Comparing Google Glass with Tablet-PC as Guidance System for Assembling Tasks

Matthias Wille, Sascha Wischniewski

Unit Human Factors, Ergonomics
Federal Institute for Occupational Safety and Health
Dortmund, Germany
wille.matthias@baua.bund.de
wischniewski.sascha@baua.bund.de

Philipp M. Scholl, Kristof Van Laerhoven

Embedded Sensing Systems
Technische Unversität Darmstadt
Darmstadt, Germany
scholl@ess.tu-darmstadt.de
kristof@ess.tu-darmstadt.de

Abstract — **Head mounted displays (HMDs) can be used as an guidance system for manual assembling tasks: contrary to using a Tablet-PC, instructions are always shown in the field of view while hands are kept free for the task. This is believed to be one of the major advantage of using HMDs. In the study reported here, performance, visual fatigue, and subjective strain was measured in a dual task paradigm. Participants were asked to follow a toy car assembly instructions while monitoring a virtual gauge. Both tasks had to be executed in parallel either while wearing Google Glass or using a Tablet-PC. Results show slower performance on the HMD but no difference in subjective strain.**

Keywords — Head mounted displays; assistance; strain; performance; assembling task; monitoring task; dual task paradigm; Google Glass; Tablet-PC

I. INTRODUCTION

The idea to support manual tasks with a head mounted display (HMD) is by far not new. The mobility of an HMD system and the fact that information is displayed within the field of view while hands are kept free for the manual task makes an HMD theoretically an ideal companion to support workers in e.g. construction, assembling, maintenance or car-workshops. Through the years, many projects and studies have investigated in the use of HMD in industrial tasks (e.g. projects ARVIKA [2] or wearIT@work [4]; for an overview on related work see also [7]). However, for a long period the advantages of an HMD were in contrast to some ergonomic constraints, especially regarding their wearing comfort. HMDs were heavy, mounted on a helmet-like head carrier and connected with a lot of wiring to a device worn on the body. Those HMDs were expensive, industrial products, and supported only a limited amount of special applications. In recent years however, companies started to develop wearable, affordable, consumer-grade HMDs. HMDs are transforming from "head mounted displays" to "data glasses" and are now on the cusp of mass-market and therefore, many new applications and fields of use will emerge.

II. SCOPE OF THE PROJECT

The Federal Institute for Occupational Safety and Health (BAuA) in Germany started a project which investigates HMDs as an assistance system for industrial manual tasks. The project consists of different work packages with diverse goals: a task analysis in the field points out tasks or task characteristics that benefit most by the support from an HMD. A laboratory study focused on the physical strain (visual and muscular) while assembling and dissembling a real car engine where instructions were shown on either an HMD or a wall mounted monitor [5, 6]. Another study focused on the psychological strain in a dual task paradigm where subjects have to assemble a toy car and monitor a virtual gauge in parallel. The content was presented on either an HMD or Tablet-PC [8]. Both studies have in common that they investigate prolonged working over four hours to get insights into the strain development over time with different assistance technologies. Even though there are many studies showing the benefits of using HMDs, the period of wearing was mostly short and yet there is little known about long term effects. The project is also not looking into augmented reality (AR) approaches, mainly because the required tracking opens a complete additional field. While AR would be a nice feature for assistance on assembling real world objects we do not see an area-wide market penetration of that technology within the next years based on the technical and organizational aspects of marker based tracking. So the focus of the project is to display work instructions and checklists with an HMD, but without any AR applications.

III. STUDY

The study reported here is a replication of the laboratory study with the dual-task paradigm [8]. This replication was done with Google Glass as one of the new lighter consumer HMDs while the original study was conducted with a MAVUS HMD of the Heitec Company which is a typical example of heavier industrial HMDs. A comparison might show which parts of the results about strain and performance are based on the technology itself and which parts are more based on the implementation of this technology into concrete hardware. Due to organizational aspects, this time the participants worked only about 30 minutes in the described dual-task paradigm. Before they executed two other studies using Google Glass which are not reported here. So the overall wearing time of the HMD was about two hours with the study taking place in the last half hour. The study is still work in progress, so results are initial.

