

The Influence of a User-Centred Design Focus on the Effectiveness of a User Interface for an Agricultural Machine

Aadesh Kumar Rakhra, Mitchell Kyle Green, Daniel Delmar Mann 

Department of Biosystems Engineering, University of Manitoba, Winnipeg, Canada

Email: danny.mann@umanitoba.ca

How to cite this paper: Rakhra, A.K., Green, M.K. and Mann, D.D. (2020) The Influence of a User-Centred Design Focus on the Effectiveness of a User Interface for an Agricultural Machine. *Agricultural Sciences*, 11, 947-965.

<https://doi.org/10.4236/as.2020.1111062>

Received: September 11, 2020

Accepted: November 6, 2020

Published: November 9, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

As agricultural machines become more complex, it is increasingly critical that special attention be directed to the design of the user interface to ensure that the operator will have an adequate understanding of the status of the machine at all times. A user-centred design focus was employed to develop two conceptual designs (UCD1 & UCD2) for a user interface for an agricultural air seeder. The two concepts were compared against an existing user interface (baseline condition) using the metrics of situation awareness (Situation Awareness Global Assessment Technique), mental workload (Integrated Workload Scale), reaction time, and subjective feedback. There were no statistically significant differences among the three user interfaces based on the metric of situation awareness; however, UCD2 was deemed to be significantly better than either UCD1 or the baseline interface on the basis of mental workload, reaction time and subjective feedback. The research has demonstrated that a user-centred design focus will generate a better user interface for an agricultural machine.

Keywords

User Interface, User-Centred Design, Agricultural Machine, Situation Awareness, Mental Workload

1. Introduction

A user interface is a means by which the user interacts with the target system. It is an important facet that helps the user to monitor, control and alter the target task environment [1]. The efficiency and effectiveness of the operation, as well as the workload and safety of the operator, depends on the information displayed to the user. Designers and researchers recommend that the typical user interface

should encapsulate and emphasize the critical features of the target environment [1] [2] [3] [4]. Using the paradigm of situation awareness, [5] proposed the goal-directed task analysis to determine the critical information needs of the user. Based on the situation awareness information needs of the operator, designers and engineers can design the user interface by focusing on the essential interface elements to accomplish operational goals of the operator.

There are two important considerations for designing a user interface: 1) information requirements of the user, and 2) information presentation to the user. Information requirements of the user specify the quantity, type or variety of the information deemed necessary for the user to achieve job-related goals. For example, a car driver requires knowledge of the current speed of the car; without the knowledge of this critical information, it would be difficult to drive safely and lawfully. Information presentation demonstrates the form, look, feel and mode of the information communicated to the user so that the processing and utilization of the information can be efficient. Referring back to the previous example, the current speed of a car can be presented to the driver in many forms: using numeric text, by showing the movement of a needle on a dial gauge, using a combination of both numeric text and animated needle, or by any other means based upon the imagination and resourcefulness of the designer.

In this study, we have designed and evaluated a driver interface for a tractor air seeder system. This work was accomplished in two phases. During the first phase, individual elements of the driver interface were designed and evaluated on the basis of mental workload invoked and level of situation awareness that was enabled. Experimental results confirmed that the metrics of mental workload and situation awareness can be used by the designer to select individual interface elements [6]. During the second phase of the study, individual elements of the tractor air seeder interface were further modified based on knowledge gained in the first phase of the study and then integrated into a complex interface. Two versions of a user interface were developed and evaluated against a pre-existing interface that had been used previously as part of a tractor-air seeder simulator. This manuscript discusses the findings of the second phase of the study.

2. Background Information Relevant to User Interfaces

2.1. Evaluating User Interfaces

Design of a user interface according to the user's goals and information requirements is half the battle towards building an effective interface. Although designers apply many interface design guidelines based on human factors principles, it is still common that certain aspects of the interface may not work as intended. Evaluation of a user interface helps designers to identify ineffective features and other issues to further improve the interface. Commonly recommended interface evaluation methods include heuristic evaluation, cognitive walk-through, usability testing, and guidelines/standard inspection [7] [8]. Heu-

Heuristic evaluation and usability testing are considered the most effective methods for improving user interfaces [7]. Heuristic evaluation requires many experts to evaluate the interface based on expertise gained over many years of professional practice. Usability testing yields data based on both objective and subjective evaluations which makes it more suitable for research and scientific studies.

Multiple metrics can be used for usability testing. Situation awareness may be considered as a primary means for evaluation, as this metric is a widely accepted criterion of evaluation in many domains [9]. Reference [10] defined situation awareness as “*the perception of the elements of the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future*”. From this definition, we can see that situation awareness consists of perception, comprehension, and projection. These are referred to as the three levels of situation awareness. Poor situation awareness could lead to ineffective, inadequate operational outcomes (perhaps leading to dangerous conditions). Reference [11] described several catastrophic airline crashes (Northwest Airlines MD-80 in 1987, US Air B-737 in 1989, or Korean Airlines Flight in 1983) which were directly or indirectly related to poor situation awareness of the operators of the automated flight system. A common reason for poor situation awareness relates to the presence of automation in the system. Automation may shift the role of the operator from “active participant” to “passive user or supervisor” [12]. In partially automated driving scenarios, adequate situation awareness is essential for the safety of the driver. During partially automated driving scenarios, human drivers are expected to take control of the situation whenever the situation demands attention (*i.e.* due to technology failure or technology limitation). This can be problematic as the driver has not likely been actively involved in decision-making leading up to the point of technology failure, and therefore, lacks complete understanding of the situation. It is for such reasons that it is critical to design a user interface that adequately supports the situation awareness of the user. Reference [13] developed and evaluated three interfaces for regenerative life support systems using the “ecological interface design” which considers the user-centered approach to better support the situation awareness of the operators. Results of the study have indicated that the interfaces which presented “situation-rich” information helped in better decision making.

Mental workload is another evaluation metric that has been used widely in human factors studies [14]. Mental workload can be defined as “the amount of cognitive capacity required to perform a given task” [15]. Evaluation of mental workload provides critical insights into design considerations and operational outcomes. As described by [14], most mental workload evaluation techniques can be categorized as analytical or empirical. The primary premise for this division is that analytical techniques (such as mathematical models and simulation models) do not require the operator to perform the task under investigation, while empirical techniques require the operator to perform the task under investigation. We can categorize empirical techniques into four divisions: primary task

performance (e.g. time or error related), secondary task techniques (e.g. loading or subsidiary task), physiological or psychophysiological techniques (such as cardiac or brain activity, eye function), and operator opinion or subjective techniques. Further details about all these techniques can be read in [14]. Subjective techniques (self-reports by the operators) are more “sensitive” and “accurate” [16] and operators show better judgment about their workload among varied task conditions [17]. Reference [16] developed and tested a unidimensional mental workload scale called the Integrated Workload Scale (IWS). This scale has shown advantages such as simplicity, ease of administration, speed of use, and minimal obstruction with the task.

