An Evaluation of New Console Technology – Large Display – in Process Control Display

Benjamin Noah, Jingwen Li, and Ling Rothrock(✉)

Harold and Inge Marcus Department of Industrial and Manufacturing
Engineering, Penn State University, State College, PA 16801, USA
lxr28@engr.psu.edu

Abstract. The objectives of this study were to test the effect of display layout/screen type on performance in a process control task (managing a tank farm). The study compared the following two conditions: (a) 4K-resolution 55″ screen with keyboard/mouse versus (b) 6-pack screens with keyboard/mouse. A within-subject experiment was conducted among 20 college engineering students. A primary task of preventing tanks from overfilling as well as a secondary task of manual logging with situation awareness questions were designed. Primary Task performance (including tank level at discharge, number of tank discharged and performance score), Secondary Task Performance (including Tank log count, performance score), system interaction times, subjective workload, situation awareness questionnaire, user experience survey regarding usability and condition comparison were used as the measures. The 6-pack setup was found to be slightly outperformed the 4K setup in tank discharge percentage. Detection+Navigation time was approximately one second shorter in the 4K condition compared to the 6-pack condition. On the other hand, the 6-pack condition outperformed the 4K-screen in the time to enter the values by two seconds. It was also found that the total time it took participants to properly discharge tanks was not significantly different between these conditions. In terms of the subjective feedbacks, participants felt equally about the 4K-screen and 6-pack conditions. More experiments need to be conducted to resolve some of the issues and come to a clearer conclusion.

Keywords: Large display technology · Bazel · Process control display

1 Introduction

In the process control room, operators need to supervise dynamic processes, recognize unplanned disturbances and/or predict before they occur so that the proper corrective measures can be carried out in order to ensure steady state operation. Thus, process control display has attracted much attention from industry and academic since operators rely on these display design to receive information. While many challenges are being overcome through good human-centered display design and proper human-machine system evaluations, the constant arrival of new technology provides both solutions and alternative challenges. New technology may be capable of providing useful features which were not previously available, but–at the same time– usability and safety requirements need to be verified prior to implementation. For example, head-mounted

D. Harris (Ed.): EPCE 2017, Part I, LNAI 10275, pp. 343–361, 2017.
DOI: 10.1007/978-3-319-58472-0_27

displays (HMD) have been shown to provide potential benefits within multiple domains [1, 2]. Heads up displays (HUD) is a similar technology which can be used for improving control in robotics [3]. Gesture control is being researched for navigation and basic control (e.g. [3, 4]). Going a step further, physiological (brain and body) techniques for control experiments have also been conducted (e.g. [5]). While these technological advancements present opportunities for improving the human-machine system, implementation of these newer technologies within the domain of process control requires careful analysis. The introduction of new forms of interaction within a field that relies on familiarity and usability could backfire if not done properly.

This study investigates the introduction of large high resolution displays that are currently available and are potentially ready for serious implementation considerations within process control. In this study, it's tested against the traditional, ubiquitous, use of multi-monitor consoles using standard keyboards and mice for interaction. The following sections present a brief review of literature on the factors which are directly relevant to the current study, based on which we will propose our hypothesis.

Spatial layout of information, i.e., where and how to place visual objects for the users [6], is one of the key factors that has been studied to address some of the aforementioned challenges. It was determined by Vincow and Wickens's study [7] that as more information integration was required, performance was negatively affected as the spatial distance between pieces of information increased, therefore, they suggested that items should be grouped closer together during higher levels of information integration, working memory load, and stress. This finding is in agreement with earlier, and more fundamental, studies which suggested that people are better able to recall information when presented with many attributes of a few objects, rather than a few attributes of many objects [8, 9]. Hess et al. [10] found that screen layout had a direct effect on the cognitive demands of a task, which was measured in the accuracy and response times of participants. The findings were supported by other studies like: Kandogan and Shneiderman [11] found that it is better to use hierarchically organized displays that are simultaneously presented, in a tiled layout, during dynamic task-switching work environments; Jang [12] also found that the layout of information affects performance, concluding that it is important that users are presented with multiple sources of information when integration tasks are being conducted. Simonin et al. [13] found that the radial layout outperformed the others in the visual search task. In summary, layouts need to be properly oriented in order to achieve optimum performance.

Screen/monitor arrangement is directly associated to the information layout, and both multiple-monitor and large screen setups are commonly used for control environments. Czerwinski et al. [14] demonstrated that there is a significant performance advantage to use very large, multiple monitor surfaces while carrying out complex, cognitively loaded productivity tasks. Generally, the multiple monitor layout is better in terms of inducing a cognitive layout or mental map for the user. It is assumed that users adopt a cognitive layout of the type of information to be presented and the relationships among the windows or screens and the information that they contain [15]. While it is clear that the layout of information is significant, the optimized layout is going to be dependent on the domain and the task requirements.

Within-screen Factors-Bezels: While it may at first seem that bezels are troublesome– and some research identifies potential bezel- related issues– most research has largely shown the contrary. Robertson et al. [16] found that bezels help organize work into different activities. This finding was supported by Ball and North's study [17], which proved that bezels between the monitors acted as natural dividers to help orient the participants, preventing them from getting lost in the display space in a target search task. It could be that bezels acted as dividers between different displays, and that these dividers are useful to have. However, bezels can also have some drawbacks. Going back to the study by Robertson et al. [16], they additionally found that bezels could present some issues if information is cut-off, creating visual discontinuity. This is often experienced when users are moving a mouse cursor between screens, where there is no compensation for the physical space that exists between the virtual spaces. The bezels also made reading tasks and perceiving image patterns more difficult. In a more recent study, Bi et al. [18] found that increasing the number of divisions by bezels to not effect performance in visual search and target selection tasks as long as objects are not split by the interior bezels. Bi et al. [18] found that splitting objects to have a negative effect on search accuracy and that bezels hinder performance on a straight tunnel steering task (i.e., click and drag between parallel lines between displays). Thus, interior bezels often constrain the sorts of possible layouts that are possible without splitting display objects across bezels. White space (or unused space) within a display has also been studied and found to not impact performance for search tasks [19]. In summary of the aforementioned studies, it can be suggested that bezels–or separation between displays– is important and helpful as long as the displays provide unique (individual) pieces of information. On the other hand, bezels can hinder performance if using the mouse for precision between screens is needed.

