Smooth Operator

A Virtual Environment to Prototype and Analyse Operator Support in CCTV Surveillance Rooms

Jonas De Bruyne[1§[0000-0002-6077-6084]], Jamil Joundi[2§[0002-3437-1972]], Jessica Morton[1§[0000-0002-2677-478X]], Niels Van Kets[3§[0001-5495-2240]], Glenn Van Wallendael[3§[0001-9530-3466]], Durk Talsma[4§[0002-3435-2317]], Jelle Saldien[2§[0003-2557-3764]], Lieven De Marez[1§[0001-7716-4079]], Wouter Durnez[1§[0001-8045-8801]], and Klaas Bombeke[1§[0003-2056-1246]]

[1] imec-mict-UGent, Department of Communication Sciences, Ghent University, Miriam Makebaplein 1, 9000 Gent, Belgium
[2] imec-mict-UGent, Department of Industrial Systems Engineering and Product Design, Ghent University, Technologiepark 46, 9052 Zwijnaarde, Belgium
[3] imec-IDLab-UGent, Department of Electronics and Information Systems, Ghent University, Technologiepark-Zwijnaarde 126, 9052 Zwijnaarde, Belgium
[4] Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Gent, Belgium

Abstract. Operators in closed-circuit television (CCTV) control rooms have to monitor large sets of video feeds coming from an ever increasing number of cameras. To assist these operators in their demanding day-to-day tasks, AI-driven support systems accompanied by user-centric interfaces are being developed. However, prototyping these support systems and testing them in operative control rooms can be a challenge. Therefore, in this paper, we present a virtual reality (VR) control room which can be used to investigate the effects of existing and future support systems on operators' performance and behaviour in a fully controlled environment. Important assets of this VR control room include the possibility to subject operators to different levels of cognitive load and to monitor their cognitive-affective states using not only subjective but also behavioural and physiological techniques.

Keywords: virtual reality · user testing · design review · virtual training · WoZ testing · cognitive load

1 Introduction

To assist human operators in modern closed-circuit television (CCTV) control rooms, AI-driven support systems are increasingly deployed to facilitate or automate certain aspects of the operator's task. In parallel with the continuously increasing number of security cameras in modern cities [9,11], control room operators are charged with monitoring an ever growing amount of video feeds. As a result, new control room interfaces – leveraging the power of computer vision algorithms – are designed to display video feeds in such a way that they facilitate the operator's work [12]. However, in order to effectively design such intuitive, AI-driven interfaces, as well as validate the impact they offer towards operator performance and cognitive load, extensive user testing is required.

The construct of cognitive load plays a profound role in cognitive ergonomics. The level of cognitive load in human actors is defined by multiple antecedents [16]. These antecedents are elements of either cognitive work demands or human cognitive architecture. As an example, increased task complexity can increase operators' cognitive load as it places higher demands on their limited working memory capacity. However, the effects of these antecedents on cognitive load can be moderated by numerous factors both related to the individual (e.g., task experience) and to the task (e.g., involving the use of assistive technology). Especially in work contexts, (sub-)optimal cognitive load can impact the actor's work behaviour (e.g., speed and accuracy) and thus overall work performance.

Cognitive load as a multi-dimensional construct [20] is assessed through a wide range of procedures. In literature, adapted versions of the NASA-TLX questionnaire [8] are regularly used as a method to obtain a measure of perceived cognitive workload (e.g., [6,5]). However, other than obtaining a subjective measure through questionnaires, researchers have studied physiological correlates of cognitive load ranging from electrical brain activity [1] to electrodermal activity [13] to pupil dilation (see, [19]). As a study of Vanneste et al. [17] demonstrates, the accuracy of cognitive load assessment increases by using a multimodal approach that includes multiple measures.

To effectuate a multimodal approach to assess cognitive load during CCTV monitoring, immersive virtual reality (VR) appears an advantageous testing environment. By means of a head-mounted display (HMD), people can watch and interact with an immersive virtual environment. Building VR simulators offers multiple advantages over traditional approaches. First, with regards to the present project, it is less time-consuming to prototype and evaluate simulated operator support systems in VR than it is to build fully operational supportive systems in a physical environment. Moreover, virtual environments allow the implementation of the Wizard of Oz prototyping approach [3] to test initial ideas without the need of developing automated systems. This means that researchers can simulate automated systems by manually steering in-scene events so that the participant believes these events to occur automatically. Second, by using VR, the experimental environment is fully controlled. Researchers can control for lightning conditions, background noise, the presence of colleagues, etc. This allows investigating the effects of experimental manipulations (e.g., the addition of supportive systems in surveillance rooms) during numerous different circumstances. In addition, it is possible to log all the participant's actions and interactions. Third, state-of-the-art HMDs with built-in eye-trackers continuously log eye-related indices. As an example, insights on participants' gaze as well as indirect indicators of cognitive load (pupil dilation and blink rate) can be derived from this data. Finally, results of a study by Tauscher et al. [14] demonstrate that, with some optional minor modifications, it is possible to combine EEG and

VR. Furthermore, researchers have already been able to discriminate between different levels of cognitive load using a classical n-back task in an interactive VR environment regardless of the increase in muscle tension and activity as a result of the interactive environment [15].

