Target Detection Performance for Head Mounted Indirect Vision Displays

Jeffrey T. Hansberger*[a], Victoria Blakely[b], Jared McArthur[b], James Smith[b]
[a]U.S. Army Combat Development Capabilities Command, Army Research Laboratory, Huntsville, AL 35816;
[b]University of Alabama in Huntsville, Huntsville, AL 35816

ABSTRACT

Participants used a helmet-mounted display (HMD) and a touchscreen monitor to detect targets across 3 different approaches to display 360º video for indirect vision display systems. A within-subject design was used with targets ranging from dismounted, mounted, and aerial targets. The number of targets detected, workload, and usability of each condition was measured. The HMD condition produced significantly more targets detected overall and within each type of target type compared to the monitor condition. HMD use also produced a significantly lower level of mental workload as measured by the NASA TLX and achieved a significantly higher level of usability compared to the other conditions. Possible reasons for these differences are discussed along with discussion on future studies using HMDs and mixed reality technology for indirect vision display systems.

Keywords: Indirect vision display, head mounted display, mixed reality, field of view, 360-degree video

1. INTRODUCTION

Indirect vision systems for driving and target detection are becoming an increasing area of interest for current and future military vehicles. Indirect vision displays allow Soldiers in tanks and infantry fighting vehicles to view the outside environment using camera and display technology instead of protective glass or exposing personnel outside the vehicle. The use of indirect vision displays (IVD) has several advantages such as greater protection, wider field of view (FOV) than small protective glass portals, and vision enhancement capabilities such as thermal imaging or automatic target recognition capabilities. This study will explore the potential benefits of 2 IVD approaches for target detection and compare them with a traditional protective glass viewing approach.

Past IVD research has focused on the effects of users in a driving task. Smyth[1] measured natural and direct vision for a driving task and found higher mental workload demands for the IVD drivers. Wu, Wu, and Bao[2] also examined drivers using IVD but examined the effect of field of view (FOV). They found that drivers with a narrower FOV had higher levels of workload. Related to FOV, *Ragan et al.*[3] examined the effects of FOV in virtual reality (VR) and found that greater FOV generates better performance in a visual search task. VR is a growing domain that presents an alternative means to display camera data from an IVD system as opposed to traditional monitors.

Virtual reality systems have been used to train and assess individuals in a wide range of settings and situations[4 5 6]. Rao, Pramod, Chandra, and Dutt[7] compared an IVD using a 2D monitor and a VR presentation of a search and shoot underwater simulation. They found that users in the VR condition performed better compared to the IVD condition using 2D monitors. These results involving VR suggests that there may be benefits for an IVD system by using an HMD to present the IVD information to the user. This could be done with a mixed reality solution. Mixed reality (MR) is a blending of real and virtual elements where digital and real objects and visualizations exist together[8]. Crew members of a vehicle without any physical viewports could wear an HMD with MR technology. Their view would consist of 2 parts, from the shoulders up, they would see real-time 360º video from cameras mounted on the exterior of the vehicle. From the shoulders down, they would see the physical controls of their vehicle. Head tracking capability would allow the user to move their head to look at different aspects of the outside environment in the same way they would if they were outside the vehicle looking around.

*jeffrey.t.hansberger.civ@mail.mil; phone 1 256 273-9895; h2ilab.com

Traditionally, IVD information has been presented to the user in a set of screens using one or more monitors. Past research with IVD systems have used different screen configurations that mix front, rear, and panoramic views[9 10]. Ruddle, Payne, and Dylan[11] examined a navigation task using an HMD and a desktop display but found no performance differences. After looking at several screen configurations, Metcalfe *et al*.[9] concluded that target characteristics (like distance/size) were more influential than any of the screen configurations they examined for target detection.

Past IVD research has occurred across 3 types of tasks: 1) driving, 2) navigation, and 3) target detection. This study attempts to take past research across these 3 tasks and explore the differences across presenting IVD information using an HMD and a monitor for a target detection task. Past research involving VR and IVDs have used a virtual environment with the VR. The use of 360º streaming video and not a virtual environment presents a novel way to display and assess IVD information that has not been a major area of study before. It is hypothesized that the monitor IVD will produce lower workload due to the 360º information being presented in a small, single area of the monitor. It is also hypothesized that the HMD IVD approach will produce better target detection performance and higher levels of user satisfaction. It is thought that the larger display area within the HMD, greater and more natural display of 360º information and FOV will produce higher performance and usability.

