International Journal of Industrial Ergonomics 83 (2021) 103108

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: http://www.elsevier.com/locate/ergon

Context enabled decision aids to support the situation awareness and performance of risk engineers carrying out loss prevention surveys

Sruthy Agnisarman[a,*], Kapil Chalil Madathil[a,b], Jeffrey Bertrand[b], Kalyan Piratla[a], Dhaval Gajjar[c], Anand Gramopadhye[b]

[a] Glenn Department of Civil Engineering, Clemson University, United States
[b] Department of Industrial Engineering, Clemson University, United States
[c] Department of Construction Science and Management, Clemson University, United States

ARTICLE INFO

Keywords:
Risk inspection
Situation awareness
Decision aids
Visualizations
Performance

ABSTRACT

Routine inspection by insurance companies at their clients' facility, also known as loss prevention survey, help identify the best strategies to minimize damages when there is a high-speed wind event. More specifically, wind vulnerabilities associated with a building are evaluated using a process known as windstorm risk inspection. This routine inspection helps clients reduce the extent of damages caused by high-speed wind events including hurricane and tornado. Risk engineers make use of their subjective and analytical deduction skills to successfully carry out the inspection tasks. In this research the researchers investigated the effect of context-based visualization strategies on situation awareness and their understanding of the situation. The study examined how different types of information contribute towards the three levels of situation awareness. Following a between-subjects study design, 65 participants completed the study. Each session lasted 90–120 min. A checklist based and predictive display-based decision aids were tested and found to be effective in supporting the situation awareness requirements as well as performance of risk engineers. However, the predictive display only helped with certain tasks such as understanding the interaction among different components on the rooftop. For remaining tasks such as perceiving obvious issues like membrane tear, clogged drains and vegetation growth, checklist alone was sufficient. This study helped the understanding of the advantages and disadvantages of the decision aids tested. More specifically, these decision aids can improve the mental model of novice risk engineers. Additionally, this study provided insights that could help design training materials for infrastructure inspectors.

1. Introduction

Over the past ten years, on an average, United States experienced 170 wind-related fatalities (NWS Analyze Forecast and Support Office, 2018). Such fatalities as hurricanes, tornados and thunderstorms affect people and society as well as the economy (Tokgoz, 2012). The effect of these disasters range from direct damages such as physical destruction of assets and capital to indirect damages (Khazai et al., 2013). More specifically, property damage costs billions of dollars in losses (Fernández, 2001). In 2017 alone, such weather events resulted in a cumulative cost of $306.2 billion (Hurricane Costs, 2019). To limit the extent of these damages, wind vulnerability assessments are conducted to identify and mitigate damage and to minimize disruption (Smith Thomas, 2011). Specifically, insurance companies conduct routine inspection tasks or loss prevention surveys in their clients' facility to reduce the frequency and severity of such damages (Schlesinger and Venezian, 1986). Though this windstorm loss prevention survey or risk inspection can benefit both the clients and insurance company, the validity of the inspection findings and conclusions depend on the individual capabilities of the personnel carrying out the survey (Agnisarman et al., 2018; Agnisarman et al., 2019a,b).

A previous qualitative study investigating the sensemaking process and situation awareness of loss prevention inspection professionals identified the lack of a standardized survey protocol as one reason for the disparity in their findings (Agnisarman et al., 2018). Furthermore, individual differences in the ability and experience level of these engineers contribute to this subjectivity (Agnisarman et al., 2018), with the latter being one of the most important factors contributing to the accuracy of the inspection report. Experienced engineers are better equipped to comprehend the elements in the environment to assess the

* Corresponding author.
E-mail address: sorozhi@clemson.edu (S. Agnisarman).

https://doi.org/10.1016/j.ergon.2021.103108
Received 12 October 2020; Received in revised form 14 January 2021; Accepted 11 February 2021
Available online 1 April 2021

current state and predict potential damages and structural changes to the building than their novice counterparts who, due to their lack of experience, may find it challenging to perceive and comprehend information to better understand the infrastructure system (Agnisarman et al., 2018).

Automation-assisted technologies and Artificial Intelligence (AI) based decision aids have been used by researchers and practitioners to improve the accuracy of the infrastructure inspection process (Agnisarman et al., 2019a,b). Such smart decision aids can facilitate decision making by reducing the mental demand on the risk engineers by assisting them with the preliminary data analysis and cue the engineers to look for relevant information when completing the risk inspection task. However, such technologies are not without limitations. Though, these technologies can assist in conducting infrastructure inspection, the engineers' ability to interpret and synthesize the data is important (Agnisarman et al., 2018), especially since the performance of the person operating such systems is affected by factors such as drop in vigilance, automation complacency and loss of situation awareness (Endsley, 1999; Endsley and Kiris, 1995).

