Comparison of Collaborative Display Technologies for Team Design Review

Ming Hou [*], Andrea Barone [#], Lochlan Magee [*], and Mike Greenley [#]
* Defence Research & Development Canada # Greenley & Associates Inc.

A comparative evaluation of collaborative display technologies was conducted to explore their ability to support a pair of participants conducting a collaborative workspace design review. Five review media were compared: 2D CAD model on CRT, 3D CAD model on CRT, 3D CAD model on a Curved plasma display, a large DataWall display, and a CAVE environment. Participants reviewed a model depicting an in-vehicle navigation system installed within the front dash of a vehicle and detected design flaws. Performance measures (number of detected flaws and detection time) and usability measures (display, design review, and collaborative quality) were collected. The main findings were: a) flaw detection was better for 3D displays than the 2D display; b) flaws detection was progressively reduced with more immersive 3D displays; c) speed-accuracy tradeoffs were observed such that detection time was less for the 2D than the 3D displays, and decreased with the degree of immersion; d) using the standard CRT is more cost-effective than using the Curved, DataWall, or CAVE displays.

INTRODUCTION

Collaborative work is becoming more important (Beevis, Vallerand, and Greenley, 2001), and a collaborative display provides a means to enhance collaborative design work, especially with the development of immersive and virtual reality technologies. Collaborative display media vary along a continuum of fidelity: a Computer-Aided Design (CAD) model can be presented in the form of a two-dimensional (2D) blueprint, a three-dimensional (3D) image, or an immersive virtual environment. It is difficult to provide users with a sense of immersion given a 2D representation on a flat Cathode Ray Tube (CRT) screen. On the other hand, a fully immersive technology can provide a variety of features that provide a greater sense of presence, such as: 3D viewing, real-time dynamics, ego-centered frame of reference, multimodal interaction, and a wide field of view (Wickens & Hollands, 2000). An example of fully immersive technology is the computer animated virtual environment (CAVE), in which users freely move and interact with stereo images projected on the four (to six) walls of a room with unlimited field of view (Owen, 1999; Dynott, 2002). Partially immersive technologies include 3D representations on flat CRTs, large flat DataWall projectors with wide field of view, and some curved display systems with a more compelling depiction of space and depth (e.g., a curved hemispheric plasma projector).

Although immersive technologies show great promise for collaborative design work, selecting a collaborative display is often done without considering its benefits, or whether other display options might be better suited to the situation at hand. In addition, the effect of collaborative display on team performance is not clearly understood, despite a fairly significant research investment over the past few years (Kline & McGrath, 1999). Furthermore, although research often does not examine the influences of technology on team performance (e.g., Gwynne et al., 1996; Hoeksema, 1998; Bolstad & Endsley, 1999), the sustained interest in the use of collaborative displays is driven by the belief that group decision-making can outperform individual decision making (Beevis, Vallerand, and Greenley, 2001). In other words, collaborative displays are likely to be acquired and implemented regardless of whether or not their use has been validated empirically. The question, therefore, from a human factors engineering perspective, is not whether collaborative displays should be adopted, but rather how to use the technology most effectively.

To address this question and provide input to the design of efficient and effective collaborative displays, we conducted an experiment that explored the utility of different types of collaborative display technologies for supporting a pair of participants performing a collaborative design review. We predicted that performance and usability ratings should improve with the degree of immersion. This has two implications. First, 3D displays should be superior to 2D displays because 3D environments can provide greater presence with real-time dynamics and an ego-centred frame of reference. Second, performance should improve as the level of immersion increases (because of the wide field of view and multimodal interaction). This study involved a comparison of five different types of collaborative displays: 2D CAD model on a CRT, 3D CAD model on a CRT, 3D CAD model on a Curved plasma display, 3D CAD model on a DataWall projector, and 3D CAD model on a CAVE. Many factors covary with changes in display type, such as display size, resolution, input device, and interaction mode. It was not technically possible to keep these factors equivalent. The results should nevertheless indicate the relative performance advantages of commercially available display systems for collaborative design review.

METHOD

The design space was a CAD model of a navigation system located within the front dash of a vehicle. Eight review teams reviewed the design space in each of the five display conditions. Each team consisted of two participants (one driver and one passenger). Each team used only one of the five collaborative displays in this experiment, giving the experiment a one-way, between-subjects design. The review task was to detect design flaws or installation problems (embedded in the system) that contradicted published navigation system installation recommendations (Green, et al., 1994; Commission of the European Communities, 2000; Japan Automobiles Manufacturing Association, 2000). The design flaws were detectable in both 2D and 3D models and are listed in Table 1. Three separate views of the model were used for exhibiting the navigation system: dash only (without DVD and screen showing), DVD tray deployed (screen not showing), and screen in upward position (DVD not showing). These views were optionally available for participants for the design review in each experimental condition, and some examples are illustrated in Figures 1 and 2.