2. METHOD

2.1 Participants

Twenty-six participants (*N*= 26) volunteered from the University of Alabama in Huntsville and local organizations. Due to VR motion sickness, 1 participant did not complete the study and their data was excluded from the study resulting in a final sample of *N*= 25. The mean age of participants was 34.8 with an age range of 20-67 and the median age of 29. Female participants comprised 34% of participants while 76% of participants were male. Among the participants, 36% had no previous experience with VR, 32% had little experience, 32% had moderate experience and none had extensive experience with VR. Of the 25 participants, 8 participants (32%) served in the Military. Half of the military participants served in the Army while 25% served in the Marines, 12.5% served in the Navy and 12.5% served in the Air Force. Of the participants with Military experience, 37.5% served as a Gunner of a Military vehicle and 37.5% served as a Commander of a Military vehicle.

2.2 Equipment

A desktop computer was used with an Oculus Rift HMD for the HMD and baseline conditions while a 27" touchscreen monitor was used for the monitor condition. The monitor was positioned at an almost flat horizontal position to avoid hand/arm fatigue by the participants. The target detection simulation was created using the Unity game engine and the NASA TLX[12] and System Usability Scale (SUS)[13] was administered on the desktop computer.

2.3 Procedure

The experimental design was a within-subject design with 3 conditions: 1) baseline, 2) monitor IVD, and 3) HMD IVD. A Latin square design was used to counter-balance the order each participant would engage with the 3 conditions to mitigate any learning effects. The baseline condition represented the view from inside a combat vehicle with 3 short but wide viewports to the exterior environment (Figure 1A). The interior of the combat vehicle was simulated using the Unity game engine and provided the participant the partially obscured view that exists in many current heavily armored vehicles.

The monitor IVD condition represents a 360º view from cameras mounted on the exterior of the vehicle and presented it on a single monitor. The view was segmented into a 3x3 grid that displayed the front, side, and rear environment (Figure 1B) as past research has shown that segmented views for IVD can be beneficial[10]. The dimensions of the 3x3 grid was 14" wide and 9" tall representing a 15" touchscreen monitor. The target and environment graphics were the same as the baseline and HMD IVD conditions.

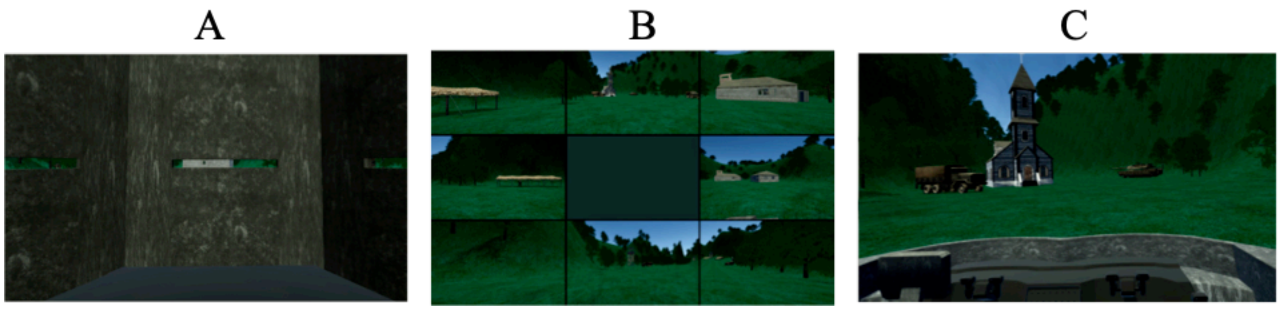

Figure 1. Three different viewing conditions. Figure 1A shows the baseline condition with the 3 small viewports. Figure 1B shows the 3x3 grid view for the screen IVD condition and Figure 1C shows a portion of the HMD view. The user can move their head around to change the viewing angle to see different aspects of the environment.

The HMD IVD condition was similar to the baseline condition using the HMD that simulated the interior of armored vehicle. Instead of the 3 small viewports, they had 360º visual access to the outside environment from the shoulders up that also included full view of the sky (Figure 1C). This simulated a mixed reality view where they could see 360º video from the exterior camera system while still seeing their interior controls. Head tracking allowed them to view the 360º video scene by looking around them in a very natural manner. The increased viewing area that HMDs provide allowed for the 360º view to be larger than the monitor IVD condition and for targets to retain their original size.

Participants were told they would interact with 3 scenarios and take a series of questionnaires after each. The task was to identify as many targets during the scenario as possible. There were 3 types of targets that included 1) soldiers, 2) vehicles, and 3) aircraft. There were 3 different scenario terrains that were equivalent in timing and targets but had different path through the terrain to avoid a learning effect. The scenario was a rural scene with hills. The duration of each trial was 3 minutes. The participant viewed the environment from inside a virtual vehicle that moved through the terrain on a set path and speed, which the participant had no control over. For the baseline and HMD IVD conditions, the participant could control a targeting reticle by moving their head. They were given an Oculus wireless remote and asked to click the remote button to tag the target when the reticle was on the target. To target an object for the monitor IVD, participants would tap the desired target on the touchscreen. Participants had a 1 minute practice session after verbal instructions on targeting before each trial. After each trial, participants were asked to complete the NASA TLX and SUS questionnaires. The experiment concluded with a short interview session. The experiment duration was approximately 1 hour.