Within-screen Factors-Resolution: Transitioning to screen resolution, there is some evidence in support of using higher resolution displays. Ball and North's study [17] showed that high-resolution displays can be a benefit in that they significantly improve performance time for basic visualization tasks in finely detailed data, and they help people find and compare targets faster (up to twice as fast), feel less frustration, and have more of a sense of confidence about their responses. The same study found that there was more physical navigation (physical bodily movement) for high-resolution displays while more virtual navigation (i.e., zooming or panning in) in low-resolution displays. Also, there appeared to be a greater amount of frustration when dealing with pan + zoom as opposed to physical navigation. These findings are supported by a later study, which found that increased physical navigation on larger displays correlates with reduced virtual navigation and improved user performance [20]. These two studies favor high resolution displays. A review of using this technology within process control rooms is needed to determine if similar, and/or other, advantages exists.

Within-screen Factors-Size: Lastly, display size is another factor which has been previously investigated. The work by Andrews et al. [21] provides a very good overview and discussion of large display technology. Some have found that the larger displays should be used for higher cognitive load tasks so that less switching occurs [14, 22]. Large format displays have been shown to provide value in multiple domains, such as in military applications (e.g. [23]) and medical applications (e.g. [24]). Perceptual tasks benefit most from large displays as they allow for quicker navigation, such as navigating a map or a visualization of genes–as within the two examples cited above. The process control industry uses visualizations of similar complexities and more testing is needed to determine if these benefits can be realized within this domain. There are some challenges that exist with using larger displays. For instance, the displays typically require that they are positioned further away from the users, which can make it more difficult to control [21]. This can present issues with controlling, selecting, navigating, and linking [21]. It is common to lose sight of the mouse cursor, for example, while interacting with a large display [25]. It is suggested that, due to these interaction challenges, an alternative method to interact needs to be incorporated when using large displays [21, 26]. In this study, the use of a touchscreen is investigated with the large 4K resolution display. This may allow users to overcome some of the aforementioned challenges. It should also be noted that the current study confounds high resolution with a large display screen, simply due to the nature of the technology.

The system investigated within this study was primarily a hardware change– using a large 55 in. 4K flat screen compared with traditional multi-monitor console environments. The comparison confounds display layout with screen type as it was necessary to modify the layout of displays within the newer system, and which could not be easily replicated within the traditional system. However, effort was made to ensure that equal information was available between these systems. Therefore, the objectives of this study were to Test the effect of display layout/screen type on performance in a process control task.

2 Methods

2.1 Industrial Background and Displays

This section describes the industrial background for the primary task in this study. A customized version of the industry-leading supplier's Console Station software was developed. The primary task (A full description can be found in the *Tasks* section) involved monitoring three tank farms within three independent areas: Crude Area, Blending area, and Product area. Participants needed to prevent tanks from overfilling by detecting when a tank began filling and manually discharging the filling tanks. Three levels of displays that were used in both monitor types were described as follows, with Level 1 displays used for monitoring, Level 2 for identification of a filling tank and for navigation, and Level 3 for discharging tanks.

Level 1-Tank Farm Overview Display: As shown in Fig. 1, the Tank Farm Overview Display, providing the operator with a high-level picture of the state of the processes, shows three individual and independent tank farms: Crude Area, Blending Area, and Product area. Each area was designed with the same process structure, each containing exactly 16 individual tanks, meaning while a single tank begins to fill, and whether it overfills or is properly discharged, there was no dependency between tanks. There are 16 vertical level indicators in each area (total 48). The plot immediately below the level

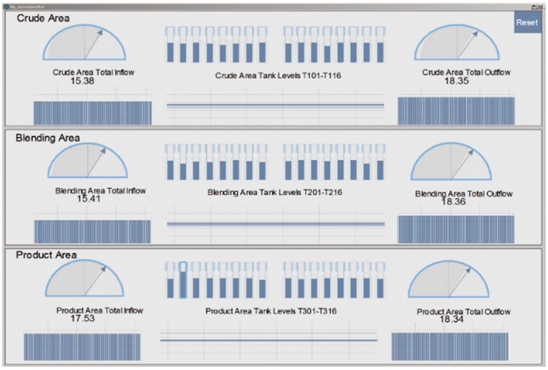


Fig. 1. The L1 tank farm overview display (Level 1, simulation is off)

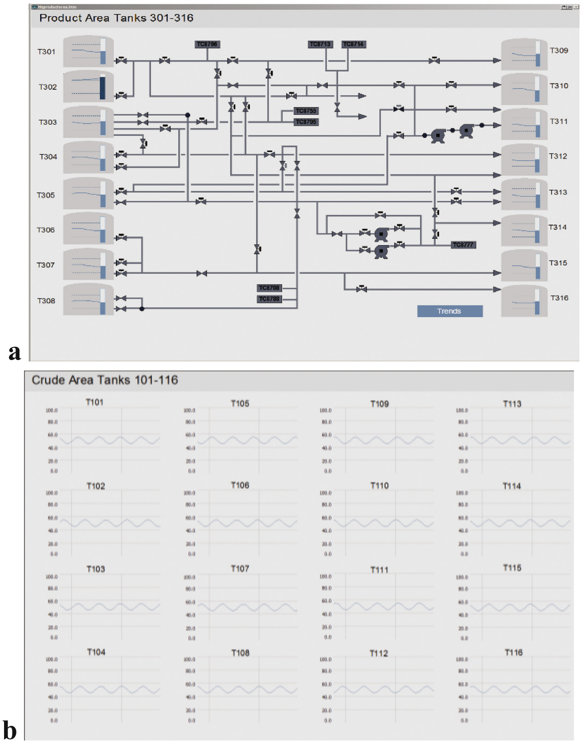


Fig. 2. Level 2 display (a) the product tank area overview (b) the crude tank area trend

indicators is a trend line chart which shows the level of all 16 tanks simultaneously, with the vertical axis representing the tank level in percent capacity (0–100%), and the horizontal axis as the time in seconds. The most current levels are plotted on the right and these levels are then updated every second, shifting the graph to the left. Total inflow and outflow indicators (including analog, dial type, flow gauge and trend bar chart) were placed on the left-hand side and right-hand side of the display respectively. The overall time range shown in both chart types is about one minute.