IEEE
computer
society

A. Design

This study follows a between subjects design where half of the people worked with Google Glass and the other worked with a Tablet-PC. All participants had worked for about one hour with Google Glass in different studies prior to the beginning of this trial. The data was analyzed in SPSS 21 using ANOVA with mixed design: "display" was a between subject factor while the different timestamps for questionnaires were within subject factors.

B. Participants

20 subjects participated in the ongoing study so far, aged between 18-67 years. 10 subjects, aged between 21-59 years (Mean = 37.20, SD = 13.323, 4 male / 6 female), were using Google Glass. The other 10 subjects were using a Tablet-PC and were aged between 18-67 years (Mean = 40.00, SD = 18.667, 7 male / 3 female). Age is distributed equally in both groups while sex is not. 18 of the subjects already participated in the previous study [8] and 2 subjects (both within the Tablet-PC group) had no experience with the task and technology.

C. Tasks

This study uses a dual task paradigm where participants had to fulfill two tasks in parallel. Both tasks were introduced as equally important and should be handled as fast and accurate as possible.

On one side, the participants had to assemble a toy car based on Lego-Technic where the instruction was given step-by-step in a graphical form: On each step some building bricks were added to a model which was getting more and more complex. The number of completed assembling slides within the first 25 minutes was an indicator for performance. This can be done because the complexity of all assembling slides is comparable.

To emphasize the mobility aspects, subjects had to pick up the building bricks in one location and assemble them in another place a few meters beside. So while working with a Tablet-PC participants had to carry it manually with them each time and while working with an HMD this was done automatically.

On the other side, participants had to observe a monitoring task presented in parallel to the right of the assembly slides. Three bars slowly but continuously varied their length and switched color from time to time between blue and red. See work content presented in figure 1. All this variation took place randomly. On average every 94.45 seconds (SD = 61.968) a reaction to length was requested and every 106.44 seconds (SD = 48.639) a reaction to color.

Participants had to react to a color-change and to a change in position of the longest bar by either saying "bar changed" in Google Glass group or a double tap in the Tablet-PC group. Reaction to color-change is a typical monitoring task and includes also a perceptional "pop-out" effect because a large area of the display changes its color at once. The reaction to changes in position of the longest bar has no such pop-out effect and is harder to detect. A written feedback was given above the bars indicating the last confirmed color and position.
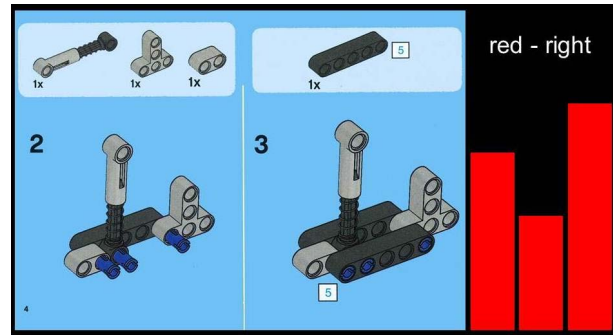

Figure 1. Work content as displayed on Google Glass or Tablet-PC: Assembly slide based on Lego Technic (left) and monitoring task (right).

D. Interaction

All interactions on the Tablet-PC were done via touch: A swipe to the left for the next assembling slide or to the right for the previous slide and a double tap as reaction to the monitoring tasks. All interactions on Google Glass were done with speech commands: "next slide" and "previous" for changing assembling slides and "bar changed" as reaction to the monitoring task. Additionally, subjects had the possibility to zoom into the assembling image with "zoom image". The zoom showed a two times enlarged image while the presented part was chosen with head movement measured by internal sensors. To shrink the image the speech command "scale down" was used.