3. RESULTS

The results of this study will be described beginning with overall performance results followed by the performance results based on each type of target. Overall workload results will be discussed followed by the individual workload subscales and then user satisfaction and subjective feedback will be discussed. Limited mean substitution was used for missing data.

3.1 Overall Target Detection Performance

The overall number of targets in a trial was 54 targets, which consisted of 36 soldiers, 13 vehicles, and 5 air targets in the form of helicopters. Performance was measured as the number of correctly identified targets within the 3-minute trial. When participants selected a target, the target would blink twice to indicate it had been successfully tagged. Each target could only be tagged once. A repeated measures ANOVA was used to analyze the overall performance data. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 0.41$, $p = .82$. A significant difference was present for overall targets identified, $F(2, 48) = 31.28$, $p < .01$. Figure 2 displays the means for each condition. Posthoc tests using Bonferroni correction showed that the HMD IVD ($M = 28.49$) condition produced a significantly higher level of performance compared to the baseline ($M = 22.39$, $p < .01$) and monitor IVD ($M = 21.11$, $p < .01$) condition.

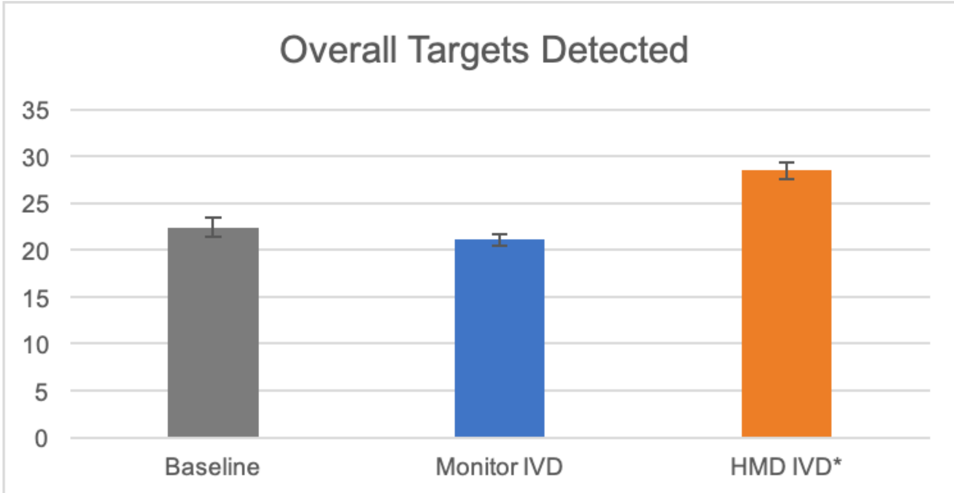

Figure 2. Overall targets detected with SE for each condition. HMD IVD condition was significantly higher than the baseline and monitor IVD conditions (*$p < .05$).

3.2 Soldier Target Detection

There was a total of 36 soldiers in each scenario who were dressed in typical military clothing and carrying weapons. There were no civilian or non-target individuals present in the scenarios. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 0.81$, $p = .67$. The repeated measures ANOVA determined that there was a significant difference among the conditions regarding the number of soldier targets detected, $F(2, 48) = 5.91$, $p < .01$. The posthoc tests using Bonferroni correction revealed that a significantly greater number of targets were detected in the HMD IVD condition (M=15.23) compared to the baseline condition ($M = 12.58$, $p < .05$) and the monitor IVD condition ($M = 12.44$, $p < .05$).

The soldier targets were divided into 3 groups that related to the level of difficulty to identify them. Each soldier difficulty group consisted of 12 soldiers and were randomly placed throughout the scenarios. The "easy" group were in plain view without any concealment. The "medium" difficulty group were partially concealed by other objects or structures in the environment (e.g., partly behind a tree). The "hard" difficulty group was mostly concealed within the environment (e.g., placed inside an outdoor structure with only their head showing or among dense woods). Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 0.30$, $p = .86$. Figure 3 shows the same trend seen for the overall soldier targets with the highest number of soldiers targeted in the HMD IVD condition ($M = 8.99$) compared to the baseline ($M = 7.92$) and monitor IVD conditions ($M = 7.88$). However, this difference was not significant $F(2, 48) = 1.62$, $p = .21$ for the easy soldier targets.

The same analyses were done for the "medium" soldier targets. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 3.86$, $p = .145$. The repeated measures ANOVA revealed a significant difference among the conditions for "medium" soldier targets, $F(2, 48) = 4.00$, $p < .05$. Posthoc tests using Bonferroni correction showed that the monitor IVD ($M = 3.88$, $p = .05$) and HMD IVD ($M = 4.16$, $p < .05$) conditions had significantly more medium soldier targets detected than the baseline ($M = 3.00$).

