complexity as described by Ranjan, Birnholtz and Balakrishnan [24], which varied in terms of graphical steps (i.e. 5-9), number of pieces (i.e. 18-27) and unique difficult-to-describe pieces (i.e. 9-12). We counted the number of instructions and divided by the number of minutes for each pair, and compared these across our four tasks. The differences were not statistically significant,  $F(3,27) = 1.154$ ,  $p = .341$ .

Considering instructors communicated their build plan as instructions to the workers, determining how instructions are phrased can attribute to worker's comprehension for task performance. As such, performance was measured in terms of the number of instructions as well as number of questions by the instructors in each step of the building process. We defined a metric (Instruction Comprehensibility) to measure the worker's comprehension in terms of the instructions communicated. The metric is based on a coding scheme involving both instructions and questions from the instructor that focuses on the communication and inquisitional functions. We analyse the communication established through grounding [5], which can be described as a collective process that ensures the worker has understood a previous utterance before the instructor proceeds with the next instruction. Thus, questions related to issuing and learning new instructions are important in monitoring for common ground [9] and analysis at the utterance level provide information on common effort about the pairs contribution during the collaborative task.

Instruction Comprehensibility is characterized as (number of instructions)/(number of instructions + number of questions by instructors). An Instruction Comprehensibility score of 1 indicates maximum comprehension (i.e. instructions are correctly comprehended by the worker and successfully translated to the construction operation). A score below 0.5 indicates that the instructors communicated more questions than instructions to the worker. This performance-based measure is grounded on the assumption that any increase in task difficulty will lead to an increase in questions by the instructors, hence decreasing the task performance.

The mean number of instructions for each condition was: Worker Task View with 18.3 (SD = 1.252), Worker Facial View with 21.4 (SD = 1.398), Worker Biofeedback View with 20.4 (SD = 1.476), and Worker Combined View with 27.5

(SD = 3.496). The corresponding mean Instruction Comprehensibility scores for the conditions were: Worker Task View with 0.849 (SD = 0.047), Worker Facial View with 0.887 (SD = 0.049), Worker Biofeedback View with 0.906 (SD = 0.057), and Worker Combined View with 0.854 (SD = 0.084). Post-hoc analysis using Tukey HSD revealed that Instruction Comprehensibility was significantly lower for the Worker Task View than both Worker Facial View ( $t(9) = -2.294$ ,  $p = .047$ ) and Worker Biofeedback View ( $t(9) = -3.238$ ,  $p = .01$ ). On the contrary, Instruction Comprehensibility was significantly higher with the Worker Biofeedback View than with the Worker Combined View,  $t(9) = 2.758$ ,  $p = .022$ . There is no significant difference on the instruction comprehensibility for Worker Biofeedback View and Worker Facial View.

In other words, participants in the *Worker Biofeedback View condition yield significant higher task performance compared to the Worker Combined View condition*. A significant lower task performance was reported for the Worker Biofeedback View condition than the Worker Task View condition.

### Perceived Stress using DSSQ

Participants' responses on the pre- and post-DSSQ were used to calculate outcome scores on eleven scales, where the normative means and standard deviation used were from Matthews et al. [18]. We analysed the scores for both instructors and workers using a 4 (conditions)  $\times$  11 (scales) repeated measures factorial ANOVA on the change scores. Figure 6 portrayed the change scores in three primary categories.

#### Perceived Stress of Instructors

For the instructors, the repeated measures factorial ANOVA revealed significant main effects in the four conditions,  $F(3, 27) = 6.23$ ,  $p = .002$ , and scales,  $F(2.972, 26.74) = 5.265$ ,  $p = .006$ , as well as a significant conditions  $\times$  scales interaction,  $F(7.42, 66.704) = 2.554$ ,  $p = .002$ .

Further analysis on the conditions  $\times$  scales interaction by means of T-tests, using an alpha level of .05 and the Bonferroni correction, showed that Worker Biofeedback View had a significantly lower DSSQ change scores than both Worker Facial View ( $p = .048$ ) and Worker Combined View ( $p = .048$ ). There was a significant drop in DSSQ change scores for Worker Facial View when compared to the Worker Combined View,  $p = .001$ . We also found significant correlation of task performance and perceived stress for the instructors in two conditions. For the Worker Task View condition, we found moderate correlation ( $p < .05$ ) for Instruction Comprehensibility and Success Motivation with  $r = .628$ , implying that instructors who performed better using Worker Task View felt more engaged with the task. For the Worker Biofeedback View condition, there was a moderate correlation for the Instruction Comprehensibility and Energetic Arousal with  $r = .503$ , implying that instructors who performed better using Worker Biofeedback View felt more engaged with the task.