Level 2-Tank Area Overview Display and Trend Display: If click any of the three areas (Crude, Blending and Product) in Level 1 display, the Level 2 would be shown. Figure 2a shows the Tank Area Overview Displays for the Product and Fig. 2b shows the Tank Area Trend Display for Crude. The only differences between the three areas are the title and tank numbers (tank labels). The schematic lines and objects in the middle of the Area Overview Display (Fig. 2a) were not used in the experimental scenarios. There are 8 tanks on the left edge and 8 on the right edge, ordered numerically. Each includes two indicators: an immediate line trend chart and a level bar. These tank shapes could be clicked to access the lower level tank detail display (Level 3 display in Fig. 3). The 'Trends' button, located in the bottom right, could be clicked to access the Level 1 display for that respective area. In summary, if a tank filling event begins and is detected, the participant would need to then interact with the Tank Area Overview Display to navigate to the proper tank which is filling. The Tank Area Trend Display (Fig. 2b) shows what the trend line chart shows within the Tank Farm Overview Display (Level 1), but it separates each tank so that there are 16 individual charts. The title of each chart could be clicked to access the respective tank detail display, as an alternate navigation path to the Level 3 displays. The purpose of this display was to provide a secondary monitoring display to see what has recently happened in a single area. It could be stated that using these displays was optional as it was not mandatory to use them in order to perform well. Another type of display will be described later in the Design of Experiments.

Level 3-Tank Detail Display: Figure 3 shows a Tank Detail Display (Level 3) of Tank 101 within the Crude Area. As each area contained 16 tanks, there were 48 individual Tank Details Displays. This is the final display that users were required to navigate to in order to properly discharge a tank to prevent overfilling. On the left edge of this display, a navigation pane allowed users to click tanks within the same area

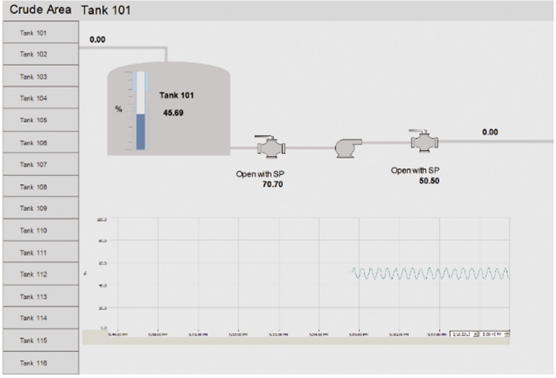


Fig. 3. The tank detail display for tank 101 (Level 3)

which would bring up the Tank Detail Display that was requested. The bottom portion of the display contains a line trend chart which shows tank level over a longer period of history than all other trend charts used in other displays. Additionally, the horizontal axis is labeled with the actual time in minute increments. The tank is shown with the tank shape in the top left, which also contains within it a level indicator and the numerical value of that level. This value is the percent that the tank is currently at with respect to its maximum capacity. Finally, to the right of the tank shape there are two valves and a pump along an exit pipe. These objects needed to be clicked, each individually, in order to call up faceplates which needed to be interacted with. Essentially, the two valves required to be opened at the set points (randomly set) shown immediately below each object using the number keypad on a keyboard. The pump needed to be turned ON using a dropdown selection panel by using a mouse. Once all three were properly opened, the tank immediately began to discharge–completing the required actions for that specific tank fill event.

2.2 Design of Experiments

Independent Variables. The study compared following two conditions: 4K large screen with keyboard/mouse versus 6 standard screens with keyboard/mouse (Fig. 4). The purpose was to investigate the effect of screen type and information layout within those screens.

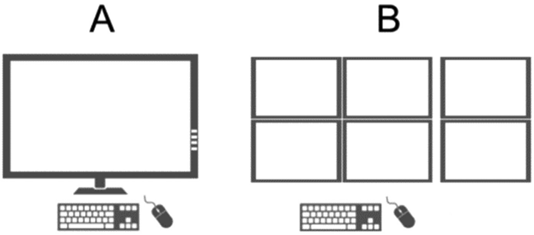


Fig. 4. Study tested case A (4K screen) versus case B (6-pack screen)

Therefore, the independent variable is the screen/layout variable, and it is a confound between screen hardware and the information layout of displays. The 4K screen used the layout that is shown in Fig. 5a. The 6-Pack screen setup used the layout that is shown in Fig. 5b. Noticing that in the 6-Pack screen layout, the right bottom one is Tank Farm Trend Display which only exist in this condition. This display shows trend line charts for all tanks in all three areas. The reason of including the display is to overcome the difference in information availability between these two conditions. As the 4K screen provided the user with the ability to view all of the three Tank Area Trend Displays simultaneously (48 tanks), the 6-monitor experimental condition only allowed for viewing of a single Tank Area Trend Display. Thus, the Tank Farm Trend Display was present in a dedicated monitor screen at all times for the 6-pack condition.

This display could not be interacted with in any form and was made to be used as a visual reference, similar to the Tank Farm Overview Display (Level 1).