In the present paper, we present a VR environment to test existing and next-generation AI-based supportive systems or new interfaces in CCTV control rooms. To this extent, we developed a virtual CCTV control room in which an operator's job is simulated using immersive VR technologies. In a pilot experiment, we tested whether we can influence participants' cognitive load in order to create different working conditions. To do so, we introduced a dual task paradigm that is commonly used by experimental psychologists to gain insight into multitasking processes. Participants – who wore an HMD in order to watch the virtual control room and interact with it – had to perform a simplified monitoring task (primary task) while from time to time they were interrupted by auditory requests which they had to respond to (secondary task). This secondary task either consisted of low demanding task rules and long response-stimulus intervals (RSIs; i.e., the time between a participant's response to a trial and the presentation of the stimuli of the next trial) or high demanding task rules and short RSIs. We will refer to these conditions as the low and high demand condition respectively. Manipulating these two features (task difficulty and RSIs) has previously been shown to increase cognitive load [7,18]. Therefore, we hypothesised higher cognitive load in the high demand condition compared to the low demand condition. To investigate this with a multimodal approach, task performance, pupil size, blink rate and subjective reports on perceived cognitive load (using an adapted version of NASA-TLX) were selected as measures for cognitive load, as they are frequently used to measure cognitive load and represent behavioural, physiological and subjective techniques.

2 Virtual Reality CCTV Control Room

The test environment was developed in Unity (version 2019.4.3f1) using a pre-existing police control room asset, which was modified according to the experiment's design needs. The VRTK framework (vrtoolkit.readme.io) was used for in scene interactions. The resulting virtual control room is equipped with a videowall consisting of 8 large screens and two desks with 3 monitors each (fig. 1). One of the monitors of the operator's (i.e., the participant) personal workspace is used as a response screen with buttons that can be pressed using a pointer and the trigger button of the controller. On the remaining screens, a wide range of various types of content can be rendered such as (interactive) city maps, surveillance camera footage and visualisations of data streams. Moreover, the screen set-up can easily be modified according to assistive technology requirements and the study's needs. Also, other than prerecorded camera footage, real-time video streams can be integrated in the scene through embedded web browsers and virtual desktop screens in Unity.

To manipulate cognitive load in order to test supportive systems during different circumstances, a secondary task with varying difficulty and response-stimulus intervals (RSI) is introduced. In the scene, a walkie talkie radio is present, which is the audio source for the presentation of the auditory stimuli of the secondary task. This secondary task consists of the auditory presentation of digit sequences (max. 6 digits) which require a different response according to the condition participants are in. In the low demand condition, they are asked to click on the on-screen response button corresponding to the last heard digit. In contrast, in the high demand condition, the response depends on the number of digits in the sequence. When the sequence consists of an odd number of digits, participants have to click the last two heard digits. When the number of digits in the sequence is even, participants have to click the first two digits they heard. In addition, the RSI of the secondary task varies over both conditions. In the low demand condition, the RSI varies ad random between 25 and 30 seconds whereas in the high demand condition the RSI varies between 2 and 7 seconds. During testing, performance on this task is instantly calculated. Furthermore, all operator actions and UI interactions are automatically logged. Additionally, eye gaze, pupil dilation and blink rate are measured continuously.

Fig. 1: Overview of the virtual control room.

3 Pilot experiment: Cognitive Load Manipulation

3.1 Method

Participants. Thirteen participants with normal or corrected to normal vision participated to the experiment (3 female, M_{age} = 21.92, SD_{age} = 0.95). All participants signed informed consent and participated voluntarily to the study. Therefore, they were not credited nor paid.

Materials and Equipment. The VR setup consisted of a computer running SteamVR (v.1.14.16) and an HTC VIVE Pro Eye. The HMD and the controllers were tracked by two Vive SteamVR Base Stations 2.0. The HMD's built-in eye-tracker and the Vive Eye-tracking Software Development Kit (SDK) SRanipal was used to obtain eye-tracking measures (incl. pupil sizes). This built-in eye-tracker had a sampling rate of 120 Hz. However, in this experiment, we recorded the eye-tracking data at 50 Hz. Perceived workload was assessed at the end of every block using an adapted version of the NASA-TLX [8]. The NASA-TLX is a well-known assessment instrument which results in an indication of perceived workload on six domains of task requirements (e.g., mental demand, physical demand etc.).