#### Perceived Stress of Workers

The repeated measures factorial ANOVA of the change scores for the workers revealed significant main effects in the four conditions,  $F(3, 27) = 22.438$ ,  $p = .001$ , and scales,  $F(3.773,$

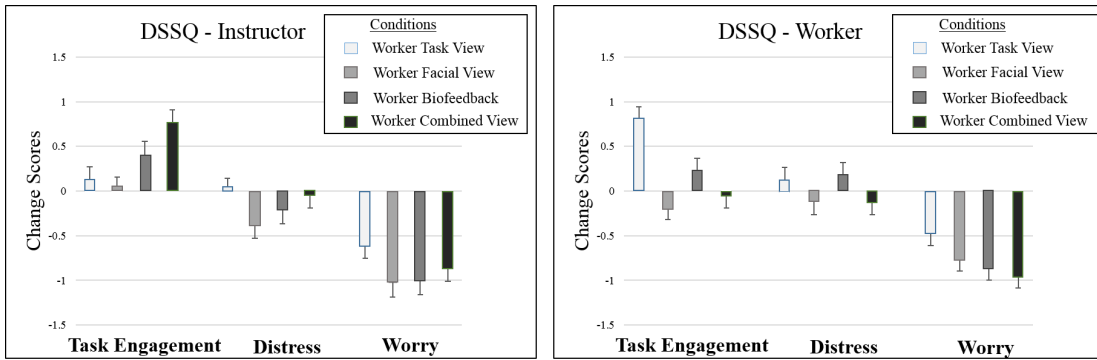


Figure 6. Change scores in three DSSQ categories for the different conditions: Error bars indicate standard deviation of errors. Instructors using Worker Biofeedback View to complete remote collaborative task resulted in significant lower stress for the instructors and workers as compared to both usages of Worker Combined View and Worker Facial View.

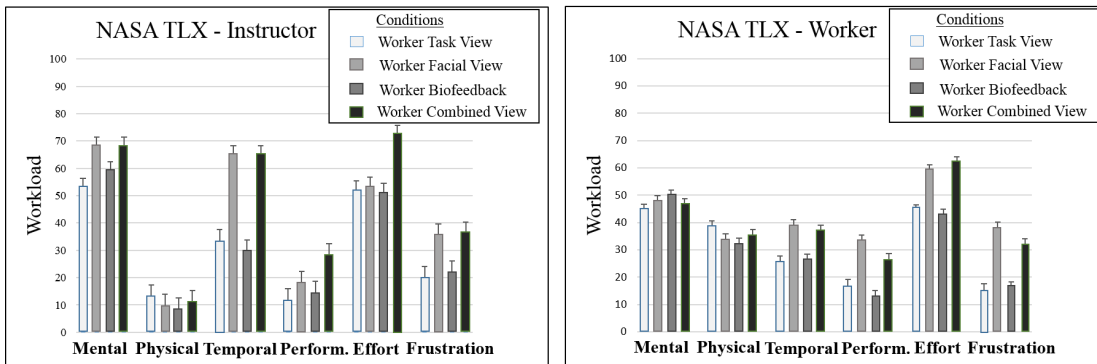


Figure 7. Workload ratings for the four conditions: Error bars indicate standard deviation of errors. Instructors using Worker Biofeedback View resulted in significant lower mental workload for the instructors and workers as compared to both usages of Worker Combined View and Worker Facial View.

33.958) = 2.659,  $p = .05$ , as well as a significant conditions  $\times$  scales interaction,  $F(5.442, 48.974) = 2.689, p = .028$ .