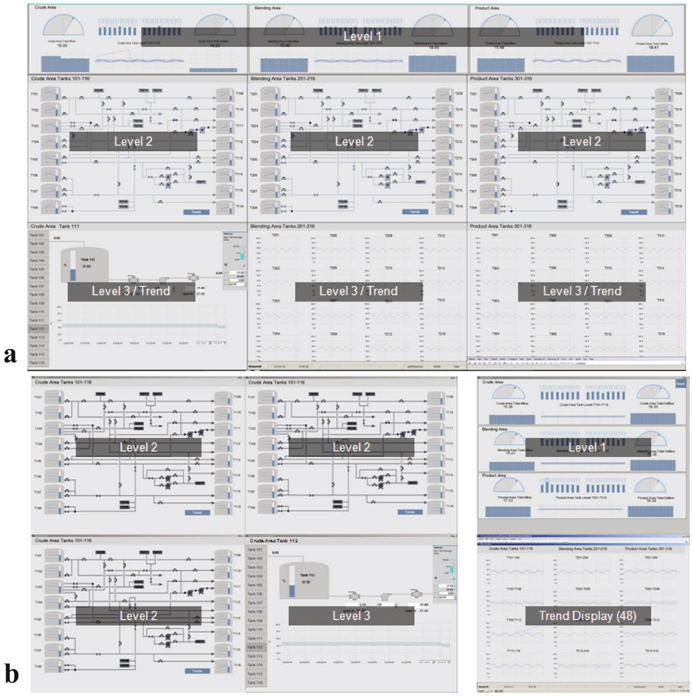


Fig. 5. Layout for two conditions (a) 4K screen (b) 6-pack screen

Tasks. Two tasks were used in the experiments. The primary task was monitoring and discharging tanks within a simulated tank farm. The secondary task was a manual logging task which used paper and pen to maintain a continuous log of tank levels. Two tasks were given equal priority.

Primary Task: The goal of the primary task was to prevent tanks from overfilling. *In* order to accomplish this, manual discharge was required through a series of actions. First, the tank that was undergoing filling needed to be identified and navigated to (to the Level 3 display). Then two valves needed to be opened and a pump needed to be turned ON. Navigating and interaction with the valves and pump required mouse clicks. Opening the valves required typing in the specified set points on the number keypad within the keyboard. Turning ON the pump required using the mouse to select the pump and then to select the 'ON' option from a dropdown menu (which only contained two choices: ON or OFF). Discharged tanks would drain down back to 50% level and the simulation automatically closed the valves and turned the pump off.

Scenario Scripts for Primary Tasks: The simulated scenarios followed pre-randomized fixed scripts with the following constraints: First, the total duration of a scenario was 20 min and scripts were broken down into 20 individual minute-blocks. Second, there were a total of 15 tanks which started filling with no overlap. Third, the time for tank filling and the time for reaction was kept within these minute-blocks. A tank fill was randomly determined to occur within the first 22 s of each minute-block. The fill rates were constant and equal between all tanks, filling from 55% level to 100% level in 38 s. Fourth, the first minute of each scenario was kept free of any event for the purpose of allowing participants to orient themselves. Fifth, four situation awareness questions were randomly preselected within individual minute-blocks, avoiding the time that a tank was filling. Figure 6 shows a breakdown of an example script, where SA refers to blocks for situation awareness questions. Scripts were created for both the experiment and also the practice scenario to train participants.

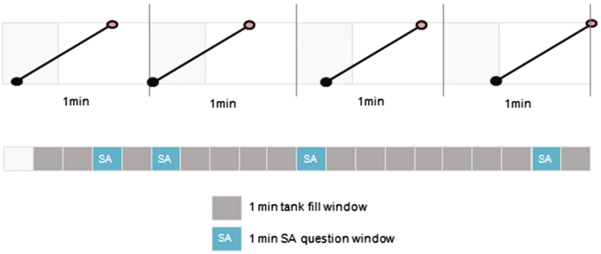
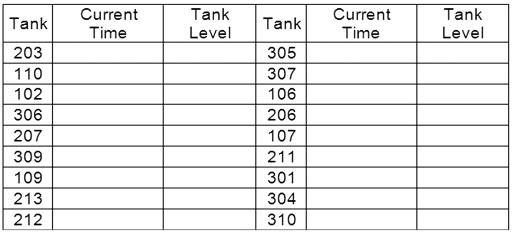


Fig. 6. An example of the structure of the experimental scenario scripts. (Color figure online)

Color indicators for primary task: The displays included warning and alarm indicators for tank fills as: no visual indicators for normal operating level range of 45%–55%; light-blue-colored, abnormal-high-warning indicators for range of 70%–79%; yellow-colored, high-alarm indicators for range of 80%–89%; red-colored, high-high-alarm indicators for range of >90%. If a tank was discharged, the colored indicators would remain present until the tank level fell back down below the aforementioned ranges. Finally, if a tank was not discharged prior to reaching 100% level, the simulation automatically reset that tank to 50% level and normal fluctuation resumed (This would be recorded as a missed event).

Secondary Task: In order to provide additional workload, the secondary task was designed as a manual logging task, which required participants to fill out a log using paper and pen as many as possible. The log listed tanks from all three areas in a randomized order (an example is shown in Fig. 7). The participants were instructed to start with the first column (the left side column) and proceed to complete each row on the log for the respective tank. The secondary task began and ended at the same times as the primary task. The final performance score on the secondary task is calculated by counting the number of tanks logged within the experimental time.



Tank	Current Time	Tank Level	Tank	Current Time	Tank Level
203			305		
110			307		
102			106		
306			206		
207			107		
309			211		
109			301		
213			304		
212			310		

Fig. 7. The manual tank logging task.