The Primary Task. A simplified video monitoring task was used as the primary task. Participants watched surveillance videos that were presented on the eight screens of the videowall and on the left and centre monitor on the personal desk. These videos merely served to create a realistic operator setting and, thus, were irrelevant to the task. However, one of these screens turned green after a variable interval (5 – 10 seconds). The task was to press the 'detected' button on the response screen whenever this happened before the response deadline of 4 seconds was exceeded. Both accuracy and reaction times were measured.

Procedure. Participants signed informed consent upon arrival and were then given a short instruction on the experiment. After the brief introduction, they were helped to put on the headset. Throughout the whole experiment, participants were seated and wore one controller in their hand of preference. Next, all instructions were presented on the centre monitor of the personal workspace in the virtual control room. After reading these on-screen instructions in VR, each participant performed a practice block on the primary task. Next, half of the participants were instructed on the low demand task rules and the other half on the high demand task rules, followed by a practice block on the secondary task. Subsequently, a practice block for the dual task was provided. This was followed by three experimental blocks of approximately 5 minutes each, separated by self-paced breaks. After these blocks, the task rules of the secondary task changed. Participants who first performed the low demand secondary task were now instructed on the high demand secondary task and vice versa. Next, a practice block to get familiar with these new task rules was performed. Finally, participants were presented with another set of three dual task experimental blocks, now using the new task rules. At the start of every block, participants performed an eye-tracker calibration procedure. At the end of each block, participants were asked to unmount the HMD and fill in the questionnaire on perceived cognitive load.

3.2 Results

The data was analysed with one within-subject factor (i.e., task demands – high or low). All data pre-processing was performed in Python 3 and models were constructed in R using the lme4 package [2] specifying a random intercept for each participant. Reported p-values for the linear mixed effects models are corrected values using the Kenward-Rogers correction [10].

Performance. In the analysis of accuracy on the primary task, we found that people performed better in the low demand condition (M = 93%, SD = 2%) compared to the high demand condition (M = 99%, SD = 6%), $F(1, 12) = 22.02$, $p < 0.001$. Moreover, participants were slower to detect the highlighted screen in the high demand (M = 1194.53 ms, SD = 709.88 ms) compared to the low demand condition (M = 905.93 ms, SD = 514.64 ms), $F(1, 12) = 8.32$, $p = 0.014$. The same trend was found for accuracy on the secondary task where participants performed worse in the high demand (M = 67%, SD = 13%) compared to the low demand condition (M = 95%, SD = 7%), $F(1, 12) = 165.00$, $p < 0.001$.

Cognitive load. Subjective cognitive load as measured by the adapted version of NASA-TLX was higher in the high demand condition (M = 52.26, SD = 13.74) than in the low demand condition (M = 19.70, SD = 11.86), $F(1, 12) = 81.24$, $p < 0.001$. This effect was found for pupil size as well. Participants had larger mean pupil sizes during the high demand condition (M = 3.17 mm, SD = 0.42 mm) compared to the low demand condition (M = 2.99 mm, SD = 0.28 mm), $F(1, 12) = 11.79$, $p = 0.004$. This effect reflects higher cognitive load when the task rules of the secondary task were more difficult. Blink rate also differed significantly between conditions. Blink rate was higher when the task rules were difficult (M = 18.10 blinks/min., SD = 12.65 blinks/min) relative to the low demand condition (M = 10.50 blinks/min., SD = 4.82 blinks/min.), $F(1, 12) = 7.74$, $p = 0.017$.

4 Discussion

In this project, we developed a VR CCTV control room in order to test the influence of supportive systems and interfaces on operators' performance, behaviour and cognitive load while they are subject to different levels of cognitive load. This allows researchers to gain insight into the effects of assistive technology during different cognitive states of the operator. To validate the cognitive load manipulation, we conducted a pilot experiment in which we investigated whether we can manipulate cognitive load by manipulating difficulty and temporal features of a secondary task. In line with previous research, the results of this experiment highly suggest an increase in cognitive load when the task rules for the secondary task became more difficult and the frequency of secondary task trials was higher. In particular, we found a significant main effect of task demands on reported cognitive load. Moreover, the same effect was found in the behavioural results. Specifically, performance (i.e., accuracy and reaction time) on the primary task decreased with increasing task demands for the secondary task. Since the primary task was identical across conditions, these results indirectly reflect an increase in cognitive load as a result of the demand manipulation in the secondary task. Next, the same effect was found in the analysis of pupil size. As larger pupil sizes reflect higher cognitive load, this result also suggests an increase in cognitive load driven by the secondary task manipulation. For blink rate, which we included as an exploratory measure, a main effect of task demands was found. Specifically, blink rate was higher in the high demand condition compared to the low demand condition. However, since multiple drivers can underlie this effect, we cannot surely attribute this effect to either an increase in mental activity needed to perform the dual task or an increase in using mental rehearsal [4] as a strategy to accomplish the secondary task goals.