Further analysis on the conditions  $\times$  scales interaction by means of T-tests, using an alpha level of .05 and the Bonferroni correction, showed that Worker Biofeedback View had a significantly lower DSSQ change scores than both Worker Facial View ( $p = .030$ ) and Worker Combined View ( $p = .024$ ). On the contrary, Worker Biofeedback View had significantly higher DSSQ change scores when compared to the Worker Task View ( $p = .024$ ). There was also a significant drop in DSSQ change scores for Worker Task View when compared to both Worker Facial View ( $p = .003$ ) and Worker Combined View ( $p = .002$ ). In addition, we could not find any significant correlation of the task performance with the perceived stress scales for the workers.

In summary, the DSSQ results indicated that *instructors using Worker Biofeedback View to complete remote collaborative task resulted in significant lower stress for the instructors and workers as compared to both usages of Worker Combined View and Worker Facial View*. On the contrary, only workers perceived significantly higher stress for the collaborative task when instructors are using the Worker Biofeedback View as compared to their usage of Worker Task View.

### Perceived workload using NASA TLX

The effects of the four conditions on task-related workload were assessed via standardized scores on each of the six NASA TLX scales for both instructors and workers. The mean weighted ratings are presented in Figure 7. The weighted ratings data were subjected to a 4 (conditions)  $\times$  6 (scales) repeated measures ANOVA.

#### Perceived Workload of Instructors

For the instructors, the repeated measures factorial ANOVA revealed significant main effects in the four conditions,  $F(3, 27) = 11.781, p < .001$ , and scales,  $F(4.74, 42.664) = 23.853, p < .001$ , as well as a significant conditions  $\times$  scales interaction,  $F(7.247, 65.223) = 3.396, p = .003$ . The average overall workload for instructors (in Figure 7) showed slightly above middle ratings for the Mental demand and Effort scales. The overall workload clearly indicated that there was relatively low rating for Physical Demand and Performance scales. The highest workload experienced was using the Worker Combined View followed by the Worker Facial View.

Further analysis on the conditions  $\times$  scales interaction by means of T-tests, using an alpha level of .05 and the Bonferroni correction, revealed that Worker Biofeedback View had significantly lower NASA TLX ratings than both Worker Facial View ( $p = .049$ ) and Worker Combined View ( $p = .023$ ). There was also a significant drop in NASA TLX ratings for

Worker Task View when compared to both Worker Facial View ( $p = .017$ ) and Worker Combined View ( $p = .016$ ). We found a significant correlation of task performance and perceived workload for the instructors in two conditions. For Worker Task View condition, there was an inverse correlation ( $p < .05$ ) for Instruction Comprehensibility and Performance level with  $r = -.612$ , implying that instructors who performed better using Worker Task View felt less work pressure on their performance for task completion. For Worker Combined View, a significant correlation  $r = .584$  was observed with the Instruction Comprehensibility and Mental demand, implying that instructors who performed better using Worker Combined View felt higher demand on their mental workload.

#### *Perceived Workload of Workers*

The repeated measures factorial ANOVA of the workload ratings for the workers revealed significant main effects in the four conditions,  $F(3, 27) = 5.645$ ,  $p = .004$ , and scales,  $F(1.996, 17.96) = 6.364$ ,  $p = .008$ , as well as a significant conditions  $\times$  scales interaction,  $F(7.125, 64.125) = 2.366$ ,  $p = .032$ . The highest workload experienced was using the Worker Facial View followed by the Worker Combined View.

Further analysis on the conditions  $\times$  scales interaction by means of T-tests, using an alpha level of .05 and the Bonferroni correction, showed that Worker Biofeedback View had a significantly lower NASA TLX ratings than both Worker Facial View ( $p = .05$ ) and Worker Combined View ( $p = .014$ ). We found a significant correlation of task performance and perceived workload for the workers in three conditions. For the Worker Task View condition, there were inverse correlations ( $r = -.609$  and  $r = -.579$ ) for the Instruction Comprehensibility with Physical demand and Performance level respectively, implying that workers who performed better when instructors are using Worker Task View felt less demand on physical object handling while having higher work pressure on their performance for task completion. For the Worker Biofeedback View condition, a significant strong correlation  $r = .715$  was observed for Instruction Comprehensibility and Performance level, implying that workers who performed better when instructors are using Worker Biofeedback View felt less work pressure on their performance for task completion. For the Worker Combined View condition, we observed an inverse correlation with the Instruction Comprehensibility and mental Effort level,  $r = -.668$ , implying that workers who performed better when instructors are using Worker Combined View felt less mental effort is demanded for the task.