The final difficulty category for soldier targets also did not violate the sphericity assumption, $\chi^2(2) = 1.96$, $p = .38$. The repeated measures ANOVA detected the strongest significant difference among all soldier target comparisons ($F(2, 48) = 10.67$, $p < .01$). The Posthoc tests revealed that the monitor IVD ($M = 0.68$) produced significantly fewer targets detected than both the baseline ($M = 1.67$, $p = .01$) and the HMD IVD ($M = 2.08$, $p < .01$) conditions. The results related to detecting the soldier targets shows a performance advantage for the HMD IVD condition as it enabled the greatest number of targets detected overall and in each of the 3 difficulty levels.

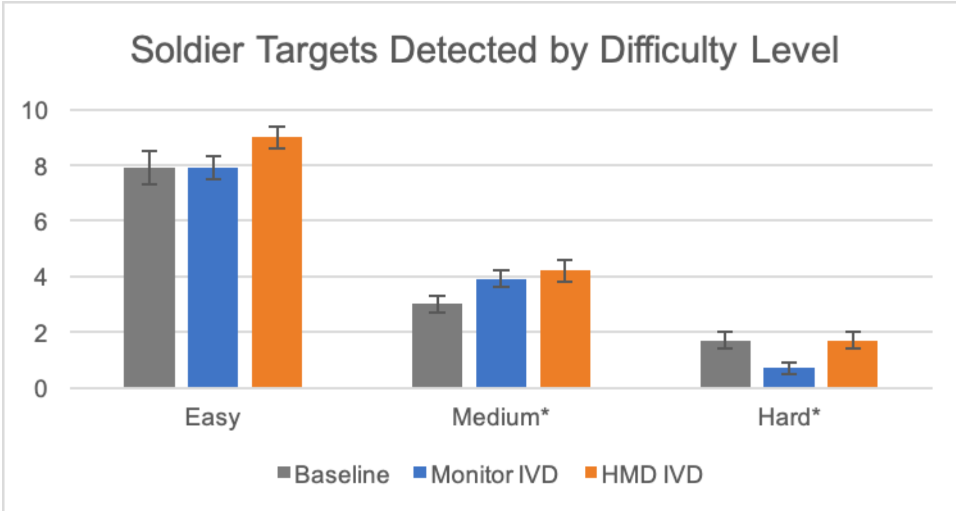

Figure 3. Target detection among soldier difficulty levels, easy, medium, and hard (*$p < .05$).

3.3 Vehicle Target Detection

There was a total of 13 vehicles in each of the scenarios. These vehicles were a mix of military Humvees and tanks that were randomly placed throughout the simulated environment. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 0.132$, $p = .94$. The repeated measures ANOVA determined that there was a significant difference among the conditions regarding vehicle detection, $F(2, 48) = 3.82$, $p < .05$ (Figure 4). Posthoc analyses using Bonferroni correction showed that the monitor IVD condition ($M = 7.63$) had significantly lower vehicle detections compared to both the baseline ($M = 9.04$, $p < .05$) and the HMD IVD ($M = 9.40$, $p < .05$).

3.4 Air Target Detection

There were a total of 5 air targets in the form of helicopters flying overhead for each scenario. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 0.64$, $p = .94$. The repeated measures ANOVA determined that there was a significant difference among the conditions regarding vehicle detection, $F(2, 48) = 75.13$, $p < .01$ (Figure 4). Posthoc analyses using Bonferroni correction showed that the HMD IVD condition ($M = 3.86$) had significantly higher air target detections compared to both the baseline ($M = 0.76$, $p < .01$) and the monitor IVD ($M = 1.05$, $p < .01$).

