R WE FOOLING OURSELVES:
DOES THE OCCLUSION TECHNIQUE SHORTCHANGE R ESTIMATES?

Christopher A. Monk and David G. Kidd
George Mason University
Fairfax, VA, USA
E-mail: cmonk@gmu.edu

Summary: The occlusion technique was originally used to evaluate the cognitive demands of the roadway. Recently, the occlusion technique has been used as a cost-effective tool for assessing the visual demand of in-vehicle devices. Occlusions simulate glances from an in-vehicle device to the roadway by interrupting visual sampling. However, occluding the in-vehicle device does not impose any additional cognitive demand on the participant like true glances back to the roadway. The purpose of this study was to compare standard no-task occlusions with occlusions requiring participants to perform a visual-motor tracking task. Results suggest that overestimates of resumability may result by not including a task during occlusions. Furthermore, estimations of visual demand based on individual post-occlusion resumption times may provide a more precise measure of transition costs and resumability than measures based on Total Shutter Open Time.

INTRODUCTION

The visual occlusion technique was originally developed by Senders et al. (1967) to assess driver workload. Recently, the technique was adopted as a means to estimate the visual demand of in-vehicle devices (see Gelau & Krems, 2004) and has been introduced by the International Standards Organization (ISO) as a standard for assessing visual demand from in-vehicle systems (ISO, 2007). With the increasing incorporation of advanced driver assistance technologies into the driver's visual, cognitive, and physical sphere, the need arose to provide designers and evaluators with a means of assessing whether these new devices imposed too great a demand on the driver's visual resources. For example, designers who needed to determine whether new destination entry interface logic would allow drivers to complete the task more efficiently.

Most implementations of the visual occlusion technique that assess driver distraction do not impose cognitive demands on participants during the occluded periods. It is our assertion that the no-task occlusion method disregards an important aspect of estimating task "resumability" as outlined by the ISO standard (ISO, 2007). The objective of this study was to compare the standard no-task occlusion method with a condition requiring participants to perform a visual-motor tracking task during occluded periods. A difference in results between these conditions would suggest that the standard method is missing an important aspect of estimating resumability. In addition, a more detailed examination of the resumption costs associated with occluded performance was conducted to further refine estimation of visual demand.

Before introducing the theoretical motivations for incorporating a task during the occluded periods, a more detailed explanation of the visual occlusion method as specified by the ISO

standard is required, along with the recent research findings that support its use for estimating visual demand. The rationale for using an estimation method rather than direct measures of visual demand relates to the ease, cost, and timeliness of conducting studies that measure eye movements. Designers and evaluators of advanced in-vehicle devices typically do not have the resources or time to conduct expensive and time-intensive eye-tracking studies that would directly measure visual demand. The visual occlusion technique was adopted as a tool for designers and evaluators to gain estimates of visual demand with a cheaper and faster method that maintained some level of methodological rigor. In fact, the visual occlusion method was sought out partly in response to dissatisfaction with the so-called "15-second Rule" (Green, 1999) for its lack of supporting data (see Bauman, Keinath, Krems, & Bengler, 2004) and face validity. The recent publication of the ISO Standard (ISO, 2007) further demonstrates the human factors community's commitment to the visual occlusion technique as an important screening tool for developers of in-vehicle systems.

The intent of the occlusion method is to simulate drivers' visual sampling of the road while interacting with an in-vehicle device. This simulation is accomplished by occluding vision every one to two seconds while the participant executes the task. When performing in-vehicle tasks while driving, visual attention is shared between the two tasks, resulting in brief samples of the roadway and the interface in an interleaved fashion. In the occlusion technique, goggles that have great temporal control over lens transparency are typically worn to enable precise occlusion intervals. The periods in which the lenses are occluded represent the times when the driver would be looking at the road. If a task can be completed efficiently with intermittent brief glances, then it is considered to be easy to resume after interruptions. The assumption is that highly resumable task interfaces are therefore reasonably suitable for use while driving.