A simple characteristic such as response time can also be used to evaluate a user interface. Reference [18] (cited by [19]) used the user’s response time in answering the questions related to information presented on an interface as a means for inferring the situation awareness attained by the user. Higher response time was associated with lower situation awareness. In another study related to workload, [20] reported that “accuracy decreased and reaction time increased as the difficulty of information processing requirements was increased”. Higher levels of subjective workload were associated with a lower level of performance and increased reaction time. In a study comparing two interfaces in an intensive care unit [21], lower response time and higher situation awareness was observed for an “integrated” display compared to the traditional display. Reference [22] mentioned that “an increase in task load led to lower situation awareness and higher mental workload, reduced mission success and increased mission times”. Overall, it can be concluded that higher response time can be associated with lower situation awareness, higher mental workload, or lower performance.

2.2. Interface Design Considerations Associated with Automation

The current trend is for automation to be incorporated into agricultural machines. Engineers are using sensor technology, combined with control systems, to automate various tasks that were previously completed manually. Although the operator may not be responsible for completing these automated tasks, the operator is still responsible for the overall operation of the machine. This implies that the operator should be provided with information that will enable him/her to fully understand the status of the machine—including the status of tasks that were completed autonomously. Therefore, the designer of a user interface should not forget to incorporate status information on tasks completed autonomously. If not, there is a chance that the operator will suffer from the so-called “out-of-the-loop” syndrome [5].

3. Research Method

3.1. Research Objective

There is a wealth of information that has been published in numerous textbooks

on the design of information displays. In-depth review of several different design principles was conducted by [23]. From this base of knowledge, an experimental study was completed in a controlled laboratory environment in which multiple versions of display elements (pictorials/symbols) relevant to the monitoring of an agricultural air seeder were evaluated [6]. Using metrics of situation awareness and mental workload [6], we were able to identify preferred display elements. The objective of this research, therefore, is to investigate the potential benefit to the operator associated with using a display interface designed from a user-centred perspective.

3.2. Design of an Air Seeder User Interface

Reference [6] identified 12 elements or functions that are most vital to the efficient operation of an air seeder. These elements were: seed level status (tank levels), fertilizer level status (tank levels), fan RPM, seed application rate, seed depth (tool depth), fertilizer application rate, fertilizer depth (tool depth), tool pressure, blockage status, desired path of the unit, desired location of the unit, and current speed of the unit. In an earlier study completed in the Agricultural Ergonomics Laboratory, the interface shown in **Figure 1** was developed for use with a tractor-air seeder simulator that was being developed for research purposes. At that time, minimal attention was given to the design of the symbols or pictorials chosen to represent the air seeder functions as the researchers were focused on development of a functioning simulator. Thus, the interface shown in **Figure 1** will be used as the baseline interface for purposes of comparison.

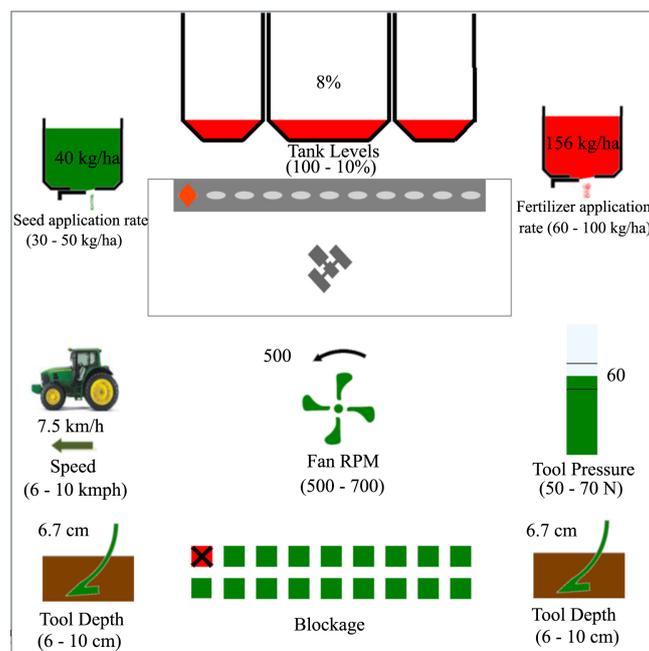


Figure 1. Air seeder interface originally developed for tractor-air seeder simulator in the Agricultural Ergonomics Laboratory at the University of Manitoba. This interface will serve as the baseline for comparison.

The individual elements that were compared in the experimental study by [6] are displayed in **Figure 2**. For the ease of presentation in this manuscript, the elements are displayed together.

To be able to achieve the objective of the current study, it was necessary to develop an integrated air seeder interface to compare against the baseline interface from the simulator (depicted in **Figure 1**). Individual elements from the study completed by [6] were used as the starting point. The first version of the interface designed according to user-centred design principles, hereafter referred to as UCD1, essentially consisted of the individual elements evaluated by [6], but with minor modifications to the elements used to display tool pressure, tool depth, and blockage (**Figure 3**). Furthermore, an element was added in the centre of the interface to provide guidance information.

A second version of an interface, hereafter referred to as UCD2, was developed based on the results and feedback from the work reported by [6] (**Figure 4**). The specific modifications are:

1) Tool Depth element: Some participants had difficulty making sense of a stationary tool with soil levels changing because this does not realistically reflect what would be happening with the machine. Participants also indicated a preference for the scale starting from the top to the bottom for the tool depth element. This feedback was used to develop an alternate element for depicting tool depth.

2) Blockage element: Most of the participants indicated a preference for green color over light blue color as an indication of the correct state (*i.e.* non-blockage state). Therefore, the color of the blockage elements was changed from blue to green.

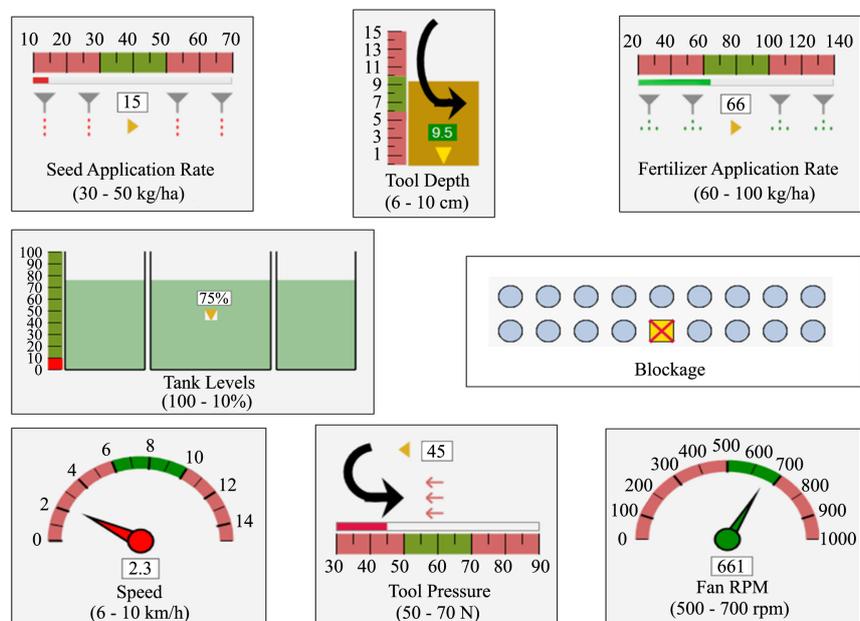


Figure 2. Air seeder monitoring elements that were evaluated in the study conducted by [6].