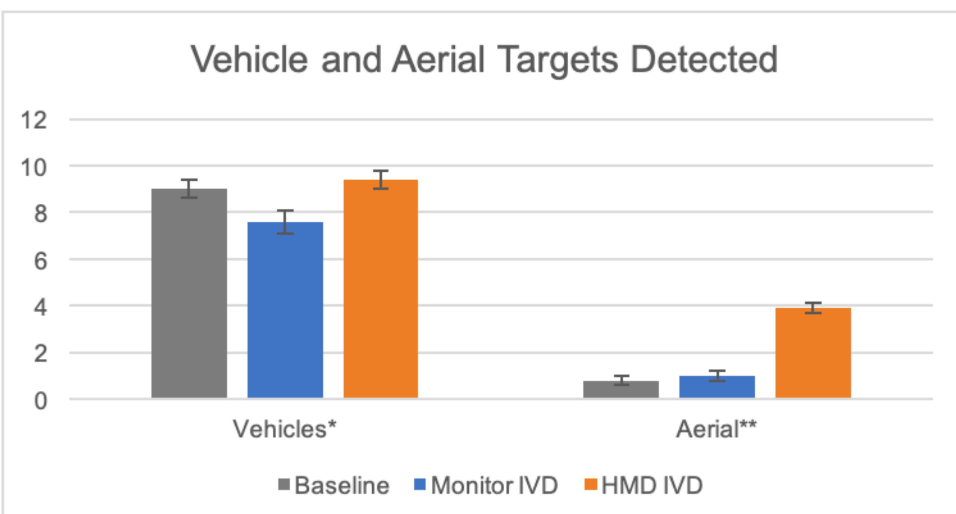

Figure 4. Target detection performance among vehicle and aerial targets (*$p < .05$).

3.5 Workload

The raw NASA TLX scores on the 21-point scale were used after each scenario in order to assess the workload associated with that condition. The overall workload will be described followed by the relevant workload subscales. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 1.84$, $p = .40$. The repeated measures ANOVA determined there was not a significant difference overall for the NASA TLX ($F(2, 48) = 2.09$, $p = .14$) but that a similar trend found among the target detection analyses was evident for overall workload. The baseline ($M = 54.2$) and monitor IVD ($M = 56.6$) conditions resulted in higher reported levels of workload than the HMD IVD ($M = 49.1$) condition (Figure 5).

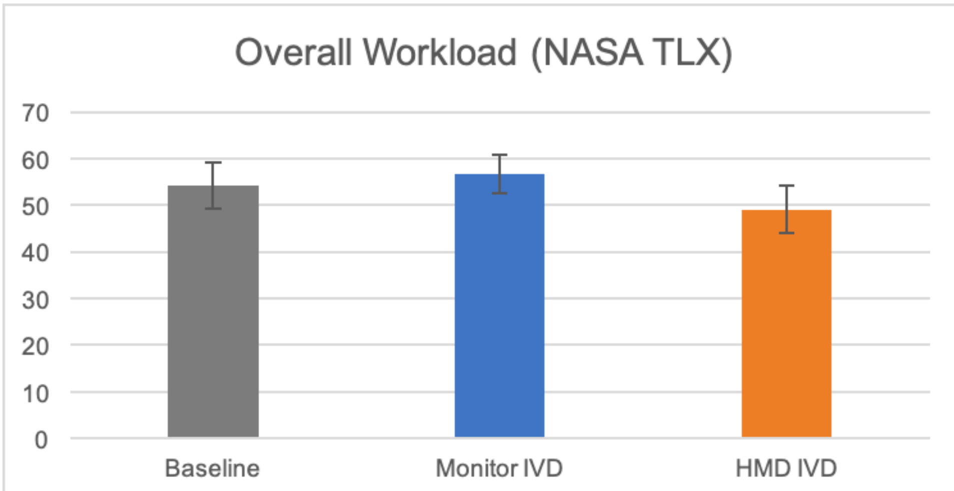

Figure 5. Overall mean workload results with SE across all conditions.

Two of the NASA TLX subscales did show a significant difference among the conditions. The Mauchly's Test of Sphericity indicated that the assumption of sphericity was violated for the mental ($\chi^2(2) = 8.26$, $p < .05$) and the physical ($\chi^2(2) = 8.27$, $p < .05$) workload subscales. The Greenhouse-Geisser correction was used due to the sphericity violation. The repeated measures ANOVA determined that there was a significant difference for both the mental ($F(2, 48) = 3.99$, $p < .05$) and physical ($F(2, 48) = 4.52$, $p < .05$) workload subscales. The mental workload was lower for both the baseline ($M = 10.94$, $p = .28$) and HMD IVD ($M = 10.24$, $p = .07$) conditions compared to the monitor IVD ($M = 12.68$) condition but neither reached statistical significance. The HMD IVD did approach statistical significance compared to the monitor IVD ($p = .07$) and the lack of significance for either pairwise comparison is likely due to insufficient sample size. The physical workload index however showed a different pattern. The posthoc analyses using Bonferroni correction indicated that the monitor IVD ($M = 3.04$) had significantly lower physical workload than the baseline condition ($M = 5.40$, $p < .01$) but no difference compared to the HMD IVD condition ($M = 4.00$, $p = .25$). Figure 6 displays the results across all the NASA TLX subscales. The consistent pattern among the subscales is that the HMD IVD produced the lowest levels of workload while the monitor IVD had the highest with the exception of the physical workload dimension.