In summary, our analysis on the NASA TLX ratings revealed that *instructors using Worker Biofeedback View resulted in significant lower mental workload for the instructors and workers as compared to both usages of Worker Combined View and Worker Facial View*. There are no significant differences for Worker Biofeedback View and Worker Task View in terms of mental workload for both instructors and workers.

#### **DISCUSSION**

We systematically investigated the mental workload and stress for each pair of participants performing remote collaborative tasks. The user study revealed support for Hypothesis H2 that workload and stress were perceived to be

lower for the instructors and workers in the Worker Biofeedback View condition than the Worker Facial View condition during video-mediated collaboration. Based on the change scores in Dundee Stress State Questionnaire, we found significant results. However, hypothesis H1 was not supported as the Worker Combined View resulted, in fact, in higher stress for both instructors and workers as compared to the Worker Biofeedback View condition. There was no support for Hypothesis H3 that the Worker Biofeedback View condition yields less stress for both instructors and workers than the Worker Task View condition.

Based on the NASA Task Load Index results, there was support for H2 but not H1. Similar to the findings from DSSQ, Hypothesis H1 was not supported as the Worker Combined View condition showed significantly higher workload for both instructors and workers than the Worker Biofeedback View condition. Conversely, there was no support for H3.

In summary, instructors using Worker Biofeedback View for video-mediated collaboration resulted in pairs of participants to perceive lower workload and less stressful when compared to conditions of instructors using both Worker Facial View and Worker Combined View. The increase in mental workload and stress perceived in the latter two conditions could be explained by Galinsky et al. [8] who also found that participants reported higher levels of stress while examining facial cues in (video-) mediated communication. They concluded that the participants were presumably affected by the negative effects of attention constraint, tension, and eye-strain. Our findings are generally consistent with Whittaker's findings [28], which reported that adding additional facial features can be counterproductive (e.g. showing another person's face during interaction tends to be less effective) when the goal of the interaction is task-oriented.

#### **Implications from stress and mental workload profiles**

The stress and mental workload profiles (as portrayed in Figure 8 and 9 respectively) for each condition are generated by flagging the significant post-test changes when the confidence intervals on population variances did not encompass their initial mean values. A closer examination revealed that instructors felt more energised for the Worker Biofeedback View condition on both stress and workload profiles in terms of correlations with task performance. Similarly, the workers felt more energetic and perceived higher performance for this condition. In other words, the Worker Biofeedback View condition has led to an increase of task engagement in both instructors and workers. Although lower worry (decreasing self-focused attention and interference on task irrelevant issues as reflected in their stress profile) was perceived by the instructors, the workers had experienced a heightened sense of distress and performance level (which are reflected in their stress and workload profiles respectively). Further analysis revealed (1) moderate correlation between task performance and Energetic Arousal for instructors, which implies instructors who performed better felt more engaged with the task, and (2) strong correlation between task performance and Performance level for workers, which implies workers felt less work pressure as they perform better for the task.



DSSQ Scales (Instructors)	Task View	Facial View	Biofeed back	Combined View	DSSQ Scales (Workers)	Task View	Facial View	Biofeed back	Combined View
(TE) Energetic arousal	↑	↑	↑	↑	(TE) Energetic arousal		↓	↑	
(TE) Interest motivation		↓		↓	(TE) Interest motivation				
(TE) Success motivation	↑				(TE) Success motivation				
(TE) Concentration					(TE) Concentration	↑			
(D) Hedonic tone					(D) Hedonic tone				
(D) Tense arousal		↓			(D) Tense arousal			↑	
(D) Control & confidence					(D) Control & confidence				
(W) Self-focused attention		↓	↓	↓	(W) Self-focused attention				
(W) Self Esteem	↓			↓	(W) Self Esteem				
(W) Task relevant CI					(W) Task relevant CI				
(W) Task irrelevant CI	↓	↓	↓	↓	(W) Task irrelevant CI	↓	↓	↓	↓

Figure 8. Significant changes for the DSSQ scales for the four conditions at 95% Confidence Intervals. The primary categories are abbreviated as Task Engagement (TE), Distress (D), and Worry (W). Instructors felt more engaged with the task w.r.t performance metric using Worker Biofeedback View.