Subjects 80 participants (42 females and 38 males) ranging in age from 18 to 58 years were selected using an eligibility checklist. Each had at least two years driving experience and normal or corrected-to-normal vision. None had experience with vehicle manufacturing or in-vehicle navigation systems. Each participant was randomly assigned to a two-person team and each team was randomly assigned to one of the five display conditions (16 participants per condition).

Apparatus The CAD models were developed in 2D using Pro/ENGINEER™ 2000 i^2 (by PTC), and converted into PDF files for viewing on a Samsung SyncMaster flat screen MagicBright™ CRT monitor. The 2D models included four separate elevations of the vehicle: from the perspective of the right and left side of the vehicle, from the perspective of the back of the vehicle looking forward, and from the perspective of the top of the vehicle looking downward. They were static and monochrome when viewed on the CRT. Figure 1 shows examples of the2D CAD models. These 2D models were translated into 3D CAD models using eDrawings 2003 (by SolidWorks). The 3D models were shown in dark grey on a blue background. Figure 2 shows examples of the 3D models. The same CRT monitor was used for viewing 2D and 3D CAD models.

Table 1. Design Flaws Embedded in In-Vehicle Navigational System

1	System obstructed other vehicle controls and displays
2	Visual display too low (not positioned close to driver's normal line of sight)
3	Controls not in easy reach
4	Frequently used controls not in dominant hand position
5	Safety: location of navigation system distracted driver's attention
6	Poor angle of the screen

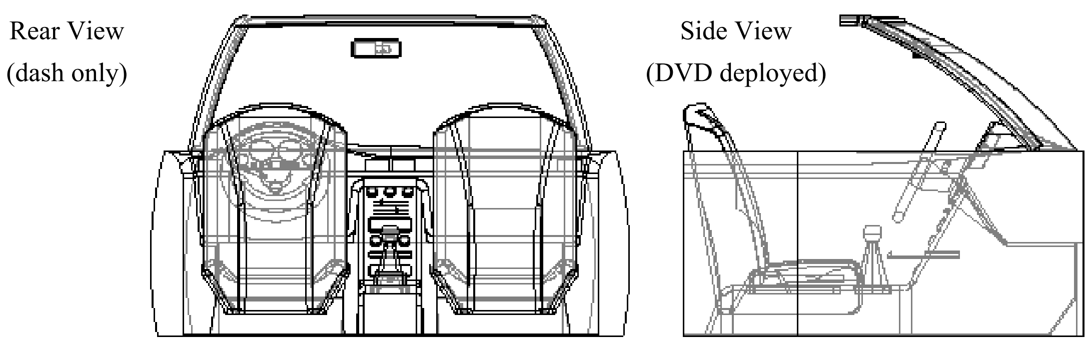
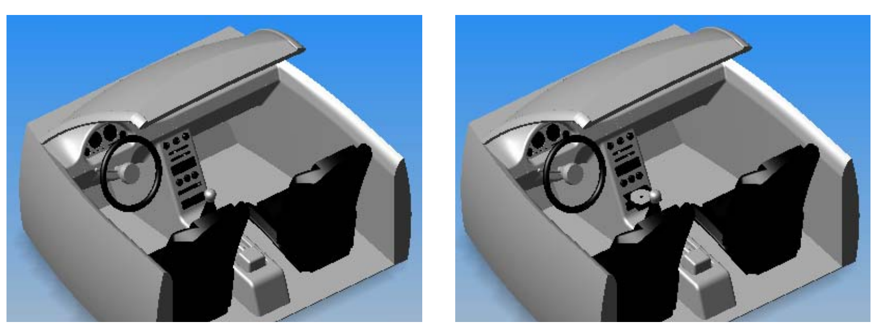


Figure 1. An example of the 2D CAD model on a CRT



Navigation system flush with dash DVD deployed

Figure 2. An example of the 3D CAD model on a CRT

The 3D CAD models were scaled to fit a Curved (hemispheric) Elumens VisionStation® plasma display, a Marquis 8500 plasma DataWall projector, and a Fakespace™ CAVE to produce three other experimental conditions. The 3D CAD models for these three conditions were generated on an SGI Onyx 3200 using Perfly for Performer. In the DataWall and CAVE environments, the models were stereoscopic and viewed by the participants wearing Stereogaphic® liquid crystal stereographic glasses. Characteristics of the collaborative display environments are outlined in Table 2.