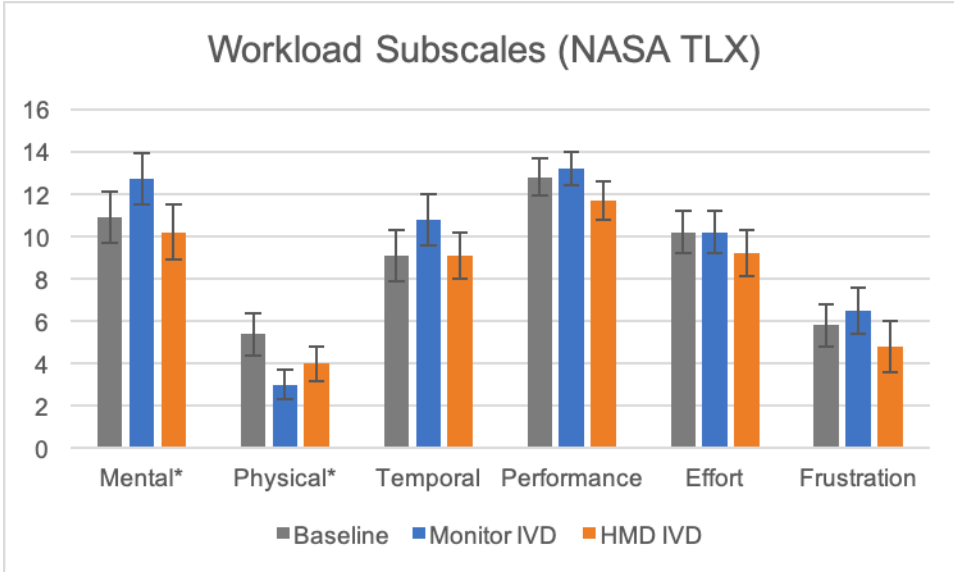

Figure 6. Mean raw workload NASA TLX overall and subscale results (*$p < .05$).

3.6 User Satisfaction

Levels of user satisfaction and usability were measured with the SUS questionnaire, which was administered after each scenario. Mauchly's Test of Sphericity indicated that the assumption of sphericity had not been violated, $\chi^2(2) = 3.63$, $p = .16$. The repeated measures ANOVA determined that there was a significant difference among the conditions regarding vehicle detection, $F(2, 48) = 16.46$, $p < .01$. Posthoc analyses using Bonferroni correction showed that the HMD IVD condition ($M = 83.3$) had significantly higher levels of usability compared to both the baseline ($M = 66.56$, $p < .01$) and the monitor IVD ($M = 63.0$, $p < .01$) (Figure 7). When these scores are compared to industry wide SUS scores, the HMD IVD obtains an "Excellent" rating because it is above an 80.3 score compared to other industry systems[14]. The baseline and monitor IVD conditions both fall into the "OK" usability rating category with a score between 51 and 67[14].

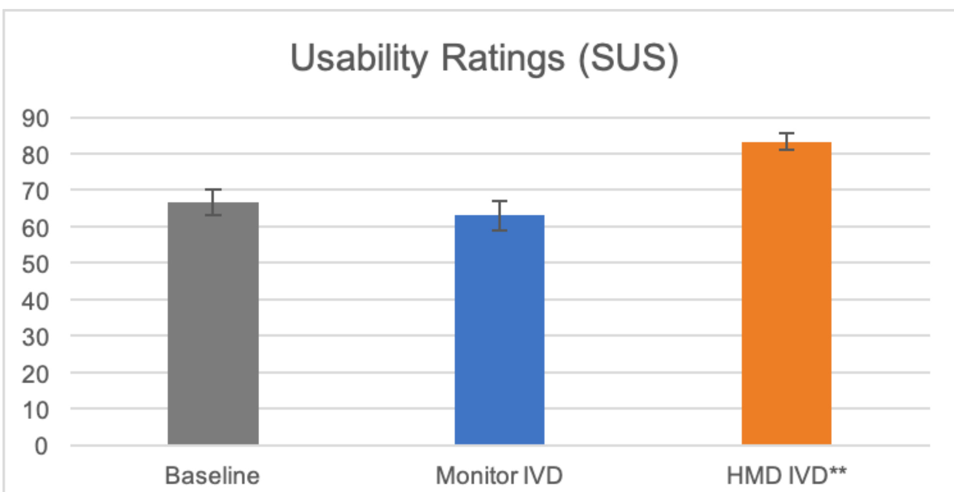

Figure 7. Usability scores with SE across all 3 conditions (**$p < .01$).

4. DISCUSSION

The potential benefits of IVDs offers increased protection for military vehicle crews but past research has shown that there can be decrements in driving performance and workload[1]. The traditional approach to display indirect vision information has been with 2D monitors and segmenting the 360° video. This study has leveraged some of the past work suggesting that increased FOV capabilities can enhance detection performance and reduce workload[2 3]. The proposed mixed reality concept in this paper can display 360º streaming video and still provide visual access to the physical controls through the use of an HMD. The results showed considerable promise for the 360º HMD IVD concept ranging from target detection and usability enhancements to reduced levels of workload.