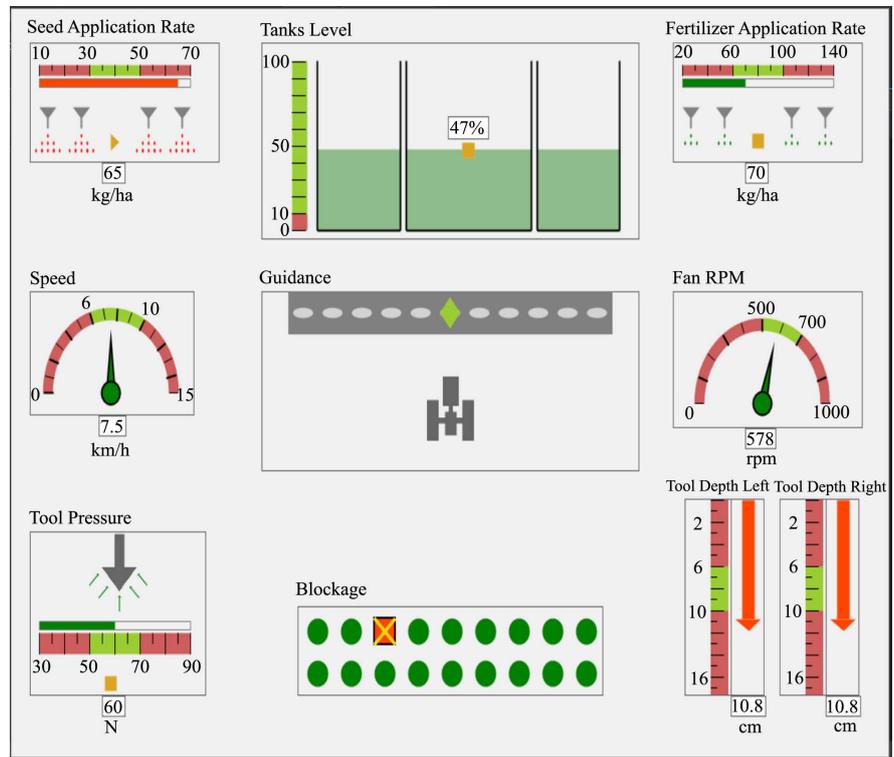


Figure 3. Version 1 of the air seeder interface, referred to as UCD1, designed from a user-centred perspective.

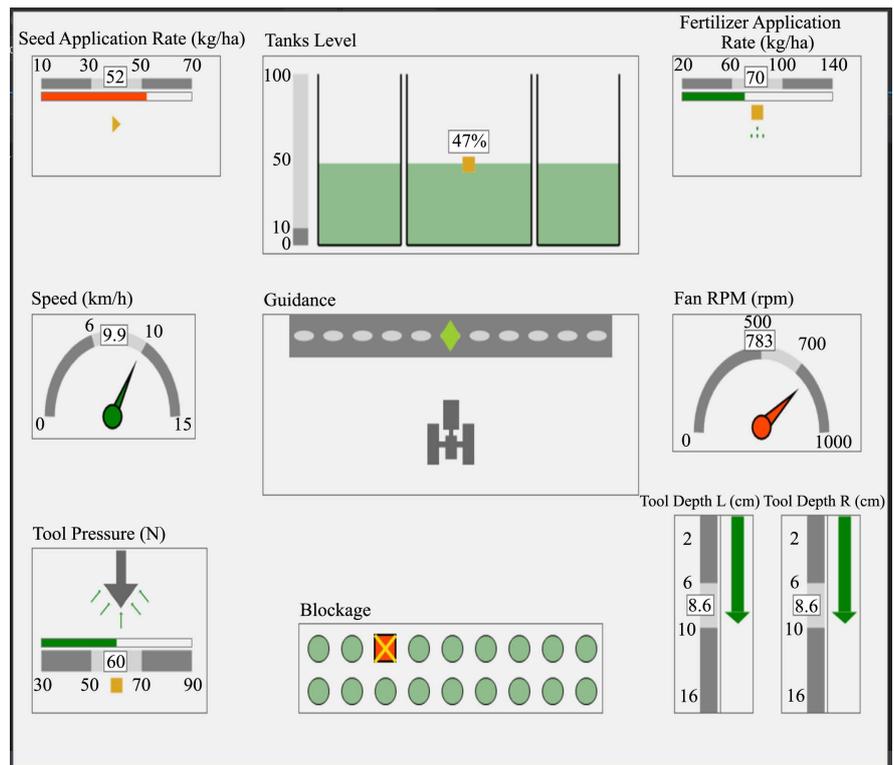


Figure 4. Version 2 of the air seeder interface, referred to as UCD2, designed from a user-centred perspective.

3) Tool Pressure element: Results from [6] indicated that the tool pressure element caused high mental workload—approximately 25% more than baseline conditions. Many study participants indicated difficulty in inferring information from the tool pressure element. Accordingly, changes have been made in the tool pressure element.

4) The coloring of the scales was modified to grayscale from red/green.

5) Marks on the scales were removed.

6) In the seed application rate, an additional animation showing the falling of the seeds was removed.

7) In fertilizer application rate, only one animation representing the falling fertilizer was displayed, instead of the four animations.

8) Another significant change regarding the placement of numeric readings was made; readings were moved from the bottom of each element to the middle of the scale.

3.3. Evaluation of Interfaces

Three interfaces (Old, UCD1, and UCD2) were compared using the metrics of situation awareness (levels 1 - 3), mental workload, and response time in the lab environment. A simulation was developed in the Visual Basic programming language using Microsoft Visual Studio Express 2013. The simulation was constructed in such a way that values on the user interface fluctuated at random intervals while the participants monitored the values on the interface. When the simulation stopped, queries were presented to the participant on the screen to assess the participant's recall of the status of the various parameters and to determine the perceived level of mental workload. This study was completed in the Agricultural Ergonomics Laboratory in the Department of Biosystems Engineering at the University of Manitoba. The experimental protocol received human ethics approval from the University of Manitoba Education/Nursing Research Ethics Board. Participants were also asked to provide subjective feedback.