A standard computer mouse was used to manipulate model orientation for the 3D CRT monitor, the Curved display, and the DataWall projector. In the 2D condition and all 3D conditions except for the CAVE, participants were able to switch between the three views of the model (dash only, DVD deployed, screen deployed) by clicking the right mouse button. In the 3D conditions (except CAVE), the models could also be zoomed and rotated in x, y, and z planes with pitch, roll, and yaw manipulations using the mouse. In the CAVE environment, a Fakespace™ Dataglove was used for manipulation of the 3D models (including pitch, roll, and yaw). The metal contacts on the glove's fingertips were used to switch views and manipulate the models. The 2D CAD models did not allow rotating or zooming. The control of the models was given one of the participants only – the driver in each team for each experimental condition. The passenger was responsible for data recording as a means for collaboration.

Procedure Each review team was briefed about the use of the in-vehicle navigation system and how to conduct the experiment after the screening process. Each team received 5 to 15 minutes of familiarization time with the display media and the input device before they started a trial. They then had a maximum of 20 minutes to detect design flaws by reviewing the models. They were encouraged to discuss possible flaws with each other during the experiment. The trial was completed and data collected when the participants indicated that they could no longer identify any design flaw. Informal observations of participants' behaviour were also noted. After the trial, participants were debriefed and completed a usability questionnaire. The experiment, including the practice, took approximately 30 to 45 minutes for each team.

Measures During the experiment, two performance measures were taken: *number of flaws detected* (number of design flaws identified) and *detection time* (total time spent on task). In addition, three subjective measures were obtained from the usability questionnaire. *Display usability* looked at the how easy it was to understand the CAD model, identify vehicle features, and imagine interacting with the navigation system. *Design review usability* examined the ease of performing the design review, and how comfortable and confident the participants were while conducting the design review. *Collaborative quality usability* assessed the level of immersion and co-presence the participants experienced within the vehicle environment, and the degree to which the review medium facilitated collaboration and communication between two participants in a review team. Ratings were averaged across items for each measure (6, 5, and 5 items, respectively).

RESULTS

The above measures were analyzed using an one-way between-subjects analysis of variance (ANOVA). Simple comparisons were conducted when appropriate. An alpha level of .05 was used for all statistical tests.

Number of Flaws Detected Figure 3 shows the number of flaws detected in each condition. Display type affected the number of flaws detected, $F(4, 35) = 3.68$, $p < .05$. Simple comparisons revealed that participants detected fewer flaws with the 2D than the 3D displays, $F(1, 35) = 7.33$, $p < .05$, and that more flaws were identified with the 3D CRT display than with the other 3D displays (Curved, DataWall, and CAVE), $F(1,35) = 4.68$, $p < .05$.

Table 2. Characteristics of Collaborative Display Media

	Collaborative Review Media				
Display Character	2D CAD on CRT	3D CAD on CRT	Curved Plasma Display	DataWall Projector	CAVE
Level of Immersion	Non-Immersive	Partially Immersive			Fully Immersive
Display Size	17''	17''	62.5" (W) x 57"(H) x 21" (D)	96" x 72"	4 x 120" x 120"
Aspect Ratio	4:3	4:3	~ 5:4	4:3	1:1
Curved	No	No	Radius: 33"	No	No
Input Device	Optical Mouse	Optical Mouse	SGI Mouse	SGI Mouse	Fakespace Dataglove
3D Glasses	No	No	No	Yes	Yes
Model Manipulation	No	Yes	Yes	Yes	Yes
Display Cost (in US $)	< $1,000	< $1,000	~ $29,000	~ $200,000	$350,000 - $450,000

Detection Time Figure 4 shows detection time in each condition. Display type influenced detection time, $F(4, 35) = 3.98$, $p < .05$. Simple comparisons indicated that participants spent more time with the 3D CRT than with the other 3D displays (Curved, DataWall and CAVE), $F(1, 35) = 5.14$, $p < .05$.