Situation awareness is vitally important to military operations. The prospect of IVD is improvement of situation awareness with augmented views (e.g., thermal imaging) and capabilities such as automatic target recognition while improving the safety and survivability of the crew. The ability to detect targets in the environment is one important component of situation awareness for the crew of a military vehicle. The results of this study showed a consistent pattern for the HMD IVD condition for target detection. HMD users were able to detect more targets overall and were able to detect more targets of each type to include soldiers, vehicles, and helicopters.

Among soldier targets, the HMD IVD condition detected more overall targets than the baseline and monitor IVD conditions. When examined by the level of difficulty, the HMD IVD capabilities performed the best for "hard" targets that had the most concealment. This is arguably the most important and realistic type of target as it reflects typical enemy targets that are intentionally trying to avoid detection. Detection of easy and medium targets with the HMD were not reduced at all given that the HMD users also detected the most targets for these as well, even if some of those differences were not statistically significant. There are a few possible reasons why the HMD IVD showed higher target detection. The first possible reason is simply that the targets were larger in the HMD due to the greater amount of display space compared to the monitor IVD. This is consistent with the findings from Metcalf et. al.[9] where they found that detection performance declined as range increased. The baseline condition, which shared the same target size as the HMD condition also produced significantly more hard target detections compared to the screen IVD condition and provides additional support for the size of the target hypothesis.

The second possible explanation of the higher detection performance could involve the field of view related to the HMD. The potential benefit regarding the FOV may not have to do with the amount of FOV but how it is structured. The HMD IVD is similar to our everyday FOV where we have our central front view that represents the user's primary interest that is complimented by periphery vision. The periphery vision can trigger a switch of attention by movement or other visual cues all while maintaining situation and spatial awareness. Ruddle et al.[11] found that the use of an HMD did promote more visual scanning of the environment from "looking around" compared to viewing the same environment on a desktop monitor. The presentation of 360º video by the monitor IVD is segmented and arranged in a 3x3 grid to represent each of the video feed locations to their grid location. There are not many instances of video being presented in this manner, let alone moving video from a vehicle. Several participants mentioned that the grid view in the monitor IVD condition was slightly disorienting. It is possible that additional training with this particular view is needed to reach its full capability.

The results for air target detection were similar to the soldier data with the HMD condition producing significantly more detection compared to the other 2 conditions. In this case, the performance difference is almost certainly due to the greater overhead FOV and is consistent with past research that shows correlation between FOV and performance[2 3]. Participants commented positively about the freedom and ease the HMD IVD condition provided in scanning and detecting the aerial targets. The results for the vehicle targets were the same as the hard soldier targets with the HMD and baseline conditions performing significantly higher than the monitor IVD condition. It is surprising that the baseline condition outperformed the screen IVD condition given its highly-occluded view. Given that the size of the vehicle targets is much larger than the soldier targets, the size hypothesis mentioned earlier does not seem applicable, which leaves how the FOV is structured as a possible explanation. Confusion or difficulty interpreting the grid format for 360º video for a moving vehicle could also negatively impact the user's workload.

Our hypothesis related to workload was that the monitor IVD condition would produce lower levels of workload due to all the 360º information being located in such a small area and how efficient and familiar the touchscreen is for selecting targets. The monitor IVD condition contradicted this hypothesis and scored the highest workload overall and across every subscale except the physical subscale. Given that only one of the subscale differences was significant (mental), these results are not conclusive. The high workload trend for the monitor IVD and lower target detection performance however suggests that users had some difficulty with the monitor IVD condition. The temporal subscale result is of particular interest as several participants asked why their movement in the monitor IVD condition was faster than the other conditions. The speed of the participant's vehicle across all scenarios was verified to be the same during the experiment. If the monitor IVD format gave the illusion of moving faster, that could be an explanation for the higher temporal workload, lower performance and the higher mental workload results. Another possible contributing factor to the higher mental workload is the 3 different forms of motion being presented to the viewer at the same time. In the monitor IVD, the top row shows the environment moving toward the viewer. The 2 side images show the environment moving past the viewer and the bottom row shows objects moving away. There could be additional cognitive demands required to monitor and interpret moving video across these 3 different perspectives. This is an area for future research efforts to better understand the perceptual and cognitive implications of a grid presentation for a 2D based IVD.