The measure used to determine if an in-vehicle task is resumable is a ratio of the time the task interface is visually available to the "driver" and the total time to complete the task uninterrupted. The resulting "resumability" ratio R is therefore an indication of whether periodically shifting attention away from the in-vehicle interface (i.e., the occluded periods) results in added time to complete the task. This added time would then be considered a "cost" when performing the task while driving. There are formal terms and values specified in the standard that will be detailed as follows. *TSOT* is the "total shutter open time," or the sum time that the participant can see the interface while completing the task. TTT_{unoccl} is the total task time for completing the same task without occlusions or interruptions. R is calculated as specified in equation 1. R-values greater than 1 indicate an added cost to completing the in-vehicle task while driving—a resumption cost. R-values near 1 indicate minimal effect of performing the task when occluded compared to when unoccluded.

$$R = \frac{TSOT}{TTT_{unoccl}} \tag{1}$$

Interest in the visual occlusion method as a design and evaluation tool was sufficient to spawn a number of studies investigating the method and its predictive validity. In fact, a special issue of *Applied Ergonomics* dedicated to the technique was published in 2004 (see Gelau & Krems, 2004). In that issue, Bauman, Keinath, Krems, and Bengler (2004) applied the occlusion technique to a navigation destination entry task, and compared the results to performance on the

same task while driving. They found that whereas the results of occlusion technique did not replicate the visual demands of driving, they did provide an adequate comparison. Similarly, Noy, Lemoine, Klachan, and Burns (2004) found that the task time results using the occlusion technique were similar to those in a simulated driving condition for a range of in-vehicle tasks, including scrolling visual search, static visual search, or radio tuning. Although the task times were similar between the occlusion condition and the dual-task (driving) condition, the *R* results were different.

Pettitt, Bayer, and Stevens (2004) conducted an experiment to validate the occlusion technique as an in-vehicle device assessment tool by comparing the occlusion technique and an on-road assessment of two navigation and sound tasks. They found that the resumability ratio *R* was not related to single glance durations, but was related to the total visual demands of a device. Overall, they concluded that the trends produced by occlusion technique were similar to the results from the on-road portion of the study.

While acknowledging the need for a low-cost and accessible method, and that the prescribed version of the occlusion technique in the ISO standard (ISO, 2007) fulfills that need, it is important to recognize a key aspect of the time-sharing context that is ignored by the visual occlusion method (cf. Lansdown, Burns, & Parkes, 2004). In fact, when drivers look to the road they are processing information about the status of the roadway, any surrounding vehicles, signs, etc. The occlusion technique disregards any cognitive demand placed upon the participant during occluded periods. Admittedly, this is a conscious trade-off for the technique to remain easy and accessible since it would not be easy to incorporate a cognitive task within the occlusion goggles. However, research has shown that performing a visual-motor task during an interruption of a VCR programming task results in longer task resumption times compared to no-task interruptions (Monk, Boehm-Davis, & Trafton, 2004). This finding suggests that omitting a cognitive task during occluded periods probably results in an underestimate of the true visual demand costs for devices and tasks while driving. To test this hypothesis, an experiment was conducted to compare no-task occlusions with occlusions that include a visual-motor tracking task. It was predicted that the tracking task occlusion condition would result in larger *R*-values, as well as longer resumption times after each occlusion.

METHOD

Participants

Eight undergraduates (5 men, 3 women) from the George Mason University subject pool participated in this study for course credit. Participants ranged in age from 18 to 22 years with an average age of 20 years.

Tasks and Apparatus

An interleaved task paradigm was used in this study to emulate the visual occlusion technique. This paradigm, which has been used to study interrupted task performance (Monk, Boehm-Davis, & Trafton, 2004), required participants to program a simulated VCR while being frequently interrupted. Interruptions occurred every 1.5 seconds and were 2 seconds in duration.

It is important to note that the visual occlusion technique does not ordinarily preclude "blind operation," which allows people to continue to interact with the in-vehicle device when vision is occluded. However, "blind operation" was prohibited in this study. Thus, this paradigm was a task occlusion technique rather than a visual occlusion technique, corresponding to in-vehicle tasks using touch-screen interfaces that minimize interactions when the driver is not looking at the display.

VCR Task. The primary task was to program show information into a simulated VCR interface. The VCR interface was created using Macintosh Common Lisp and has been used in previous experimental research (Monk, Boehm-Davis, & Trafton, 2004). Participants conducted a series of mouse clicks to enter in the start time, end time, channel, and day of the week for each show into the VCR interface. This show information was presented on an index card to the right of the computer monitor and was available throughout the VCR task. To enter some of the show information, participants had to complete a series of subtasks. For example, to enter in the hour of the start time participants first had to click the column button above start hour. Next, participants had to click the radio button next to start hour. Then participants used the up or down scroll buttons to display the correct start hour. Once the correct number was displayed, participants clicked on the enter button to set the new start hour. The final step in this subtask was to deselect the current column by clicking on the column button above the "start hour" button. After every subtask was completed and the show information properly entered, participants conducted a series of clicks to set the VCR to "Record show mode" and then clicked the "Stop Trial" button to end the experimental trial.