Thirty individuals (20 male, 10 female) were recruited to participate in the study. For convenience, recruitment was focused on the University of Manitoba campus, with the majority of the participants being University of Manitoba students. Ages ranged from 19 to 52 years with a mean age of 28.4 years. Only 10 participants had prior driving experience and most of the participants (29 out of 30) had no previous experience with agricultural machines. We did not screen participants using any eligibility criteria—all respondents were considered eligible to participate in the study.

The full experiment was divided into two sessions: low-level automation, and high-level automation sessions. Automation level and interface type were considered as independent variables while situation awareness, mental workload and response time were considered as dependent variables. In the low-level automation session, the user was responsible for both observing and correcting the situation. To correct the situation, the user was required to click on the incorrect (out-of-range) parameter on the interface (e.g. see the seed application rate, block-

age, and tool depth in **Figure 3**, and seed application rate, fan rpm and blockage in **Figure 4**). Response time is defined as the amount of time between the appearance of the error and correction of the error by the participant. During the high-level automation session, the user was only responsible for observing the situation; the user was not allowed to correct the situation by clicking on the interface. Half of the participants completed the low-level automation session first, and other half of the participants completed the high-level automation session first. After every simulation, study participants were asked questions related to situation awareness and mental workload (see **Figure 5**). The responses used to assess both situation awareness and mental workload were recorded by the simulation program during the experiment.

As recommended in the Situation Awareness Global Assessment Technique (SAGAT) [5], the questions asked were related to three levels of situation awareness. Responses received under “VALUE”, “TREND”, and “STATUS” headings were correlated to level 1 (perception), level 2 (comprehension) and level 3 (projection) situation awareness, respectively [5]. The degree of situation awareness attained by the participant was inferred based on the proportion of correct responses entered by the participant.

Participants reported their mental workload on a nine-point Integrated Workload Scale [16]. User’s mental workload was inferred based on the user’s selection on the nine-point integrated workload scale varying from “Not Demanding” to “Work Too Demanding”. The numerical equivalent value of mental workload can vary from 1 to 9, where 1 represents “Not Demanding” and 9 represents “Work Too Demanding”.

Subjective feedback regarding the three interfaces was collected after each experimental session using a paper form. Users were required to rate the three interfaces regarding the various criteria mentioned in the questionnaire (**Table 1**).

SA Evaluation

Please respond appropriately for each of the parameters mentioned below:

	VALUE	TREND	STATUS	FUTURE STATUS (After 1 min)
Seed Application Rate (kg/ha)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Tank Levels (%)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Fertilizer Application Rate (kg/ha)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Speed (km/h)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Fan RPM (rpm)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Tool Pressure (N)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Blockage	<input type="radio"/> No <input type="radio"/> Yes	N/A	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable
Tool Depth (cm)	<input type="text"/>	<input type="radio"/> Increasing <input type="radio"/> Decreasing <input type="radio"/> Constant	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable	<input type="radio"/> Acceptable <input type="radio"/> Not Acceptable

Mental Workload

Not Demanding
 Minimal Effort
 Some Spare Time
 Moderate Effort
 Moderate Pressure
 Very Busy
 Extreme Effort
 Struggling to Keep Up
 Work Too Demanding

Next

Figure 5. Screen shot of the form used to collect participant responses after every simulation.

Table 1. Questions asked of the participants at the end of every experimental session to rate the three interfaces.

No.	Criteria
1	Rate the interfaces in terms of perception of the information.
2	Rate the interfaces in terms of recalling (or remembering) of the information.
3	Rate the interfaces in terms of understanding or comprehension of the situation.
4	Rate the interfaces in terms of trend identification.
5	Rate the interfaces in terms of prediction of the future situation.

3.4. Research Hypothesis

As the factors being evaluated involved the respective levels of situation awareness, mental workload and reaction time for different interfaces, the hypotheses should be constructed with these parameters in mind. Consequently, the means of each respective interface form the most reasonable base for comparison, provided that potential random variations across factors for individuals is later acknowledged and accounted for during the analysis.

H₀: Null Hypothesis: On the basis of situation awareness, mental workload and reaction time, neither of UCD1 or UCD2 will be superior in these categories to the original interface. The old model will have a larger or equal average degree of situation awareness, an average reduction in mental workload for the user and an average reduction in reaction time.

$$\mu_{(SA-Old)} \geq \mu_{(SA-UCD1/2)} \text{ and } \mu_{(SA-UCD2/1)}$$

$$\mu_{(MWL-Old)} \leq \mu_{(MWL-UCD1/2)} \text{ and } \mu_{(MWL-UCD2/1)}$$

$$\mu_{(Reaction\ Time-Old)} \leq \mu_{(Reaction\ Time-UCD1/2)} \text{ and } \mu_{(Reaction\ Time-UCD2/1)}$$

H₁: Alternative Hypothesis: Either UCD1 or UCD2 will be superior in terms of situation awareness, mental workload or reaction time to their alternate counterpart and to the original interface design.

$$\mu_{(SA-UCD1/2)} > \mu_{(SA-UCD2/1)} \text{ and } \mu_{(SA-Old)}$$

$$\mu_{(MWL-UCD1/2)} < \mu_{(MWL-UCD2/1)} \text{ and } \mu_{(MWL-Old)}$$

$$\mu_{(Reaction\ Time-UCD1/2)} < \mu_{(Reaction\ Time-UCD2/1)} \text{ and } \mu_{(Reaction\ Time-Old)}$$

As the dataset is heavily predicated on questionnaire responses from individual participants, accounting for the possibility of random variation was imperative. Linear mixed effect models were generated for each output parameter accounting for the input factors and the potential random variation per participant of each factor. Means were examined for each parameter in relation to their dominant factor using the emmeans() function in RMarkdown. Generated random effect models were used to predict values which were subsequently graphically and analytically compared to the originally measured values before being tested for the significance of random and fixed effects.

4. Results and Discussion

4.1. Situation Awareness

Examination of the boxplots for level 1 situation awareness (SA) showed consis-

tent variance across the three interfaces (**Figure 6**). Level 2 SA boxplots indicated a heavy skew towards the higher threshold values, particularly in the case of low automation. The spread of variance was larger for higher automation in both interfaces UCD1 and UCD2 than in the original interface. Variance distribution was somewhat more even in level 3 SA, increasing slightly from the original design to UCD1 and UCD2, but remaining within similar relative ranges.

Examination of the histograms shows a normal distribution for level 1 SA (perception) across all interfaces (**Figure 7**). The distributions for SA levels of comprehension (level 2) and projection (level 3) show a significantly uneven distribution, with values being weighted very heavily towards the upper end of the range of values. In the case of level 2 SA, the majority of samples fall at the highest possible value. The lack of additional degrees of stratification for the data may play a role in this distribution being so skewed, particularly in the case of level 2 where the majority of responses indicated the maximum possible value.