Dependent Variables. Table 1 lists all the dependent variables which were measured. For the primary task, a performance metric was calculated by subtracting the discharged level percentage from the maximum capacity (100%). Thus, higher scores are possible if tanks are discharged quicker. The score values are based on the average of all tanks discharged, which included tanks that were not properly discharged, as a zero value. The final performance score was then normalized to the 0–1 range, where 0

Table 1. The dependent variables in the study

Dependent Variable		Metric(s)	Detail
Primary task	Number of tanks discharged	Count (#)	Tanks could be discharged or overfilled
	Level % at time of discharge	Percentage (%)	As tanks fill, their level % increases and a quicker discharge ensures a lower %
	Performance score 1	Standardized score (#)	Range 0 to 1, dependent on tank level % at time of discharge
Secondary task	Tank log count	Count (#)	Number of tanks logged on log sheet
	Performance score 2	Standardized score (#)	Range 0 to 1, standardized on the individual with the highest number of tanks logged
System interaction times	Detection +Navigation time	Time (sec)	Time to reach the Tank Detail Display (Level 3) from the time when tank starts to fill
	Data entry time	Time (sec)	Time to interact with the Tank Detail Display in order to properly discharge filling tanks
	Tank discharge time	Time (sec)	Sum of (Det+Nav) and (Data entry) times
Other	Subjective workload	Index (#)	NASA TLX, scale: 0 to 100
	Situation awareness	Time to say 'ready' (sec)	SPAM Technique (see below), 4 questions per scenario, 8 questions total per participant
		Time to answer (sec)	
	Subjective situation awareness	Likert scale responses	4 subjective situation awareness questions
	System usability scale	Likert scale responses	14 usability statements (positive and negative)
	System comparisons	Binary responses	4 questions

means that no tanks were properly discharged and 1 means that all tanks were discharged immediately when filling started. For the secondary task, a performance score was calculated based on the number of logs and the individual participant who obtained the maximum number of log entries. These scores were also standardized within the 0 to 1 range. Finally, performance score 1 and 2 were summed for a total performance score, giving equal weight to each primary and secondary task.

Figure 8 represents the three times which were extracted from the experiment for the System Interaction Time metrics.

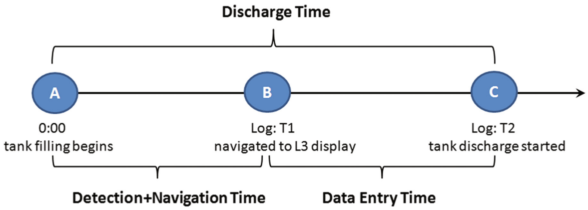


Fig. 8. System interaction times

Situation awareness was first measured with the Situation Present Assessment Method (SPAM) [27]. SPAM is an online, real-time, assessment technique to measure situation awareness. The key metric with this technique is to measure the amount of time participants take to respond to the situation awareness question. Subjective situation awareness was also measured using four Likert-scale type questions at the conclusion of each experimental scenario/condition.

2.3 Participants and Procedure

Total of 20 participants (11 males and 9 females, age range: 20–28) were recruited for this within-subject experiments. The order of the two conditions were balanced. The procedure of the experiments is shown in Fig. 9.

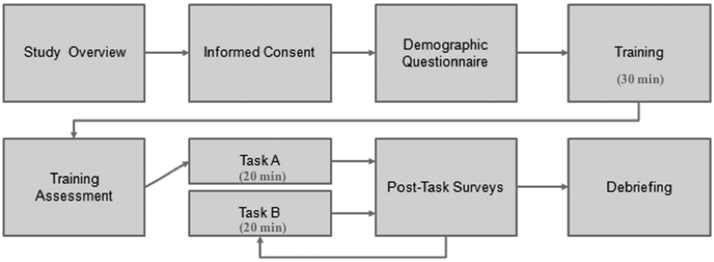


Fig. 9. Experiment protocol

2.4 Hypothesis

H1: There will be a statistical difference in performance measures between conditions

H1(a): Primary task performance scores will be different between conditions.
H1(b): Secondary task performance scores will be different between conditions.
H1(c): Detection+Navigation times will be different between conditions.

H2: There will be no statistical difference in data entry times between conditions
H3: There will be a statistical difference in subjective workload between conditions
H4: There will be a statistical difference in situation awareness between conditions
H5: There will be a statistical difference in system usability between conditions
H6: Users will prefer one condition more than the other

2.5 Data Analysis Approach

Performance data includes: primary and secondary task performance, system interaction times, situation awareness (SPAM), and subjective workload (NASA-TLX). This data was parametric and met the requirements for making the assumption of being normally distributed (tested using Anderson-Darling tests). As the comparisons were all conducted within-subject, paired t-tests were conducted to test if the means of each measure were equal (the null hypothesis). Therefore, if the paired t-tests produced results that were significant at $\alpha = 0.05$, then the null hypothesis was rejected–indicating that a difference might exist. Along with the paired t-tests, the effects size was calculated using Cohen's d parameter. From the effects size parameter, the power of the statistical test was finally calculated.

The survey data includes: the System Usability Scale, the Subjective Situation Awareness Questionnaire, and the Systems Comparisons Questionnaire. The first two questionnaires allowed for Likert scale (non-parametric) responses along a five-level scale, with the middle option always being neutral. The Wilcoxen Signed-Rank Test was used as the non-parametric version of the paired t-test to conduct similar comparisons as in the previous section (within-subject). However, two types of Wilcoxen Signed-Rank Tests were conducted. The first was a single variable test which took the median response for a question, in a single condition, and tested to see if it was equal to neutral (the null hypothesis). A significant result at $\alpha = 0.05$ would lead to the rejection of the null and indicate that the median response is not equal to neutral. Next, the Wilcoxen Signed-Rank Test was used like the paired t-test, but instead of means it compared medians between conditions. Again, the null hypothesis was that the medians were equal, and significance was determined at the same $\alpha = 0.05$ criteria. The Systems Comparison questionnaire allowed for only binary results and no statistical analysis was conducted. Instead, the direct number of responses in favor of one condition versus the other are reported.

3 Results and Discussion

3.1 Performance Results and Discussion

Results: Table 2 contains a detailed summary of the performance statistical results and they are summarized in Table 3. The statistically significant results that were better than the alternate are highlighted in green. There were four metrics which resulted in being statistically different: Tank Level at Discharge, Primary Score, Detection+Navigation Time, and Data Entry Time. No significant results were found for: performance on the secondary task, situation awareness, and workload. Below is the review of the hypothesis and the corresponding results.