Occlusion Task. On some trials, either a blank screen or a pursuit-tracking task interrupted the VCR task. When a blank screen interrupted participants the computer screen was occluded and participants were required to wait until the VCR task reappeared before resuming the primary task. The pursuit-tracking task required participants to follow an airplane around the computer screen using the mouse cursor until the VCR task reappeared. The airplane moved around the computer screen in a random pattern. Both the VCR task and the two secondary tasks were presented side by side on a 17-inch computer monitor. The VCR task was presented on the left-hand side and the two secondary tasks were presented on the right-hand side. The VCR task and the secondary tasks were not presented simultaneously.

Design

The experiment was a within-subjects design with two interruption conditions: the no-task condition and the tracking task condition. In the no-task condition, the VCR interface was completely occluded and participants were required to wait until the VCR interface reappeared before resuming the task. In the tracking task condition, participants were required to perform the pursuit-tracking task until the VCR interface reappeared. An equal number of unoccluded trials were also included to obtain TTT_{unoccl} values. Participants performed 10 unoccluded VCR trials, and 10 occluded trials (5 of each occlusion type). The three trial types (unoccluded, no-task, and tracking task) were performed in blocks, counterbalanced with a Latin Square.

Total task times were measured for the unoccluded VCR programming trials, as well as for the no-task and tracking task occlusion conditions. For the unoccluded condition, the total task times

corresponded with TTT_{unoccl} in the occlusion technique. The number of occlusions per trial was also recorded in order to calculate TSOT for the occluded trials. In addition, the inter-action intervals (IAI) and resumption lags (RL) were collected for the occlusion trials. An IAI is the average time between uninterrupted clicks on the VCR task and a RL is the time from the end of an interruption to the first click back on the VCR task.

Procedure

Each experimental session lasted about one hour. First, the use of the VCR interface was demonstrated to the participants. Participants then completed two separate pairs of practice trials on the VCR task and the tracking task. The VCR practice trials were uninterrupted and each practice trial with the tracking task lasted 60 seconds. After participants were familiarized with both tasks, they completed one trial of the tracking task occlusion condition followed by one trial of the no-task occlusion condition. Once training was finished, participants completed 20 experimental trials, each with new show information. Experimental conditions were counterbalanced across 4 blocks of 5 trials. The no-task and tracking task occlusion conditions were both presented in one block of 5 trials and the unoccluded condition was presented in two blocks of 5 trials. Once the experimental trials were completed, participants were debriefed and dismissed.

RESULTS

Mean TTT_{unoccl} was calculated for each participant based on the task completion time in the 10 unoccluded trials. Mean TSOTs for the two occlusion conditions were calculated for each participant by subtracting the summed occlusion times (number of occlusions multiplied by 2 seconds) from the mean TTT_{occl} from each block of the 5 occlusion trials. Finally, the R ratio was calculated according to equation 1.

The mean R calculations for the tracking task and no-task conditions (see Table 1) were entered into a repeated-measures ANOVA. The mean R for the tracking task condition was significantly larger than for the no-task condition, $F(1, 7) = 7.73$, $p < .05$, $\eta^2 = .52$. As predicted, this result indicates that the resumability of the VCR task as assessed by R was diminished by the presence of the tracking task during the occluded period.

Table 1. Means and standard deviations for *R* and *D*

Occlusion Task	*R*		*D*	
	Mean	SD	Mean	SD
No-task	1.23	0.19	0.16	0.08
Tracking	1.48	0.26	0.27	0.11

In an effort to better assess the impact of the different occlusion conditions on task resumption, the RLs were examined. With this more detailed measure of resumption performance, the proportion of the TSOT due to resumption delay was quantifiable. This resumption proportion of TSOT was calculated by subtracting the mean IAI from the mean RL to isolate the actual resumption delay. The mean resumption delay was then multiplied by the number of occlusion

periods and divided by the TSOT to produce a new estimated effect of visual demand. This new proportion was referred to as the resumption delay D (see equation 2, where n is the number of occlusion periods within a trial).

$$D = \frac{(\overline{RL} - \overline{IAI})n}{TSOT}$$ (2)

The mean D calculations for the tracking task and no-task conditions (see Table 1) were entered into a repeated-measures ANOVA. As with R, the mean D for the tracking task condition was significantly larger than for the no-task condition, $F(1, 7) = 15.31$, $p < .01$, $\eta^2 = .69$.