Little difference was found to exist in each level of SA across all interfaces for all levels (**Table 2**), such that no meaningful distinction can be drawn from that factor. While no strong association was derived from the SA parameters, there was a correlation present with the level of automation used during the trial. Significant variation does exist in levels 2 and 3 SA, when considering the factor of automation and how it contributes to the output variables. There is a distinct decrease in comprehension (level 2 SA), and conversely a significant increase in

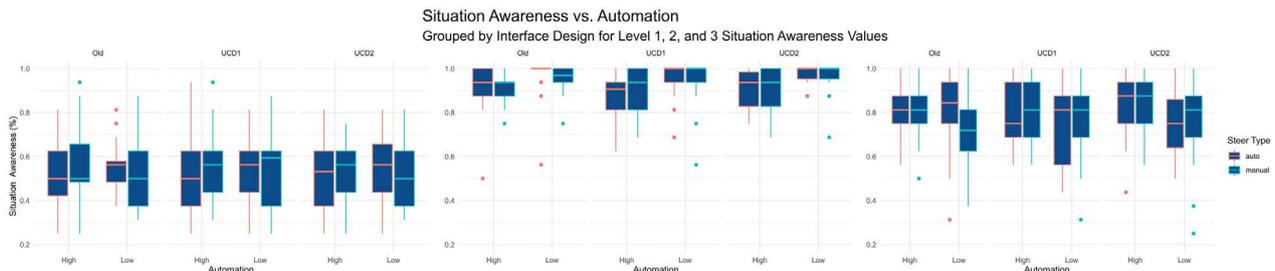


Figure 6. Box-plot of level 1, 2 and 3 situation awareness vs automation for each group, with steer types shown in red and blue.

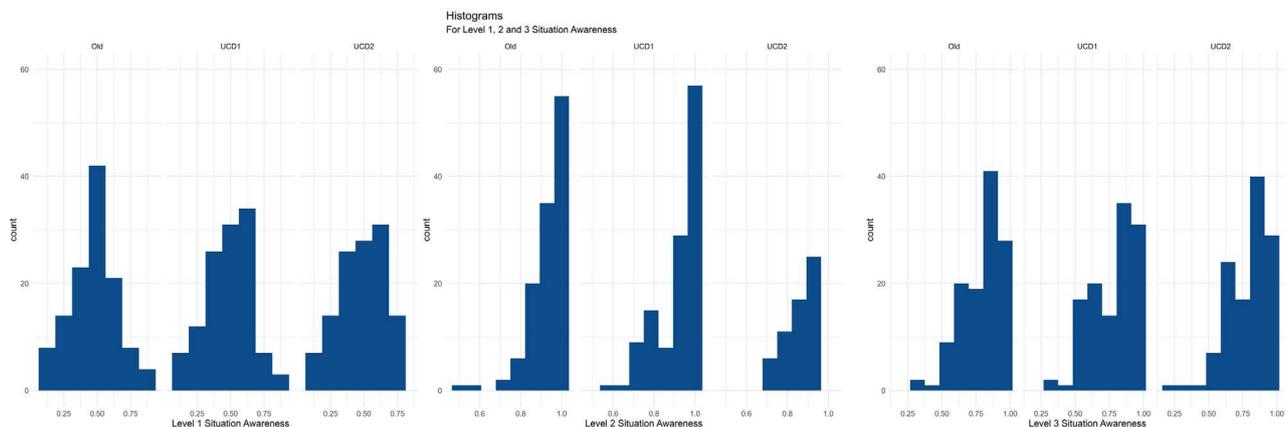


Figure 7. Histograms for level 1, 2 and 3 SA across all interfaces.

Table 2. Comparison of means by level for all interfaces across automation, showing marginal changes across interface, but more pronounced changes when moving by automation for level 2 and 3 SA.

Automation Level	Level 1 SA	Level 2 SA	Level 3 SA	Overall SA
Low Old Interface	0.49	0.96	0.76	0.74
Low UCD1	0.52	0.96	0.76	0.75
Low UCD2	0.51	0.97	0.75	0.75
High Old Interface	0.51	0.92	0.82	0.75
High UCD1	0.50	0.89	0.80	0.73
High UCD2	0.52	0.90	0.83	0.75
Low Overall	0.51	0.96	0.76	0.74
High Overall	0.51	0.90	0.82	0.74
Aggregate	0.51	0.93	0.79	0.74

projection (level 3 SA) when moving from low to high automation levels in the experiment, across all forms of interface. Analysis of the p-values for the generated linear models indicate a reasonably high level of significance for these associations (level 2 p-value = $1.578e-15$, level 3 p-value = 0.01162). Overall, the influence from automation level was seen to have a more significant influence on variations in level 1 SA ($p = 0.004643$) than variations in the interface design ($p = 0.783680$).

Automation proved to be the most significant contributor when an analysis of variance was carried out, with a sum of squares of 0.321 ($p = 1.58e-15$). Prediction of level 2 SA values based solely on automation failed to generate values resembling the originally collected data, with an R^2 of 0.448 (Figure 8). This suggests the presence of random variation significantly influenced the data collected, and the model created does not accurately capture the behavior of the experiment. It should be acknowledged that the predictions generated for all SA scenarios were over a continuous spectrum as opposed to discrete values, but this is unlikely to impact the obtained result and remediation through discrete conversion may distort the outcome. The largest source of random effects for level 2 SA came from per-participant variation ($p = 1.59e-20$).

Prediction of values for level 3 SA appeared somewhat stronger in the case of low automation, although values were still clustered closer together than what might be expected of a more accurate relationship. Analysis of variance testing confirmed that low automation was the primary contributor to a reasonably significant degree, with an estimated decrease of -0.0625 as a result ($p = 0.012$), while the largest component of random variation came from automation variations per participant ($p = 9.64e-11$). R^2 was measured to be 0.538 for the predicted values, shown in Figure 9, indicating the model does not strongly conform to the data.

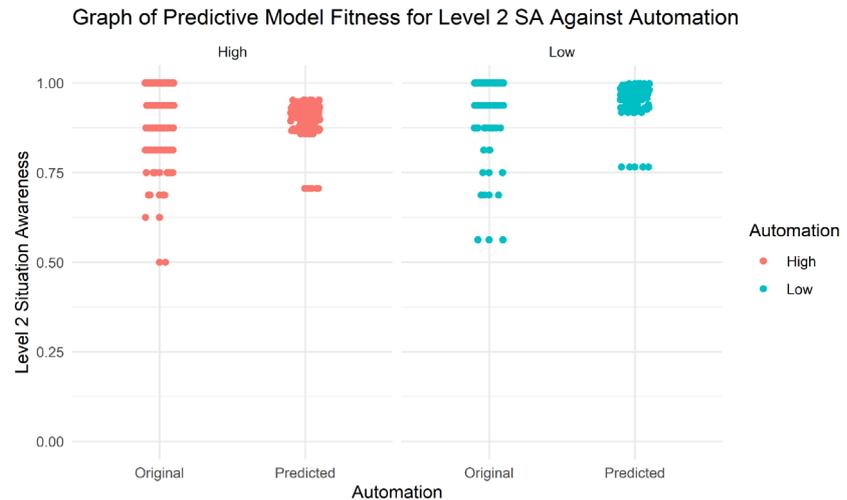


Figure 8. Graph of level 2 situation awareness vs. automation for both experimental and predicted values.