Table 2. Detailed performance results

Condition	**Primary Task Performance**			**Secondary Task Performance**		**Overall**	**System Interaction Times(sec)**			**Situation Awareness**		**Workload**
	%Tanks Discharged	Tank Level at Discharge	Primary Score	Log Count	Secondary Score	Overall Score	Detection+ Navigation Time	Data Entry Time	Tank Discharge Time	(Ready) Time	(Answer) Time	NASA TLX
4K	*m*=0.97	*m*=0.83* *sd*=0.04	*m*=0.38* *sd*=0.09	*m*=69 *sd*=26	*m*=0.34 *sd*=0.13	*m*=0.72 *sd*=0.18	🟩 *m*=13.9** *sd*=2.2	*m*=12.0** *sd*=2.0	*m*=25.9 *sd*=3.1	*m*=3.2 *sd*=1.7	*m*=13.4 *sd*=3.7	*m*=54.6 *sd*=18.0
6 Pack	*m*=0.98	🟩 *m*=0.82* *sd*=0.03	🟩 *m*=0.40* *sd*=0.08	*m*=72 *sd*=24	*m*=0.36 *sd*=0.12	*m*=0.76 *sd*=0.17	*m*=15.0** *sd*=2.4	🟩 *m*=10.2** *sd*=1.8	*m*=25.2 *sd*=3.0	*m*=3.2 d=0.9	*m*=14.8 *sd*=4.6	*m*=52.2 *sd*=14.1
Paired-T Tests		*d*=0.284 *t*= 2.14 ***p*=0.0046** *Pw*=0.23	*d*=0.270 *t*= -2.16 ***p*=0.044** *Pw*=0.21		*d*=0.127 *t*= -1.10 *p*=0.284 *Pw*=0.08	*d*=0.217 *t*= -2.00 *p*=0.060 *Pw*=0.15	*d*=0.495 *t*=-2.46 ***p*=0.024** *Pw*=0.56	*d*=1.009 *t*=-6.63 ***p*=0.000** *Pw*=0.99	*d*=0.250 *t*= 1.79 *p*= 0.089 *Pw*=0.19	*d*=0.047 *t*=0.23 *p*=0.823 *Pw*=0.05	*d*=0.345 *t*=-1.38 *p*=0.185 *Pw*=0.31	*d*=0.154 *t*=0.89 *p*=0.387 *Pw*=0.10

Notes: Scores are out of 0.50 max per task, and out of 1.00 overall
** Significant difference at a=0.05, but low statistical power;*
*** Significant difference at a=0.05, with good statistical power*
d = Cohen's d; effects size; *t* = T-value; *p* = P-value; *Pw* = Power

H1: There will be a statistical difference in performance measures between conditions
Results: The data is both in support and against this general hypothesis:

H1(a): Primary task performance scores will be different between conditions.
Results: The data supported this with Tank Level at Discharge and Performance Score metrics but both of these were accompanied by low statistical power.

H1(b): Secondary task performance scores will be different between conditions
Results: The data did not support this hypothesis.

H1(c): Detection+Navigation times will be different between conditions.
Results: The data supported this with a statistical power of 0.56.

H2: There will be no statistical difference in data entry times between conditions
Results: The data did not support this as there was a statistical difference detected, with a statistical power of 0.99

H3: There will be a statistical difference in subjective workload between conditions
Results: The data did not support this hypothesis

H4: There will be a statistical difference in situation awareness between conditions
Results: The data did not support this hypothesis

Table 3. Summary results for the System Usability Scale questionnaire.

	System Usability- (Positive Statement)							
	#1	#3	#5	#7	#9	#11	#13	#14
Condition	Visually appealing	Easy to use	Functions well integrated	Learn to use quickly	High confi-dence	Made few errors	Satisfied with system performance	Good to use for a job
4K	Agree *	Agree *	Agree *	Agree *	Agree *	Agree *	Agree *	Agree *
6 Pack	Agree *	Agree *	Agree *	Agree *	Agree *	Agree*	Agree*	Agree*

	System Usability- (Negative Statement)					
	#2	#4	#6	#8	#10	#12
Condition	System complex	Need tech support	Inconsistent	Cumbersome	Had to learn a lot	Not remember how to use
4K	Disagree*	s.disagree*	Disagree*	Disagree	Disagree*	Disagree*
6 Pack	Disagree*	Disagree*	Disagree*	Disagree*	Disagree*	Disagree

*Notes: * indicates value is significantly different than neutral;*
(**) indicates significant difference between conditions at a=0.05

Discussion: It was found that the 6-pack setup slightly outperformed the 4K-keyboard setup in tank discharge percentage (primary task performance). However, the statistical power of this comparison was low. In addition to that, the mean values were 0.82 (6-pack) and 0.83 (4K-keyboard)–indicating that the difference was not definitive or likely not present. It is possible that the layout of the 6-Pack setup may have enabled better performance, but it will be shown below how system interaction times likely had a larger role in this result. No statistical differences were found for situation awareness and workload. This suggests that the experimental changes (independent variables) had no effect on the user's situation awareness level or on their perceived workload. Finding that there is not much of a difference between systems in the metrics mentioned above could also be perceived as comforting to industry members, in that introducing new technology does not adversely affect these basic performance metrics (within the context for these experiments). The question at this point might be 'how does the implementation of new technology provide benefit, if any exists?'