E. Apparatus

The HMD in this study was Google Glass. It is a 640 x 360 pixel see through display mounted on a spectacle frame weighting 50 grams. It was connected to a battery extension pack to enable continuous displaying of information for about 2 hours. The Tablet-PC was a Samsung Galaxy SM-T210 with a resolution of 1024x600, a size of 17.8 cm (7'') and a weight of 300 grams. In the previous study [8] both devices were industrial tools, this time both devices are consumer products.

F. Procedure

In the beginning all participants were introduced to Google Glass and took part in another study where they extracted the DNA of onions and tomatoes while instructions were displayed on the HMD. This lasted for about one hour and was followed by a break where Google Glass was taken off and shut down for cooling. After the break subjects were diverted into the two groups and half of them continued with a Tablet-PC. In the Google Glass group the speech commands were practiced 5 times each. In both groups a baseline measure of the monitoring task followed where participants only reacted for 5 minutes to the bars without a secondary task. This was done to assert the reliability of the speech input modality. Google Glass users raised their hand whenever the system did not understand their bar changed command which was noticed by the investigator. Subjects with a detection quote less than 100% were excluded from this analysis. During the study several questionnaires were asked at different times to get insight into the development of strain and visual fatigue over time. Before

starting, in the break after one hour and at the end of each trial, the Visual Fatigue Questionnaire (VFQ) [1] was examined. The Rating Scale of Mental Effort (RSME) [9] where participants rate their subjective strain on a one-dimensional scale from $0 - 150$ was used 4 times during the experiment (every half hour). The NASA-TLX [3, 4] and an interview were done at the end.

IV. RESULTS

Although results show a larger number of processed assembly slides (and therefore completed assembly steps) with the Tablet-PC (see figure 2) this effect is not significant $[F(1, 18) = 1.887, p = .186]$. The direction of more processed slides on Tablet-PC is the same as in the previous study [8], where it was significant and the gap between both groups might enlarge with longer collection times. The zoom function was only available on the HMD and could be an additional time consuming factor in working. However, participants used that function only rarely (mean = 4.4; SD = 4.38) and some did not use it at all. No correlation could be found between number of zooms and processed slides ($r = .10$, $p = .783$).

In the monitoring task the hit rate on color changes shows a significant effect of dual task (compared with baseline) $[F(1, 18) = 8.692, p < .05]$. And it showed a statistical tendency on the between subject effect regarding used display $[F(1, 18) = 3.104, p = .095]$. The trend goes to better hit rates on the Tablet-PC. An interdependency between dual task and display could not be found $[F(1, 18) = .434, p = .518]$. The reaction time shows a similar effect: It is significantly higher during dual task $[F(1, 18) = 42.217, p < .001]$, but no effect of display $[F(1, 18) = .456, p = .508]$ and no interdependency between those factors $[F(1, 18) = .036, p = .852]$.

The reaction to position change of the longest bar shows no effect in the hit rate when comparing dual to single task $[F(1, 18) = 2.816, p = .111]$, no effect of display $[F(1, 18) = 2.072, p = .167]$ and no significant interdependency $[F(1, 18) = .047, p = .830]$. The reaction time shows a significant increase in the dual task compared to baseline $[F(1, 18) = 14.097, p < .001]$, but no effect of display $[F(1, 18) = .228, p = .638]$ or interdependency $[F(1, 18) = .320, p = .579]$.

It has to be stated that the hit rate on color change is not 100% trust worthy: A hidden process in Android wrote erroneous color change events into the results matrix which resulted in an increased number of misses even though the participant was not able to react. The problem occurred on both devices in the same way and may increase the misses about 10%. This way the hit rate on color change is slightly worse than the hit rate on length change (which is unusual because of the pop out effect of color change).

The VFQ shows on the item "difficulties to see clearly" significant higher ratings for the HMD $[F(1, 18) = 4.675, p < .05]$, an increase over time $[F(2, 17) = 6.915, p < .05]$ and an interdependency $[F(2, 17) = 6.596, p < .05]$ indicating a higher increase on the HMD. Furthermore the item "mental fatigue" shows similar effects: Higher values for HMD $[F(1, 17) = 5.198, p < .05]$, an increase over time $[F(2, 16) = 7.599, p < .05]$ and an interdependency

$[F(2, 16) = 4.635, p < .05]$. But all those values are on a low overall level with a maximum of 3 on a 10 point scale. Headache and neck pain showed no significant effects this time, while they had clearly significant higher values for the HMD in the previous study [8].