DISCUSSION

The present study turned the tables on the visual occlusion technique by manipulating the cognitive load during the occluded periods. It was believed that incorporating a visual-motor task into the occluded period provided a better representation of real-world driving where drivers process information when sampling the visual road scene. The results showed that R-values were significantly greater when a cognitive load was incorporated during the occlusion periods while keeping the task constant. These findings suggest that R-values underestimate visual demand in the no-task occlusion condition, which has strong implications for the veracity of the occlusion technique as specified in the ISO standard (ISO, 2007). However, it is recognized that the standard was developed with designers and evaluators in mind, and that incorporating a task into occlusion goggles would possibly impose prohibitive methodological requirements.

An important question that arises from R underestimation in the no-task condition is whether a given system's R-value would be deemed "acceptable" with the ISO standard occlusion technique application (no-task condition) but "unacceptable" if applied with a cognitive load during the occlusions as in the tracking task condition. The standard offers no guidance for a criterion R-value, stating instead that users of the technique should establish their own criteria. The present study raises the matter of accepting designs with R-values that might be above threshold if performed with a task during occlusions.

Finally, a new method for estimating visual demand costs was specified based on post-occlusion task resumption times. Despite revealing directionally similar differences between the no-task and tracking task occlusion conditions and the belief that R and D captured the same underlying cause of visual demand, the correlation between R and D was unimpressive, $r = .24$. It was difficult to specify the cause of this weak correlation with such a small data set, but the fact that there was no meaningful relationship (r < .3) suggests that R and D captured different aspects of visual demand. Ultimately, D may offer the promise of more precise estimates of resumability because it is based on resumption measures, but more data is required before that can be determined. Additionally, it is acknowledged that the measurement requirements for D make it infeasible for most users of the occlusion technique.

More than one study has attempted to validate the occlusion technique with eye-gaze measurements (Baumann, Keinath, Krems, & Bengler, 2004; Pettitt, Burnett, Bayer, & Stevens, 2004; van der Horst, 2004), but none have explored including a cognitive load during occluded

periods relative to "blank" occlusions. The present study compared no-task and tracking task occlusions and found that the tracking task condition produced higher resumability ratio R scores, indicating an underestimation of visual demand in the no-task condition. The implication of this finding is that R may shortchange visual demand estimates for in-vehicle systems because R may fail to account for additional resumption costs associated with cognitive load during occluded periods. On the other hand, D is sensitive to actual resumption times for each shift of attention, and shows promise as a better estimate of the true visual demand costs.

REFERENCES

Baumann, M. Keinath, A. Krems, J.F., and Bengler, K. (2004). Evaluation of in-vehicle HMI using occlusion techniques: Experimental results and practical implications. *Applied Ergonomics, 35*, 197-205.

Chiang, D.P., Brooks, A.M., & Weir, D.H. (2004). On the highway measures of driver glance behavior with an example automobile navigation system. *Applied Ergonomics, 35*, 215-223.

Gelau, C. & Krems, J.F. (2004). The occlusion technique: a procedure to assess the HMI of in-vehicle information and communication systems. *Applied Ergonomics, 35*, 185-187.

Green, P. (1999). Estimating compliance with the 15-second rule for driver interface usability and safety. *Proceedings of the 43rd Human Factors and Ergonomics Society Annual Meeting*. Santa Monica: HFES.

International Organization for Standardization. (2007). *Road vehicles - Ergonomic aspects of transport information and control systems - Occlusion method to assess visual demand due to the use of in-vehicle systems* (ISO Report No. 16673:2007). Geneva, Switzerland: International Organization for Standardization.

Lansdown, T.C., Burns, P.C., & Parkes, A.M. (2004). Perspectives on occlusion and requirements for validation. *Applied Ergonomics, 25*, 225-232.

Monk, C. A., Boehm-Davis, D. A., & Trafton, J. G. (2004). Recovering from interruptions: Implications for driver distraction research. *Human Factors, 46*, 650-663.

Noy, I., Lemoine, T., Klachen, C., & Burns, P. (2004). Task interruptability and duration as measures of visual distraction. *Applied Ergonomics, 35*, 207-213.

Pettitt, M.A., Burnett, G.E., Bayer, S. and Stevens, A. (2006). Assessment of the occlusion technique as a means for evaluating the distraction potential of driver support systems. *IEE Proceedings Intelligent Transport Systems, 153(4)*, 259-266.

Senders, J., Kirstofferson, A., Levison, W., Dietrich, C., & Ward, J. (1967). The attentional demand of automobile driving. *Highway Research Record, 195*, 15-33.

van der Horst, R. (2004). Occlusion as a measure for visual workload: An overview of TNO occlusion research in car driving. *Applied Ergonomics, 35*, 189-196.