System interaction times presented more interesting findings, which helped in answering the question that was asked above. These findings were more in line with the expectations as information retrieval, system navigation, and system interaction are all aspects which should be affected by differences in information layout and interaction method. It was found that the combined time to detect a tank filling and then navigate to the proper Tank Detail Display was faster in the 4K-keyboard condition compared to the 6-pack condition. It took participants approximately one second longer to do these actions when using the 6-pack condition. It is suggested that the layout of the 4K-keyboard condition either enabled better tank level deviation detection, or enabled users to navigate quicker. The design of this study could not differentiate between these two possibilities. On the other hand, when looking at the time it took participants to enter the values (for discharging tanks), the 6-pack condition outperformed the 4K-keyboard condition, enabling users to enter data about two seconds faster. As both conditions used keyboards/mice as the interaction method, it could be suggested that the layout differences must have played a significant role. However, as the interaction method was identical in both conditions, this came as a surprise. Fortunately, there is a reasonable explanation for what was observed. It should be noted that there was a noticeable delay in calling up displays on the 4K screen system, which would contribute to this difference[1]. The delay was primarily present when calling up the faceplate on the tank detail display, where values needed to be entered. Analysis of recorded video on both systems showed that the total delay in display call up to be between 0.5 to 1.0 s. This at least partially explains why participants took longer to enter data in the 4K screen condition. To sum up, it was found that the final time it took participants to properly discharge tanks to not be significantly different between the these conditions. So while the 4K-keyboard condition resulted in a faster Detection +Navigation time, the 6-pack condition resulted in a faster Data Entry time. The benefits of faster interaction times for each condition cancelled out one another when comparing the final discharge time, which was a summation of each system interaction time component (see Fig. 8). Had the 4K-keyboard condition not suffered from the display call up delays the overall result may have favored the 4K-keyboard condition, but further verification may be needed.

Regarding the findings in support of the layout used with the 4K screen (see Fig. 5a) compared to the layout within the 6-Pack condition (see Fig. 5b), there are some points that could be further discussed. While an equal amount of information was presented in both conditions, the organization of that information is important and can have an effect on performance as was observed within this study. Among many other researchers, Hess et al. [10] found that the layout of information affects the cognitive demands placed on individuals. The 4K screen contained columns for each of the three areas, whereas the 6-Pack screens did not have areas organized to easily recognized patterns. It could be hypothesized that one may be better than the other, but more research might be needed to verify such claims. There may be a larger effect which partially hid the effects of layout. Both conditions used a grid layout to separate the displays and this has been shown to provide real performance benefits [10, 11]. In addition, both conditions assigned specific screen space to displays, which is known to help with productivity tasks [14]. This study supports the findings of research done in different contexts and domains and only goes to further emphasize the importance of proper information layout. While much work has investigated the proper design of visualizations within displays, more work could be potentially be done in area of display layout to determine what organization of displays provides the most benefit to the console operator. A benefit of the 4K screen is that it allows for tremendous flexibility in layouts.

3.2 Survey Results and Discussion

Results: Table 3 summarizes the results for the System Usability Scale questionnaire. In summary of the statistical results, most responses generally agreed with the positive usability statements and disagreed with negative usability statements. In addition to that, no statistical differences were found in either of the studies.

---

[1] The delay in display call up was only discovered midway through data collection following comments made by some participants. The delay was likely due to a system configuration error in the 4K screen setup that is easily corrected.

Table 4 shows the summary of subjective situation awareness questions. There were no statistical differences found between two conditions. Three out of the four questions showed responses which were statistically different than neutral (see asterisks noting). Table 5 shows the responses for the Systems Comparisons. No statistical analysis was conducted on these responses. It can be seen that the study produced similar responses between conditions. There were issues with the use of: keyboard, mouse. These issues can be known from the optional comments which participants provided (are reviewed within the Discussion section). These were not technical in nature, but rather preference in usability. The relevant hypotheses proposed are reviewed below.

Table 4. Summary results for the four subjective situation awareness questions.

Condition	Subjective SA questions			
	#1	#2	#3	#4
	Aware of levels	Able to detect	Quickness to detect	Aware of 50% recovery
4K	Somewhat aware*	Somewhat able*	Somewhat/very quickly*	Neutral
6 Pack	Somewhat aware	Somewhat able*	Somewhat quickly*	Neutral

* *Indicates value is significantly different than neutral*

H4: There will be a statistical difference in situation awareness between conditions
Results: The data did not support this hypothesis

H5: There will be a statistical difference in system usability between conditions
Results: The data did not support this hypothesis

H6: Users will prefer one condition more than the other
Results: The data did not support this hypothesis

Discussion: The System Usability Scale questionnaire did not present statistical differences between the two conditions. For the most part, participants agreed with the positive usability statements and disagreed with the negative usability statements. The subjective situation awareness questions likewise, did not show any significant differences between conditions, and results were not anything out of the norm (i.e., no extreme responses at either end of the Likert scale). These questionnaire results could indicate that there are no significant concerns with any of the experimental conditions for the task that was used. Finally, the subjective comparisons questionnaire resulted in that participants felt equal preference to the 4K-keyboard and the 6-pack conditions. Eight participants did report that they had some issues with using the mouse. Most of these were related to difficulty in locating the mouse icon within the screens, for both conditions. As it can be seen, there are some useful design-related comments, but there was no single common concern among participants.

Table 5. Responses for the Systems Comparisons questionnaire.

Condition	Subjective Comparisons			
	#1	#2	#3	#4
	Condition preference	System preference	Issues with keyboard	Issues with mouse
4K	11	10	0 (yes)	8 (yes)
6 Pack	9	10	20 (no)	12 (no)

4 Limitation and Future Work

The limitations are along the lines of the design of the experiment. The first limitation is the laboratory setting and the participant pool being college students. A validation of these results should be tested in the field with console operators. The second limitation is the task used being relatively basic compared to real-world control room operations. Again, a field investigation could incorporate multiple tasks which require regular interaction with the equipment. Stating this, it is expected that the results are generalizable to industry as the significant effects are fundamental to human-computer interaction – expertise is not expected to have an effect on the system interaction times for example. Another limitation is in reference to the equipment that was used. As a delay in display call-up was encountered within the 4K conditions, some of the effects became less clear in the analysis. This can be overcome in future investigations through proper hardware/software implementations.

5 Conclusion

This study compared the 4K display with the 6-pack displays using the keyboard/mouse interaction method. The 4K display's layout seemed to allow for the faster time in detection of tanks filling and navigation to the tank detail displays. However, the 6-pack condition showed faster data entry time (time to discharge tank by opening two valves and a pump). This resulted in slightly quicker tank discharges for the tanks in the 6-pack condition, seen by a lower average for tank level percentage at the time of discharge. The time to enter data should, hypothetically, have been similar as the interaction method remained constant. There was noticeable delay in display call up using the 4K screen, which could explain this discrepancy. So while the 4K screen layout seemed to allow participants to better navigate and detect tank fillings, the display call up delay may have slowed down data entry within the 4K setup. Therefore, there is some evidence to support that a 4K-keyboard setup could potentially outperform the 6-pack setup in these performance metrics if the display call up delay is resolved.