The subjective strain measured by RSME shows a significant increase over time $[F(1, 17) = 10.148, p < .05]$ but no effect of display $[F(2, 17) = .715, p = .409]$ and no interdependency $[F(2, 17) = 1.068, p < .316]$. The NASA-TLX shows slightly higher values for the HMD (mean = 69.85; SD = 20.47) than for the Tablet-PC (mean = 54.45; SD = 20.27) but this fails statistical tendency sharply $[F(1, 18) = 2.857, p = .108]$.
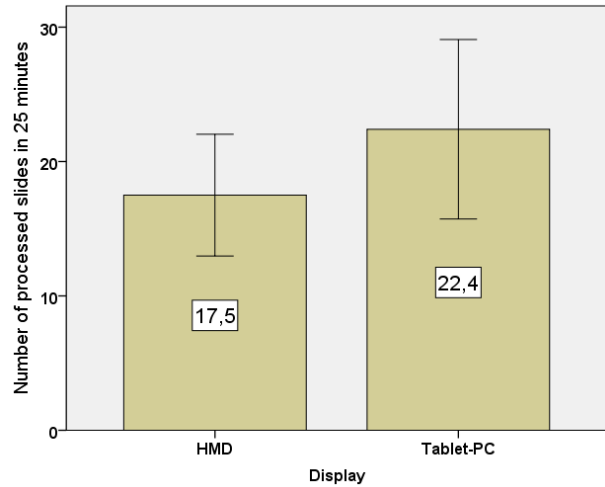

Figure 2. Number of processed assembly slides in 25 minutes on both displays. Whiskery show 95% confidence interval.

TABLE I. HIT RATE AND REACTION TIME FOR COLOR CHANGE MONITORING TASK ON BOTH DISPLAYS

Color	Google Glass		Tablet-PC	
	Mean	*SD*	*Mean*	*SD*
hit rate % (baseline)	78,21	21,35	89,16	18,44
hit rate % (dual task)	59,24	32,81	77,12	23,51
RT seconds (baseline)	2,68	1,49	1,98	0,93
RT seconds (dual task)	7,06	1,92	6,63	4,09

TABLE II. HIT RATE AND REACTION TIME FOR LENGTH CHANGE MONITORING TASK ON BOTH DISPLAYS

Length	Google Glass		Tablet-PC	
	Mean	*SD*	*Mean*	*SD*
hit rate % (baseline)	84,40	21,68	94,17	12,45
hit rate % (dual task)	75,61	13,90	82,74	24,42
RT seconds (baseline)	1,96	0,83	1,96	1,26
RT seconds (dual task)	8,16	3,06	10,37	13,09

In the interview all subjects that took part in the previous study [8] were excited by Google Glass. They point out to feel less stressed with this new HMD and also they liked the display and the positioning of the display. It is interesting to mention that those subjects who had not experienced the other HMD complained about discomfort with Google Glass.

V. DISCUSSION

These preliminary results show that some of the effects found previously [8] are rather caused by the hardware realization of HMDs than by the technique itself. While subjective strain measures were significantly higher with the industrial HMD compared to the Tablet-PC this effect does miss significance with Google Glass, although the trend of mean values is still the same. Furthermore, headache and neck pain is not caused by the HMD anymore. Participants still experienced a faster increase in mental fatigue while working with an HMD and also found it increasingly difficult to see clearly if working with it over several hours which did not happen when using a Tablet-PC.