Usability Ratings Display usability ratings were affected by display type, $F(4, 35) = 10.08$, $p < .05$. Simple comparisons indicated that participants produced higher display usability ratings with the 3D displays than with the 2D CRT, $F(1, 35)= 4.21$, $p < .05$. Design review usability ratings were higher for the 3D displays than for the 2D display, $F(4, 35) = 14.08$, $p < .05$. Simple comparisons revealed that participants produced higher design review ratings with the 3D displays than with the 2D CRT, $F(1, 35) =3.85$, $p < .05$. They produced higher ratings for the 3D CRT than for the other 3D displays (Curved, DataWall and CAVE), $F(1, 35) = 3.36$, $p < .05$. Collaborative quality usability ratings were higher for the 3D conditions than for the 2D CRT, $F(4, 35) = 11.74$, $p < .05$. Simple comparisons indicated that participants produced higher ratings with the 3D displays than with the 2D CRT, $F(1, 35) = 3.62$, $p < .05$. They produced higher collaborative ratings for the 3D CRT than for the other 3D displays (Curved, DataWall, and CAVE), $F(1, 35) = 3.22$, $p < .05$. In fact, participants were observed having difficulty communicating with each other when wearing 3D glasses in the DataWall and the CAVE conditions. Especially in the CAVE condition, most of them expressed difficulty in controlling the Dataglove to manipulate the models. In some cases, participants commented that models were unintentionally manipulated and they expressed frustration while one viewing the model and the other accidentally moved it. Many teams showed reluctance exploring the models through extensive interaction, and some of them viewed only one or two views of the models, presumably assuming that the models were static.

DISCUSSION

Participants viewing the 2D CRT display spent less time and detected fewer flaws than those using 3D displays. The results only partially support the hypothesis that the better performance and usability the more immersive technology used. Participants were better able to detect flaws with the 3D CRT than with the 2D display, but took more time, demonstrating a speed-accuracy trade-off.

Since 2D CAD models on the CRT were static and provided little depth information, they may have been more difficult to understand and interpret than 3D models (see Figures 1 and 2). In addition, the participants were unable to interactively explore the models. The flaws they did identity (i.e., low visual display, poor screen angle, and distractive location) were salient in 2D models, but other flaws (i.e., obstructed vehicle controls, controls not in easy reach, frequently used controls not in dominant hand position) may have been difficult to detect without manipulation of the model. The display usability evaluation also indicated that some participants thought 3D models more representative of the real world objects than 2D models. Thus, participants using the 2D display may have taken less time because they run out of identifiable flaws, leading to fewer flaws detected.

In addition, participants using the DataWall and CAVE displays spent less more time and detected fewer flaws than those using the 3D CRT and Curved displays. The usability testing results also indicated that the participants were not comfortable and confident to perform the task in the DataWall and CAVE displays. Although the DataWall and CAVE conditions had more immersive features, dynamic manipulation of the models may have been difficult for inexperienced users. In both DataWall and CAVE conditions, wearing 3D glasses degraded communication between team members. In the CAVE condition, interactive control of both the rotation (pitch, roll, and yaw) and translation (moving forward and backward) was through the manipulation of the fingertips on the Dataglove, which participants found difficult. This may have been why they terminated the trial more quickly than in other conditions where a mouse was used. The difficulty manipulating the models and communicating with each other may have led to short detection times and lower usability ratings.

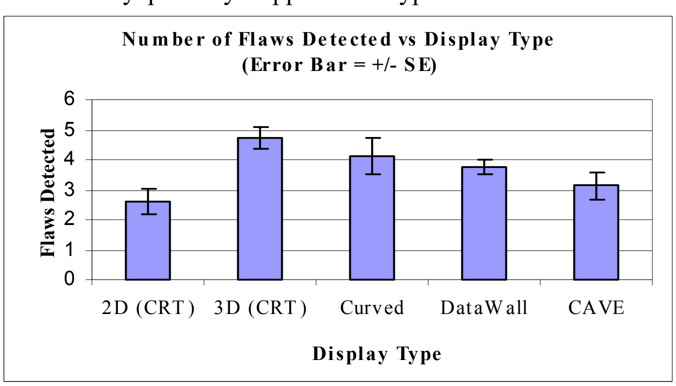


Figure 3. Effect of display type on number of flaws detected

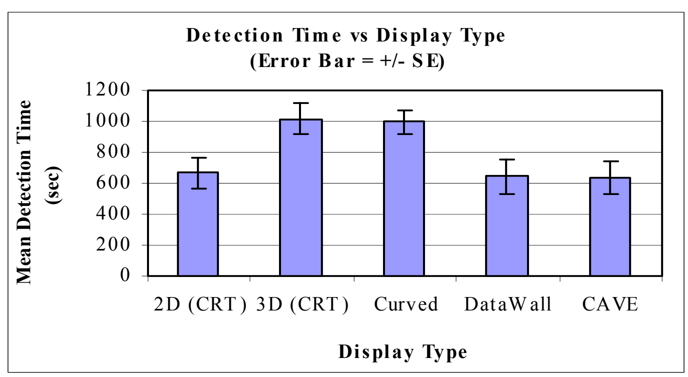


Figure 4. Effect of display type on overall detection time

Therefore, neither non-immersive technology nor fully immersive technology helped participants detect more design flaws. Only an inexpensive but partially immersive technology did facilitate the flaw detection. However, the improved detection did take longer. In other words, the level of immersion in virtual environments should be appropriately scaled to optimize users' interaction with the display media and facilitate collaborations. It is also possible that with more training users could take advantage of the more immersive virtual environment.