NASA TLX Scales (Instructors)	Task View	Facial View	Biofeed back	Combined View	NASA TLX Scales (Workers)	Task View	Facial View	Biofeed back	Combined View
Mental demand		↑	↑	↑	Mental demand	↓	↑	↑	
Physical demand					Physical demand	↑			
Temporal demand	↓	↑	↑	↑	Temporal demand	↓			
Performance	↓		↓		Performance	↓		↑	
Effort	↓	↑	↑	↑	Effort		↑		↑
Frustration	↓		↓		Frustration	↓		↓	

Figure 9. Significant changes for the NASA TLX scales for the four conditions at 95% Confidence Intervals. Statistical analysis shows that workers felt less work pressure in relation to the performance metric for task completion while instructors are using the Worker Biofeedback View.

The use of Worker Biofeedback View has caused a decline in instructors’ worry (self-focused attention and task-irrelevant cognitive interference). Thus suggesting a shift away from emotion-focused coping and avoidance coping towards problem-focused coping, which similarly has positive influence on work as reported by Bowman and Stern [4]. On the contrary, the use of Worker Biofeedback View by the instructors had caused an increase in distress (tense arousal) for the workers, which denoted emotion-focused coping was being employed by the workers to regulate their emotional consequences so that their performance for the task can be improved. This corresponds with the emotion-focused coping with high emotional distress reported by Knight et al. [15].

**CONCLUSION**

This paper reports our findings regarding the mental workload and stress effect on using biofeedback during video-mediated collaboration. The biofeedback informed the instructors about the physiological responses of their geographically distributed workers. We evaluated pairs of participants during remote task collaboration in four conditions (Worker Task View, Worker Facial View, Worker Biofeedback View, and Worker Combined View) with two self-report measures for mental workload and stress states. Significant differences were found for both measures and they supported our hypothesis H2. Although a significant difference was observed for the hypothesis H1, it was not supported as the Worker Combined View condition resulted in fact in higher workload and more stress for both instructors and workers as com-

pared to the Worker Biofeedback View condition. A third hypothesis (H3) that instructors using the Worker Biofeedback View would yield lower workload and less stress for the pairs to complete remote collaborative task when compared to Worker Task View was not supported. In summary, the Worker Biofeedback View proved to be a more effective interface in terms of stress change scores and workload ratings than the Worker Facial View and the Worker Combined View.

According to our findings, integrating biofeedback for visual stress indication and as part of the video-mediated interface can improve remote task collaboration. More specifically, the stress level (i.e. increased task engagement and decreased worry) might be reduced for the instructors as well as declining the mental workload (i.e. heightened performance) for the workers. Coupled with coping strategies, these insights will be able to advance the development of video-mediated interfaces for remote video-mediated assistance.

**REFERENCES**

1. Ayzenberg, Y., Rivera, J. H., and Picard, R. Feel: frequent eda and event logging – a mobile social interaction stress monitoring system. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems*, ACM (New York, NY, USA, 2012), 2357–2362.
2. Berntson, G. G., Binkley, P. F., Uchino, B. N., Quigley, K. S., and Fieldstone, A. Autonomic cardiac control: Psychological stress and cardiac response in autonomic