In the previous study significantly worse hit rates on both reaction tasks were seen on HMDs compared to a Tablet-PC. This is not the case anymore, although the same tendency can be seen in the mean values. The opposite, more accurate reactions on HMDs given by the fact that values are always in the field of view, is also not supported by our observations. This points out the selective nature of perception and is important for all designers of HMD-based interactions: People will not perceive the information only because you display it in front of the eye - the problem with focus of attention is the same as on a Tablet-PC or other handheld device. Reaction times however did not vary significantly between both display technologies. This was the same in the previous study.

The assembling performance did not show a significant difference between both displays. However, mean values still show more processed slides with the Tablet-PC and it is imaginable that if the sampling period is enlarged the gap will increase too and the effect will become as significant as it was in the 4 hour longing previous study. The question why people work slower with an HMD stays unanswered but it might be based on the smaller display as they report also difficulties to see clearly or because of unfamiliarity with HMDs. It is not based on the zooming function as an extra time-consuming procedure given only on the HMD because we found no correlation between number of zooms and number of processed slides.

As those new consumer HMDs perform better, the question also pops up, why / if we still need those mentioned industrial HMDs. One major difference of Google Glass is that it is not designed to continuously display information (for example during a work shift in a steel mill). Besides the energy problem it is getting warm and decreases its processing speed for cooling. Also it is not robust enough and tolerant enough to temperature and air humidity for some applications. HMDs designed for industrial environments are still necessary, but will hopefully profit from the ergonomic progress the consumer HMDs have taken during the last years.

REFERENCES

[1] Bangor, A. W. Display technology and ambient illumination influences on visual fatigue at VDT Workstations. PhD thesis, Virginia Polytechnic Institute and State University, USA, 2000.

[2] Friedrich, W., Jahn, D., & Schmidt, L. ARVIKA - Augmented Reality for Development , Production and Service. In ISMAR, 2002, pp. 3–4.

[3] Hart, S.G., & Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: P.A. Hancock & N. Meshkati (Eds.) Human mental workload. Amsterdam: North-Holland, 1988, pp. 139-183.

[4] Kunze, K., Wagner, F., Kartal, E., Kluge, E.M., and Lukowicz, P. Does Context Matter ? - A Quantitative Evaluation in a Real World Maintenance Scenario, in Pervasive '09: Proceedings of the 7th International Conference on Pervasive Computing, 2009, pp. 372-389.

[5] Theis, S., Alexander, T., Mertens, A., Wille, M. & Schlick, C. Younger beginners, older retirees: Head-Mounted Displays and Demographic Change. Applied Human Factors and Ergonomics (AHFE), Krakau, Poland, 2014, accepted.

[6] Theis, S., Alexander, T., Mayer, M. & Wille, M. Considering ergonomic aspects of head-mounted displays for applications in industrial manufacturing. In: Duffy, V.G. (Ed.), Digital Human Modeling and Applications in Health, Safety, Ergonomics, and Risk Management (DHM/HCII) 2013, Part II, Lecture Notes in Computer Science 8026. Berlin: Springer, 2013, pp. 282-291.

[7] Völker, K., Adolph, L., Pacharra, M., Windel, A. Datenbrillen – Aktueller Stand von Forschung und Umsetzung sowie zukünftige Entwicklungsrichtungen. In: Gesellschaft für Arbeitswissenschaften e.V.: Bericht zum 56. Arbeitswissenschaftlichen Kongress vom 24. – 26. März 2010 an der Technischen Universität Darmstadt, 2010, pp. 61-65.

[8] Wille, M., Grauel, B. & Adolph, L. Strain caused by head mounted displays. In: De Waard, D., Brookhuis, K., Wiczorek, R., Di Nocera, F., Barham, P., Weikert, C., Kluge, A., Gerbino, W., and Toffetti, A., (Eds.) Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2013 Annual Conference, 2013. Online available: http://www.hfes-europe.org/books/proceedings2013/Wille.pdf [May 2014].

[9] Zijlstra, F.R.H.. Efficiency in workbehaviour: an approach for modern tools. PhD thesis, University of Delft. Soesterberg, The Netherlands: Institute for Perception TNO, 1993. Online available: http://repository.tudelft.nl/view/ir/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f/ [May 2014].