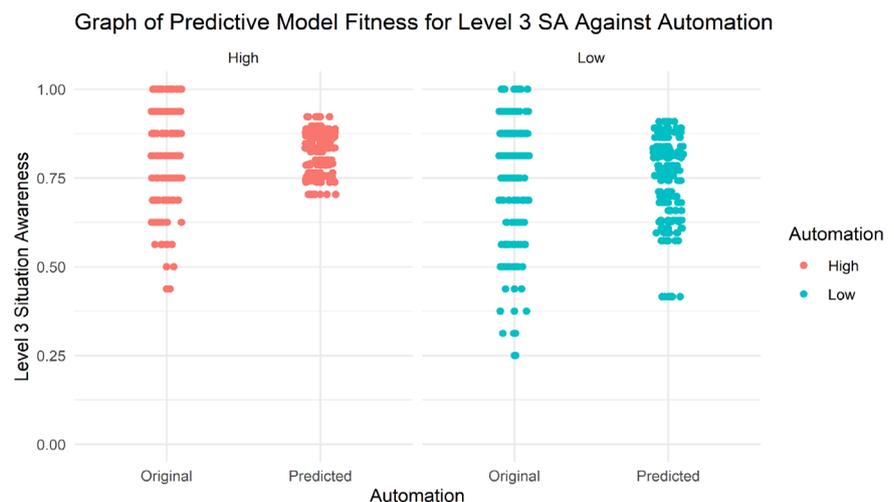


Figure 9. Graph showing predicted level 3 SA vs. automation.

4.2. Mental Workload

The mental workload values for each respective interface were separated and sorted by interface, level of automation and steer type. The resulting measurements were arranged into a comparative boxplot, shown in **Figure 10**.

Examination of the mental workload shows that there is a reduction in mental workload for UCD2 when compared with the original interface and UCD1 ($p = 0.000797$) (**Table 3** & **Table 4**). It is interesting to note that there is an apparent and unexpected slight increase in mean mental workload when comparing UCD1 to the original interface.

Prediction of the values proved to be more reasonable, with a larger distribution of values found compared to the previous predictions for SA, as shown in **Figure 11**. R^2 was measured to be 0.823, indicating a reasonably significant approximation for the data.

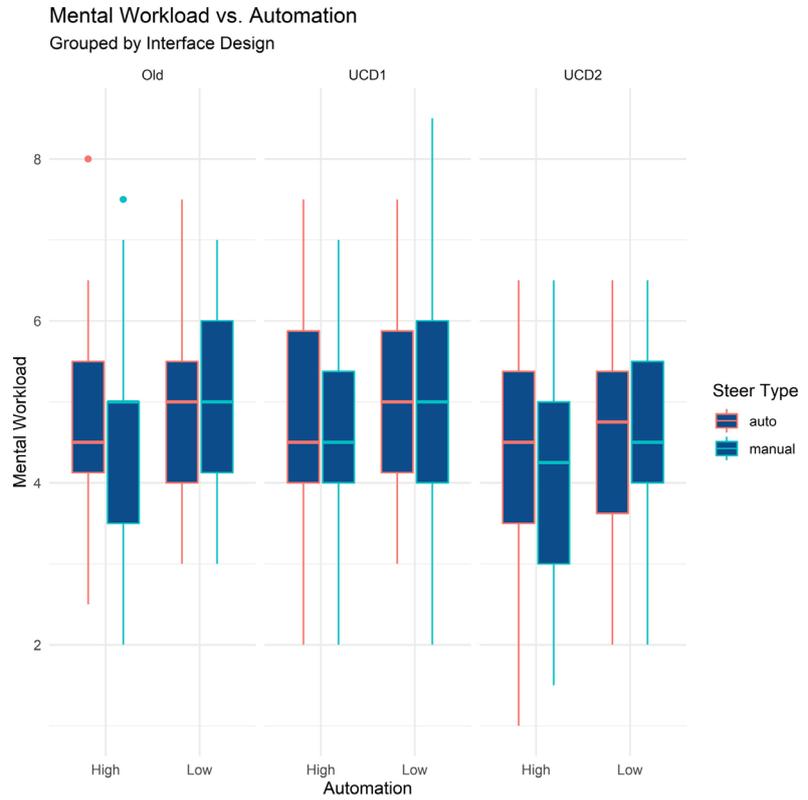


Figure 10. Box-plot of mental workload vs interface design for each group, with steer types shown in red and blue.

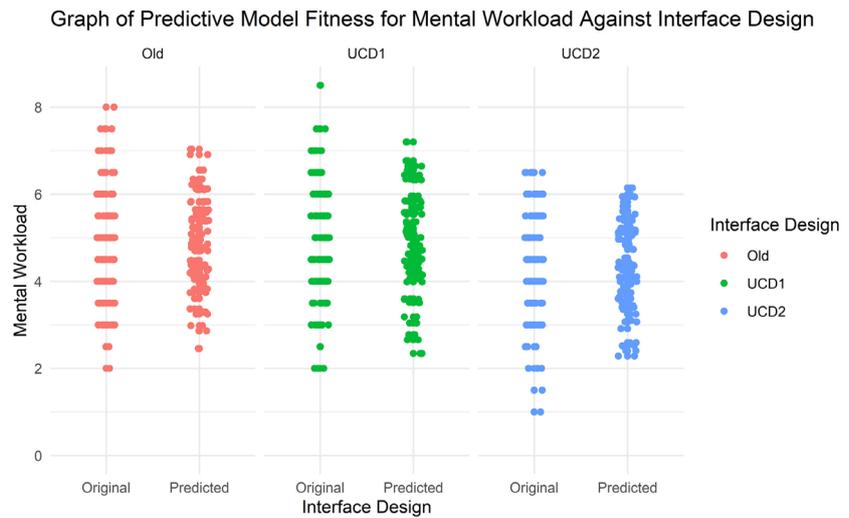


Figure 11. Graph comparing original and predicted mental workload vs. interface design.

Table 3. Mean mental workload vs. interface.

Interface	Mean Mental Workload	SE	df	lower.CL	upper.CL
Old	4.86	0.1870	29	4.48	5.24
UCD1	4.92	0.2011	29	4.51	5.33
UCD2	4.44	0.1793	29	4.07	4.80

Table 4. Summary of mental workload random analysis of variance model showing estimated change and p-values.