Future experiments within the current project will investigate the effects of a specific support system on operators' behaviour, performance and cognitive load. Furthermore, we will explore the gain of adding EEG as a measure of cognitive load. In particular, we will look into parietal alpha suppression, increases in frontal theta power and the cognitive load index [19]. Additionally, using the built-in eye-tracker, we will explore how we can gain insights in visual search and active exploration of monitored video footage using eye gaze data. An example of such a measure would be the proportion of time spent looking at suggested video footage pushed by a camera selection algorithm compared to the proportion of time spent looking at the other videos that are presented on, for example, the videowall. Additionally, from a user testing perspective, it remains important to take into account the perceived usability of professional control room operators when testing new assistive technology. Therefore, this VR tool can also be used to set up qualitative experiments tapping into the subjective experience of users. In sum, these experiments will yield a demonstration of how the VR control room can be used and investigate possible additional dependent measures that might offer insights in how control room operators are affected by introducing AI-based supportive systems.

References

1. Antonenko, P., Paas, F., Grabner, R., Van Gog, T.: Using electroencephalography to measure cognitive load. Educational Psychology Review **22**(4), 425–438 (2010)
2. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. arXiv preprint arXiv:1406.5823 (2014)
3. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of oz studies—why and how. Knowledge-based systems **6**(4), 258–266 (1993)
4. De Jong, P.J., Merckelbach, H.: Eyeblink frequency, rehearsal activity, and sympathetic arousal. International Journal of Neuroscience **51**(1-2), 89–94 (1990)
5. Di Nocera, F., Camilli, M., Terenzi, M.: A random glance at the flight deck: Pilots' scanning strategies and the real-time assessment of mental workload. Journal of Cognitive Engineering and Decision Making **1**(3), 271–285 (2007)
6. DiDomenico, A., Nussbaum, M.A.: Interactive effects of physical and mental workload on subjective workload assessment. International journal of industrial ergonomics **38**(11-12), 977–983 (2008)
7. Haga, S., Shinoda, H., Kokubun, M.: Effects of task difficulty and time-on-task on mental workload. Japanese Psychological Research **44**(3), 134–143 (2002)

8. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: Advances in psychology, vol. 52, pp. 139–183. Elsevier (1988)
9. Hollis, M.E.: Security or surveillance? examination of cctv camera usage in the 21st century (2019)
10. Kenward, M.G., Roger, J.H.: Small sample inference for fixed effects from restricted maximum likelihood. Biometrics pp. 983–997 (1997)
11. Norris, C., McCahill, M., Wood, D.: The growth of cctv: a global perspective on the international diffusion of video surveillance in publicly accessible space. Surveillance & Society **2**(2/3) (2004)
12. Pelletier, S., Suss, J., Vachon, F., Tremblay, S.: Atypical visual display for monitoring multiple cctv feeds. In: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems. pp. 1145–1150 (2015)
13. Setz, C., Arnrich, B., Schumm, J., La Marca, R., Tröster, G., Ehlert, U.: Discriminating stress from cognitive load using a wearable eda device. IEEE Transactions on information technology in biomedicine **14**(2), 410–417 (2009)
14. Tauscher, J.P., Schottky, F.W., Grogorick, S., Bittner, P.M., Mustafa, M., Magnor, M.: Immersive eeg: Evaluating electroencephalography in virtual reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). pp. 1794–1800. IEEE (2019)
15. Tremmel, C., Herff, C., Sato, T., Rechowicz, K., Yamani, Y., Krusienski, D.J.: Estimating cognitive workload in an interactive virtual reality environment using eeg. Frontiers in human neuroscience **13**, 401 (2019)
16. Van Acker, B.B., Parmentier, D.D., Vlerick, P., Saldien, J.: Understanding mental workload: from a clarifying concept analysis toward an implementable framework. Cognition, technology & work **20**(3), 351–365 (2018)
17. Vanneste, P., Raes, A., Morton, J., Bombeke, K., Van Acker, B.B., Larmuseau, C., Depaepe, F., Van den Noortgate, W.: Towards measuring cognitive load through multimodal physiological data. Cognition, Technology & Work pp. 1–19 (2020)
18. Veltman, J., Gaillard, A.: Physiological workload reactions to increasing levels of task difficulty. Ergonomics **41**(5), 656–669 (1998)
19. van der Wel, P., van Steenbergen, H.: Pupil dilation as an index of effort in cognitive control tasks: A review. Psychonomic bulletin & review **25**(6), 2005–2015 (2018)
20. Young, M.S., Brookhuis, K.A., Wickens, C.D., Hancock, P.A.: State of science: mental workload in ergonomics. Ergonomics **58**(1), 1–17 (2015)