References

1. Phipps, D.: The human–computer interaction handbook: fundamentals, evolving technologies and emerging applications. Ergonomics **56**(8), 1345–1346 (2013)
2. Ruddle, R.A., Payne, S.J., Jones, D.M.: Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays? Presence Teleoperators Virtual Environ. **8**(2), 157–168 (1999)
3. Eliav, A., Lavie, T., Parmet, Y., Stern, H., Edan, Y.: Advanced methods for displays and remote control of robots. Appl. Ergon. **42**(6), 820–829 (2011)
4. Alcoverro, M., Suau, X., Morros, J.R., López-Méndez, A., Gil, A., Ruiz-Hidalgo, J., Casas, J.R.: Gesture control interface for immersive panoramic displays. Multimedia Tools Appl. **73**(1), 491–517 (2013)
5. Nelson, W.T., Hettinger, L.J., Cunningham, J.A., Roe, M.M., Haas, M.W., Dennis, L.B.: Navigating through virtual flight environments using brain-body-actuated control. In: Virtual Reality Annual International Symposium 1997, pp. 30–37. IEEE (1997). doi:10.1109/VRAIS.1997.583041
6. Lok, S., Feiner, S., Ngai, G.: Evaluation of visual balance for automated layout. In: Presented at the Proceedings of the 9th International Conference on Intelligent User Interfaces, pp. 101–108. ACM (2004)
7. Vincow, M.A., Wickens, C.D.: Spatial layout of displayed information: three steps toward developing quantitative models. In: Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 37, pp. 348–352. SAGE Publications (1993)
8. Harwood, K., Wickens, C., Kramer, A., Clay, D., Liu, Y.: Effects of display proximity and memory demands on the understanding of dynamic multidimensional information. In: Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 30, pp. 786–789. SAGE Publications (1986)
9. Yntema, D.B.: Keeping track of several things at once. Hum. Factors J. Hum. Factors Ergon. Soc. **5**(1), 7–17 (1963)
10. Hess, S.M., Detweiler, M.C., Ellis, R.D.:. The effects of display layout on monitoring and updating system states. In: Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 38, pp. 1336–1340. SAGE Publications (1994)
11. Kandogan, E., Shneiderman, B.: Elastic windows: improved spatial layout and rapid multiple window operations. In: Presented at the Proceedings of the Workshop on Advanced Visual Interfaces, pp. 29–38. ACM (1996)
12. Jang, J.: Lost in space: the surprising role of information spatial layout (2009)
13. Simonin, J., Kieffer, S., Carbonell, N.: Effects of display layout on gaze activity during visual search. In: Costabile, M.F., Paternò, F. (eds.) INTERACT 2005. LNCS, vol. 3585, pp. 1054–1057. Springer, Heidelberg (2005). doi:10.1007/11555261_103
14. Czerwinski, M., Smith, G., Regan, T., Meyers, B., Robertson, G., Starkweather, G.: Toward characterizing the productivity benefits of very large displays. In: Presented at the Proceedings of INTERACT, vol. 3, pp. 9–16 (2003)
15. Norman, K.L., Weldon, L.J., Shneiderman, B.: Cognitive layouts of windows and multiple screens for user interfaces. Int. J. Man Mach. Stud. **25**(2), 229–248 (1986)
16. Robertson, G., Czerwinski, M., Baudisch, P., Meyers, B., Robbins, D., Smith, G., Tan, D.: The large-display user experience. IEEE Comput. Graph. Appl. **25**(4), 44–51 (2005)
17. Ball, R., North, C.: Effects of tiled high-resolution display on basic visualization and navigation tasks. In: Presented at the CHI 2005 Extended Abstracts on Human Factors in Computing Systems, pp. 1196–1199. ACM (2005)

18. Bi, X., Bae, S.-H., Balakrishnan, R.: Effects of interior bezels of tiled-monitor large displays on visual search, tunnel steering, and target selection. In: Presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 65–74. ACM (2010)
19. Zdralek, J.: White space: how much nothing should there be? National Library of Canada = Bibliothèque nationale du Canada (2004)
20. Ball, R., North, C., Bowman, D.A.: Move to improve: promoting physical navigation to increase user performance with large displays. In: Presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 191–200. ACM (2007)
21. Andrews, C., Endert, A., Yost, B., North, C.: Information visualization on large, high-resolution displays: issues, challenges, and opportunities. Inf. Vis. (2011). 1473871611415997
22. Baudisch, P., Good, N., Bellotti, V., Schraedley, P.: Keeping things in context: a comparative evaluation of focus plus context screens, overviews, and zooming. In: Presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 259–266. ACM (2002)
23. Gouin, D., Evdokiou, P.: A showcase of visualization approaches for military decision makers. DTIC Document (2004)
24. Hibbs, M., Wallace, G., Dunham, M., Li, K., Troyanskaya, O.: Viewing the larger context of genomic data through horizontal integration. In: Presented at the 11th International Conference Information Visualization, IV 2007, pp. 326–334. IEEE (2007)
25. Czerwinski, M., Robertson, G., Meyers, B., Smith, G., Robbins, D., Tan, D.: Large display research overview. In: Presented at the CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 69–74. ACM (2006)
26. Moreland, K.: Redirecting research in large-format displays for visualization. In: 2012 IEEE Symposium on Large Data Analysis and Visualization (LDAV), pp. 91–95 (2012). doi:10.1109/LDAV.2012.6378981
27. Durso, F.T., Dattel, A.R., Banbury, S., Tremblay, S.: SPAM: the real-time assessment of SA. Cogn. Approach Situat. Aware. Theor. Appl. **1**, 137–154 (2004)