Element	Estimate	Standard Error	df	p
Intercept	4.858	0.182	29	2e-16
UCD1	0.063	0.103	29	0.549
UCD2	-0.421	0.113	29	0.000841

4.3. Response Time

A significant decrease in mean reaction time is observed across interfaces as shown (Table 5). Reaction times are only listed for low automation, as that was the only experimental scenario where a response was required from participants. Analysis of the model indicated this difference was distinctively tied to the respective interface to a significant degree ($p = 9.232e-08$) (Table 6). Fixed effects anova testing for mental workload gave a sum of squares value of 6.72 for interface design, indicating significant contribution to the variance ($p = 0.000797$), while anova testing for random effects indicated both interface design and automation to be significant contributors ($p = 4.49e-03$ and $4.93e-24$, respectively).

Attempting to predict the reaction time based on the generated linear model was successful, as the predicted values appeared to closely resemble the originally measured reaction times (Figure 12). R^2 was measured as 0.790, indicating the predicted model is reasonable. Anova tests confirmed the interface design to be the primary source of variation for reaction times, with a sum of squares value of 5.04 ($p = 9.23e-8$). Random effects testing indicated the dominant contribution being as a result of interface design for individual participants, to a significant degree ($p = 0.0101$).

4.4. Subjective Feedback

At the end of each session (high-automation or low-automation), subjective feedback was collected. Participants were asked to rate the three interfaces as best, average or worst based on their experience during the experiment and were required to evaluate the three interfaces using five questions related to the perception of the information, recall, comprehension, trend, and prediction of the future state. Data were evaluated by performing Ordinal Logistics Regression of Interface-Ranking versus Interface-Design in Minitab (Table 7) to understand and quantify the participant's responses to the three interfaces. Interface-Ranking was a categorical response variable with three outcomes having an order (best, average and worst). Interface-Design was a categorical predictor variable having three levels (A, B, C). Level A represents the Old (baseline) design, level B represents UCD1, and level C represents UCD2.

Results from the Ordinal Logistics Regression (Table 7) indicate that the relationship between predictor and response variables is significant ($G = 244.062$,

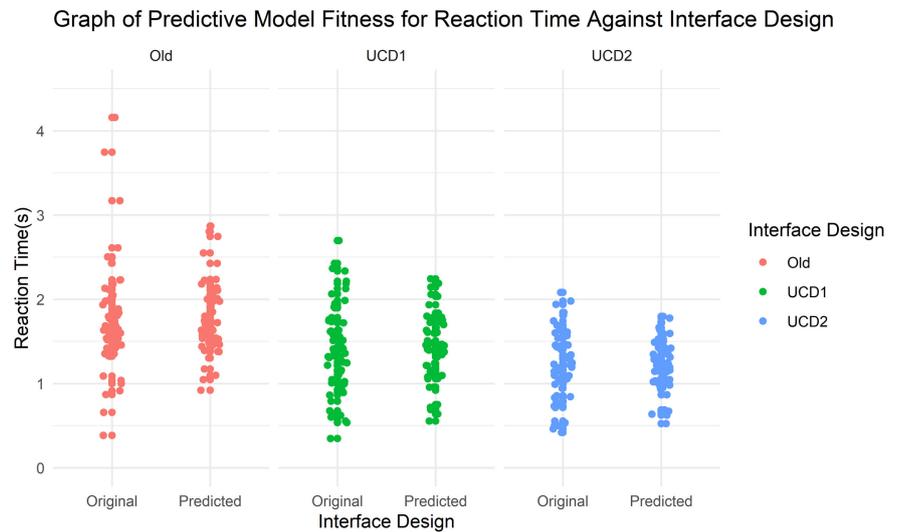


Figure 12. Reaction time vs. interface design for the original experiment and generated prediction values.

Table 5. Means for reaction time relative to the interface used.

Interface	Mean Reaction Time (s)	SE	df	lower.CL	upper.CL
Old	1.73	0.0913	29	1.55	1.92
UCD1	1.40	0.0898	29	1.21	1.58
UCD2	1.19	0.0707	29	1.04	1.33

Table 6. Summary of reaction time model showing estimated decrease and p-values.

Element	Estimate	Standard Error	df	p
Intercept	1.7325	0.0896	29	2e-16
UCD1	-0.334	0.0946	29	0.0014
UCD2	-0.545	0.0811	29	1.2e-07

Table 7. Ordinal Logistics Regression output of Interface-Ranking (best, average, worst, with a total count of 900) as dependent (response) variable versus Interface-Design (A-Old, B-UCD1 & C-UCD2) as independent (predictor) variable.

Predictor variable— Interface-Design (3 levels—A, B, C)	Coefficient	SE of Coefficient	Z-value	p-Value	Odds Ratio	95% CI	
						Lower	Upper
B (UCD1)	1.0425	0.1587	6.57	0.000	2.84	2.08	3.87
C (UCD2)	2.5589	0.1751	14.61	0.000	12.92	9.17	18.21

*Results of the test that all slopes are zero: G = 244.062, df = 2, p-Value = 0.000

Goodness-of-Fit Tests

Method	Chi-Square	df	p-Value
Pearson	2.0687	2	0.355
Deviance	2.0497	2	0.359

$p = 0.000$). From the Goodness-of-Fit Tests, we have observed p -values as 0.355 and 0.359 based on the Pearson and Deviance method. This high p -value does not provide evidence that the model is inadequate. Most importantly, it has been observed that both design B (UCD1) and C (UCD2) are significantly different from Design A (Old). For design B in comparison to A, we found $p = 0.000$, Odds ratio = 2.844 and Coefficient = 1.04251. A positive coefficient and greater than 1 odds ratio indicate that B has been rated significantly higher than A. Odds of being rated higher for B are 2.8 more than that of A. Similarly, odds for being ranked high for C are 12.9 times more than A. The high odds ratio of C indicates that subjects have placed very high confidence in the UCD2 interface.

5. Conclusions

Based on the analysis performed on the data obtained, UCD2 was found to be superior to both UCD1 and the original interface across the assessed parameters. Situation awareness changes were marginal between interfaces, and were more closely correlated to the type of automation being used in the experiment than the interface design used. While the means for perception (level 1 SA), comprehension (level 2 SA) and projection (level 3 SA) were not found to have any significant distinction across interfaces, UCD2 was shown to have lower average levels of both participant mental workload and response time than either UCD1 or the original interface. The null hypothesis can be said to have been rejected for both mental workload and reaction time. The decrease in mental workload and response time, coupled with the relative parity of all interfaces for situation awareness provide sufficient grounds to state that the second alternate interface (*i.e.* UCD2) shows general improvement when situation awareness, mental workload and reaction time are the focus of assessment.