One might argue that participants in the 2D CRT condition were at a disadvantage because the CAD models could not be zoomed in this condition but they could in other conditions. Certain design flaws could be easily identified with the face value of 2D models without zooming, but other design flaws may be difficult to detect due to the complex nature of CAD models, as illustrated in Figure 1. However, all the design flaws were implemented to be detectable in both 2D and 3D CAD models. Therefore, the existence of zooming function in the 2D CRT condition should not significantly affect the comparison results.

CONCLUSIONS

From the analyses and discussion above, it can be concluded that:

a. In general, compared to partially or fully immersive displays, participants using the non-immersive display (i.e., the 2D CRT) spent less time and detected fewer flaws, demonstrating a speed-accuracy trade-off.

b. There was another speed-accuracy trade-off between partially and fully immersive technologies. The 3D CRT partially immersive display facilitated flaw detection better than other partially or fully immersive technologies (i.e., the Curved, DataWall, and CAVE displays) at the cost of increased time. In addition, a comparison of display costs shows that using the standard CRT would probably be more cost-effective than using the Curved, DataWall, or CAVE displays as long as an increase in detection time does not present a serious obstacle.

c. For immersive displays, performance and usability were not improved as the degree of immersion increased. Fewer flaws were detected in the fully immersive environment (i.e., CAVE) because the interaction with the environment was reportedly difficult. Therefore, the trade-off between level of immersion and ease of interaction in virtual environments should be balanced in order to maximize the collaborative review performance.

ACKNOWLEDGEMENTS

This work was carried out collaboratively by Defence R&D Canada-Toronto, Greenley & Associates Inc. (Ottawa), and Virtual Environments Technologies Centre, Integrated Manufacturing Technologies Institute, National Research Council Canada (London, Ontario). Support from the Synthetic Environments Coordination Office, Canadian Forces Experimentation Centre (Ottawa) is appreciated. The authors thank Justin Hollands for critical comments.

REFERENCES

Beevis, D., Vallerand, A., and Greenley M. (2001). *Technologies for Workload and Crewing Reduction* (DCIEM TR 2001-109). Toronto, Canada: Defence and Civil Institute of Environmental Medicine.

Bolstad, C.A. & Endsley, M. R. (1999) *The Effects of Task Load and Shared Displays on Team Situation Awareness.* Marietta, GA: Air Force Research Laboratory.

Commission of the European Communities (2000). *Commission recommendation on safe and efficient in-vehicle information and communication systems: A European statement of principles on human machine interface*. Brussels, Belgium: European Union.

Dynott, R., Lemoine, P. Piechulla, W., Stefani, O., Wakula, J. & Williams. M. (2002). *Project Report: Virtual Reality Systems for perceived ergonomic quality testing of driving task and design.* Traffic Research Group, Universität Regensburg. Commission of the European.

Green, P., Levison, W., Paelke, G. & Serafin, C. (1994). *Suggested human factors design guidelines for driver information systems*. Technical Report UMTRI-93-21. The University of Michigan, Transportation Research Institute.

Gwynne, J.W., Feher, B. A., Obermayer, R. W., Smillie, R. J., Linville, J. M., & Heacox, N. J. (1996). *Evaluating Collaborative Technologies for Command and Control Teams*. San Diego, CA: Office of Naval Research.

Hoeksema-van Orden, C.Y. (1998). *Effects of Computer Support on Communication and Decision Making: A Review of the Literature (Effecten van Computer-Ondersteuning op Communicatie en Besluitvorming: Een Literatuurstudie).* Soesterberg, Netherlands: TNO Human Factors Research Institute.

Japan Automobile Manufactures Association (2000). *Guideline for In-Vehicle Display Systems.* Version 2.1.

Kline, T. J. B. & McGrath, J. (1999). A Review of the groupware literature: Theories, methodologies, and a research agenda. *Canadian Psychology, 40*, 265-271.

Owen, S. (1999). *HyperVIS: Teaching Scientific Visualisation Using Hypermedia. Report to the* Project of the ACM SIGGRAPH Education Committee, the National Science Foundation, and the Hypermedia and Visualization Laboratory, Georgia State University.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd ed.). Upper Sadde River, NJ: Prentice Hall.