- space as revealed by pharmacological blockades. *Psychophysiology* 31 (1994), 599–608.
3. Biondi, M., and Picardi, A. Psychological stress and neuroendocrine function in humans: the last two decades of research. *Psychotherapy and Psychosomatics* 68, 3 (1999), 114–50.
  4. Bowman, G. D., and Stern, M. Adjustment to occupational stress: The relationship of perceived control to effectiveness of coping strategies. *Journal of Counselling Psychology* 42, 3 (1995), 294–303.
  5. Clark, H. H., and Wilkes-Gibbs, D. Referring as a collaborative process. *Cognition* 22 (1986), 1–39.
  6. Follmer, S., Ballagas, R. T., Raffle, H., Spasojevic, M., and Ishii, H. People in books: using a flashcam to become part of an interactive book for connected reading. In *Proc. of CSCW '12*, ACM (New York, NY, USA, 2012), 685–694.
  7. Fussell, S. R., Setlock, L. D., and Kraut, R. E. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proc. of CHI '03*, ACM (New York, NY, USA, 2003), 513–520.
  8. Galinsky, T. L. *Psychophysical Determinants of Stress in Sustained Attention*. University of Cincinnati, 1991.
  9. Graesser, A. C., Person, N. K., and Huber, J. Question asking during tutoring and the design of educational software. 149–172.
  10. Hart, S. G., and Staveland, L. E. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Human mental workload* 1, 3 (1988), 139–183.
  11. Healey, J. A., and Picard, R. W. Detecting stress during real-world driving tasks using physiological sensors. *Trans. Intell. Transport. Sys.* 6, 2 (2005), 156–166.
  12. Heath, C., and Luff, P. Media space and communicative asymmetries: preliminary observations of video-mediated interaction. *Hum.-Comput. Interact.* 7, 3 (1992), 315–346.
  13. Hernandez, J., Morris, R. R., and Picard, R. W. Call center stress recognition with person-specific models. In *Proc. of ACHI '11*, Springer-Verlag (Berlin, Heidelberg, 2011), 125–134.
  14. Ishii, H., Kobayashi, M., and Grudin, J. Integration of interpersonal space and shared workspace: Clearboard design and experiments. *ACM Trans. Inf. Syst.* 11, 4 (Oct. 1993), 349–375.
  15. Knight, B. G., Silverstein, M., Mccallum, T. J., and Fox, L. S. A Sociocultural Stress and Coping Model for Mental Health Outcomes Among African American Caregivers in Southern California. *Journals of Gerontology Series* 55, 3 (2000), 142–150.
  16. Kraut, R. E., Miller, M. D., and Siegel, J. Collaboration in performance of physical tasks: effects on outcomes and communication. In *Proc. of CSCW '96*, ACM (New York, NY, USA, 1996), 57–66.
  17. Kuzuoka, H. Spatial workspace collaboration: A sharedview video support system for remote collaboration capability. In *Proc. of CHI'92*, P. Bauersfeld, J. Bennett, and G. Lynch, Eds., ACM (1992), 533–540.
  18. Matthews, G., Campbell, S. E., Falconer, S., Joyner, L., Huggins, J., Gilliland, K., Grier, R., and Warm, J. S. Fundamental dimensions of subjective state in performance settings: Task engagement, distress and worry. *Emotion* 2 (2002), 315–340.
  19. Nguyen, D., and Canny, J. Multiview: spatially faithful group video conferencing. In *Proc. of CHI '05*, ACM (New York, NY, USA, 2005), 799–808.
  20. O'Hara, K., Kjeldskov, J., and Paay, J. Blended interaction spaces for distributed team collaboration. *ACM Trans. Comput.-Hum. Interact.* 18, 1 (2011), 1–28.
  21. O'Neill, J., Castellani, S., Roulland, F., Hairon, N., Juliano, C., and Dai, L. From ethnographic study to mixed reality: a remote collaborative troubleshooting system. In *Proc. of CSCW '11* (2011), 225–234.
  22. Picard, R. W. *Affective Computing*. MIT Press, Cambridge, MA, USA, 1997.
  23. Plarre, K., Raij, A., Hossain, S. M., Ali, A. A., Nakajima, M., al'Absi, M., Ertin, E., Kamarck, T., Kumar, S., Scott, M., Siewiorek, D. P., Smailagic, A., and Wittmers, L. E. Continuous inference of psychological stress from sensory measurements collected in the natural environment. In *IPSN* (2011), 97–108.
  24. Ranjan, A., Birnholtz, J. P., and Balakrishnan, R. An exploratory analysis of partner action and camera control in a video-mediated collaborative task. In *Proc. of CSCW'06*, ACM (New York, NY, USA, 2006), 403–412.
  25. Shostak, B., and Peterson, R. A. Effects of anxiety sensitivity on emotional response to a stress task. *Behavior Research and Therapy* 28, 6 (1990), 513–533.
  26. Stroop, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18 (1935), 643–662.
  27. Vicente, K. J. *Cognitive Work Analysis: Toward Safe, Productive, and Healthy Computer-Based Work*. Lawrence Erlbaum Associates, New Jersey, April 1999.
  28. Whittaker, S. Theories and methods in mediated communication. In *Handbook of Discourse Processes*, A. Graesser, M. Gernsbacher, and S. Goldman, Eds., Erlbaum (2003), 243–286.