There are two important limitations to be noted with this research. First, the research was conducted on a user interface for an agricultural air seeder. Although the results might be generalizable to other agricultural machines, we do not have any experimental evidence to confirm that these results can be applied to the user interface for other types of agricultural machines. Second, the research results are based on the perceptions of participants with limited experience using agricultural machines. It is unknown whether the results would differ if participants would have been recruited from the population of experienced air seeder users.

Acknowledgements

The authors would like to acknowledge funding from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Liu, Y. (1997) Software-User Interface Design. In: Salvendy, G., Ed., *Handbook of Human Factors and Ergonomics*, 2nd Edition, John Wiley & Sons, New York, 1689-1724.
- [2] Norman, D.A. (1983) Design Rules Based on Analyses of Human Error. *Communications of the Association for Computing Machinery*, **26**, 254-258.
<https://doi.org/10.1145/2163.358092>
- [3] Sanchez, J. and Duncan, J.R. (2009) Operator-Automation Interaction in Agricultural Vehicles. *Ergonomics in Design: The Quarterly of Human Factors Applications*, **17**, 14-19. <https://doi.org/10.1518/106480409X415161>
- [4] Vicente, K.J., Rasmussen, J. and Member, S. (1992) Ecological Interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics*, **22**, 589-606. <https://doi.org/10.1109/21.156574>
- [5] Endsley, M.R., Bolte, B. and Jones, D.G. (2003) Designing for Situation Awareness: An Approach to User-Centered Design. CRC Press, Boca Raton.
<https://doi.org/10.1201/9780203485088>
- [6] Rakhra, A.K. and Mann, D.D. (2018) Design and Evaluation of Individual Elements of the Interface for an Agricultural Machine. *Journal of Agricultural Safety and Health*, **24**, 27-42. <https://doi.org/10.13031/jash.12410>
- [7] Jeffries, R., Miller, J.R., Wharton, C. and Uyeda, K. (1991) User Interface Evaluation in the Real World: A Comparison of Four Techniques. *Proceedings of the SIGCHI Conference on Human Factors in Computing System*, New Orleans, April 1991, 119-124. <https://doi.org/10.1145/108844.108862>
- [8] Nielsen, J. (1994) Usability Inspection Methods. *Conference Companion on Human Factors in Computing Systems*, Boston, April 1994, 413-414.
<https://doi.org/10.1145/259963.260531>
- [9] Endsley, M.R. (2015) Situation Awareness: Operationally Necessary and Scientifically Grounded. *Cognition, Technology & Work*, **17**, 163-167.
<https://doi.org/10.1007/s10111-015-0323-5>
- [10] Endsley, M.R. (1988) Design and Evaluation for Situation Awareness Enhancement. *Proceedings of the Human Factors Society Annual Meeting*, **32**, 97-101.
<https://doi.org/10.1177/154193128803200221>
- [11] Endsley, M.R. (1996) Automation and situation awareness. In: Parasuraman, R. and Mouloua, M., Eds., *Automation and Human Performance: Theory and Applications*, Lawrence Erlbaum, Mahwah, 163-181.
- [12] Byrne, E.A. and Parasuraman, R. (1996) Psychophysiology and Adaptive Automation. *Biological Psychology*, **42**, 249-268.
<http://www.ncbi.nlm.nih.gov/pubmed/8652747>
[https://doi.org/10.1016/0301-0511\(95\)05161-9](https://doi.org/10.1016/0301-0511(95)05161-9)
- [13] Taylor, H., Lee, B., Jhingory, J., Drayer, G.E. and Howard, A.M. (2010) Development and Evaluation of User Interfaces for Situation Observability in Life Support Systems. American Institute of Aeronautics and Astronautics, Reston, 1-8.
- [14] Megaw, T. (2005) The Definition and Measurement of Mental Workload. In: Wilson, J.R. and Corlett, N., Eds., *Evaluation of Human Work*, 3rd Edition, Taylor & Francis Group, Abingdon-on-Thames, 525-551.
- [15] Di Stasi, L.L., Adoracion, A. and Canas J.J. (2013) Evaluating Mental Workload while Interacting with Computer-Generated Artificial Environments. *Entertain-*

- ment Computing, 4, 63-69. <https://doi.org/10.1016/j.entcom.2011.03.005>
- [16] Pickup, L., Wilson, J.R., Norris, B.J., Mitchell, L. and Morrisroe, G. (2005) The Integrated Workload Scale (IWS): A New Self-Report Tool to Assess Railway Signaller Workload. *Applied Ergonomics*, 36, 681-693. <https://doi.org/10.1016/j.apergo.2005.05.004>
- [17] Muckler, F. and Seven, S.A. (1992) Selecting Performance Measures: "Objective" versus "Subjective" Measurement. *Human Factors*, 34, 441-455. <https://doi.org/10.1177/001872089203400406>
- [18] Durso, F.T., Hackworth, C.A., Truitt, T.R., Crutchfield, J., Nikolic, D. and Manning, C.A. (1998) Situation Awareness As a Predictor of Performance in En Route Air Traffic Controllers. *Air Traffic Control Quarterly*, 6, 1-20. <https://doi.org/10.2514/atcq.6.1.1>
- [19] Endsley, M. and Jones, D. (2016) Designing for Situation Awareness: An Approach to User-Centered Design. 2nd Edition, CRC Press, Boca Raton. <https://books.google.com/books?hl=en&lr=&id=eRPBkapAsggC&oi=fnd&pg=PP1&ots=dJIDDbVII&sig=JUoCXkSF7MIGhT-0uPgSnIwnN4o>
- [20] Hart, S.G. and Staveland, L.E. (1988) Development of NASA-TLX: Results of Empirical and Theoretical Research. *Advances in Psychology*, 52, 139-183. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- [21] Koch, S.H., Weir, C., Westenskow, D., Gondan, M., Agutter, J., Haar, M., Staggers, N., et al. (2013) Evaluation of the Effect of Information Integration in Displays for ICU Nurses on Situation Awareness and Task Completion Time: A Prospective Randomized Controlled Study. *International Journal of Medical Informatics*, 82, 665-675. <https://doi.org/10.1016/j.ijmedinf.2012.10.002>
- [22] Squire, P.N. and Parasuraman, R. (2010) Effects of Automation and Task Load on Task Switching during Human Supervision of Multiple Semi-Autonomous Robots in a Dynamic Environment. *Ergonomics*, 53, 951-961. <https://doi.org/10.1080/00140139.2010.489969>
- [23] Rakhra, A.K. and Mann, D.D. (2014) Design Guidelines Review and Conceptual Design of an User-Centered Information Display for Mobile Agricultural Machines. *American Society of Agricultural and Biological Engineers Annual International Meeting*, Montreal, 13-16 July 2014, 2917-2932. <https://elibrary.asabe.org/abstract.asp?aid=44456>