The HMD IVD condition showed a general trend for low amounts of workload across every subscale except for physical workload. Several participants requested that the reticle be controlled by eye gaze instead of head movement, which would likely reduce the physical workload associated with the HMD. It is possible that the degree of similarity presenting 360º video with an HMD to the way we typically scan and view our surrounding contributed to the lower levels of workload, especially for the mental workload index. The naturalness of this approach also was evident with the SUS scores where the HMD condition scored the highest by a large margin. Its score of 83.3 places it in the highest SUS rated category ("Excellent") based on industry-wide SUS use and scoring. The HMD workload and SUS scores are surprising given that nearly half (48%) of the participants in this study reported that they had little to no prior experience with HMDs and VR. Only 1 minute of training was provided before the 2 conditions using the HMD. This level of unfamiliarity with HMDs did not appear to have a negative effect on users. As part of the interview process at the end of the study, participants were asked which method they preferred. Eighty-four percent of the participants preferred the HMD IVD condition while 8% preferred the baseline and 8% preferred the monitor IVD condition.

This study provided initial results with the HMD IVD concept. Future efforts could expand on this research by using a more expert military population beyond the 32% that participated in this study. Expanding the type of task used beyond the high-density target detection task used in this study is also needed to include lower workload type of scenarios and different settings like an urban environment. Use of the HMD IVD concept in an actual moving vehicle is needed in order to explore the possibility of motion sickness among users while using this type of technology. A logical next step to this research is to conduct a similar experiment with real-time streaming video using mixed reality hardware instead of simulating it in virtual reality. More exploration in the use of the monitor IVD configuration is needed to validate these results and explore the higher workload finding, particularly with the higher mental workload effect, which could help explain the lower performance results within this condition.

The results indicate that using an HMD for 360º IVD has promise in terms of target detection performance and acceptable levels of workload. Participants with little training and limited prior experience with HMD technology evaluated the HMD 360º IVD concept high in usability and overwhelmingly preferred it over the other conditions. This suggests that future research efforts should continue to pursue this concept and ways to further augment performance through the use of automatic target detection and imaging enhancement approaches.

REFERENCES

[1] Smyth, C. C. (2001). Modeling Mental Workload and Task Performance for Indirect Vision Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *45*(23), 1694–1698.

[2] J. Wu, Z. Wu and J. Bao, "Study on the impact of indirect driving system on mental workload and task performance of driver," *Proceedings of 2013 IEEE International Conference on Vehicular Electronics and Safety*, Dongguan, 2013, pp. 53-56.

[3] E. D. Ragan, D. A. Bowman, R. Kopper, C. Stinson, S. Scerbo and R. P. McMahan, "Effects of Field of View and Visual Complexity on Virtual Reality Training Effectiveness for a Visual Scanning Task," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 7, pp. 794-807, 1 July 2015.

[4] ter Haar, R. (2005). Virtual reality in the military: Present and future. In *3rd Twente Student Conf. IT*.

[5] Stedmon, A.W., & Stone, R.J. (2001). Re-viewing reality: human factors of synthetic training environments. *Int. J. Hum. Comput. Stud., 55*, 675-698.

[6] Ka-Chun Siu, PhD, Bradley J. Best, PhD, Jong Wook Kim, PhD, Dmitry Oleynikov, MD, Frank E. Ritter, PhD, Adaptive Virtual Reality Training to Optimize Military Medical Skills Acquisition and Retention, *Military Medicine*, Volume 181, Issue suppl_5, 1 May 2016, Pages 214–220

[7] Rao, A. K., Pramod, B. S., Chandra, S., & Dutt, V. (2018, July). Influence of Indirect Vision and Virtual Reality Training Under Varying Manned/Unmanned Interfaces in a Complex Search- and-Shoot Simulation. In *International Conference on Applied Human Factors and Ergonomics* (pp. 225-235). Springer, Cham.

[8] C. E. Hughes, C. B. Stapleton, D. E. Hughes, and E. M. Smith. Mixed reality in education, entertainment, and training. IEEE Computer Graphics and Applications, pp. 24–30, Nov./Dec. 2005.

[9] Metcalfe, J.S., Cosenzo, K., Johnson, T., Brumm, B., Manteuffel, C., Evans, A.W., & Tierney, T.M. (2010). Human Dimension Challenges to the Maintenance of Local Area Awareness Using a 360 Degree Indirect Vision System.

[10] Boonsuk, W. Evaluation of Desktop Interface Displays for 360-Degree Video. Master's Thesis, Iowa State University, Ames, IA, USA, 2011.

[11] R. A. Ruddle, S. J. Payne, and D. M. Jones, "Navigating Large- Scale Virtual Environments: What Differences Occur Between Helmet-Mounted and Desk-Top Displays?" Presence: Teleoperators and Virtual Environments, vol. 8, no. 2, pp. 157– 168, 1999.

[12] Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908.

[13] Brooke, J. (1986). "SUS: a "quick and dirty" usability scale". In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (eds.). *Usability Evaluation in Industry*. London: Taylor and Francis.

[14] *Bangor, A., Miller, J. & Kortum, P. (2009). Determining what individual SUS scores mean: Adding an adjective rating scale. Journal of Usability Studies, 4(3), 114–123.*