In the risk inspection domain, automated decision aids are not expected to completely automate the risk inspection process. Instead, it can augment the risk engineers' decision making with the help of predictive algorithms, which generally outperform expert judgement as risk engineers' ability to predict the potential damages to the infrastructure system is limited. However, human involvement is required to make decisions about unusual situations that are not accurately modeled using historical data (Guszcza, 2018). Such situations require intelligent systems to generate anchor points for the experts to augment human decision making (Guszcza, 2018). To support this effort, there is a need to develop algorithms meeting contextual needs. The human-centered design should highlight the needs and requirements of the specific context under consideration to facilitate the optimal use of AI algorithms, emphasizing the importance of considering situation awareness in designing decision aids based on AI for risk engineers (Agnisarman et al., 2018).

1.1. Situation awareness (SA)

Situation awareness (SA) is the perception of the elements/cues in the environment (Level 1), comprehension of the current situation of the elements (Level 2) and the projection of the status of the elements and environment into the future (Level 3) (Endsley, 1995). Past studies suggested that the systems that could throw operators out of the loop can affect any of these levels (Endsley and Kiris, 1995; Khasawneh et al., 2019). The application of this 3-level situation awareness framework (Endsley, 1995) can be seen in many domains such as aviation, aircraft maintenance and surgery in an effort to improve operator performance (Endsley and Robertson, 2000; Fioratou et al., 2010; Jones and Endsley, 1996). However, our systematic literature search (Agnisarman et al., 2019a,b) did not retrieve any articles in the domain of loss prevention inspection or building inspection focusing on the SA requirements of inspectors/engineers. To address this lack of research, this study focuses on designing visual decision aids to improve the situation awareness of infrastructure inspectors.

1.2. Relevance of SA in infrastructure risk inspection

Windstorm risk inspection process involves identifying wind vulnerabilities associated with a building to reduce the extent of damage in the event of a hurricane or windstorm. Though the 3-level SA framework has been primarily used to identify SA requirements in dynamic systems, this concept is relevant to the inspection and maintenance domain as well (Endsley and Robertson, 2000). Though the infrastructure inspection process does not involve a dynamic environment, risk engineers need to assess the current state and extrapolate it to the future. However, there are a number of unknown factors such as wind speed and direction, the overall condition of the infrastructure, and other interdependencies such as the distance between potential windborne projectiles known as missiles and infrastructure system and locations of other objects that make predicting the future state of the infrastructure a challenging task. More importantly, the dynamic events and behavior patterns of the components of an infrastructure following a higher category hurricane pose a real challenge for the risk engineers.

The Level 1 SA requirements of risk inspection involve perceiving cues including, but not limited to, the type of roof, type of rooftop equipment, age of the roof, surface roughness and missile exposure. In Level 2 SA, the engineers comprehend the information perceived to evaluate the current condition of the infrastructure system. During this process, engineers may face a number of challenges, the most important one being the lack of information available. They then predict potential damages and failures based on its current state as well as other environmental conditions. The sensemaking process of infrastructure risk engineers during this process has been discussed in detail in another article (Agnisarman et al., 2018). While automated systems are used to support the windstorm risk inspection process, there is a need to understand how engineers' SA is impacted. In this research we will develop decision aids for information visualization to support the process of synthesizing cues in the environment to achieve adequate level of SA for risk inspectors.

1.3. Risk assessment

There are two primary methods currently being used for assessing hurricane structural damage: the subjective method and the analytical method (Mehta et al., 1981). The subjective method involves windstorm engineers going to a site to obtain information about the roofing system, envelope, connections, drawings and specifications, while the analytical method is based on the principles of structural mechanics and an understanding of material properties to predict wind speed and potential damages (Mehta et al., 1981). The subjective windstorm visual inspection method detailed in Agnisarman et al. (2018) formed the basis for identifying the information needed in the visualizations. In addition, analytical hurricane damage prediction models were also explored to identify the elements that need to be included in the contextual visualization.

Risk involves both the probability of risk realization and the effect of threat realization (Väisänen et al., 2018). Though human visual perception is capable of detecting anomalies and patterns, the ability of the risk engineers to predict the future state of an infrastructure is limited. Information visualization uses external aids such as computers to strengthen the cognitive capabilities of users/decision makers (Kapler and Wright, 2005). Risk visualization, which involves visualizing potential risks to enhance cognition to facilitate decision making, will potentially augment the inspector's cognition and enhance his/her situation awareness. However, presenting the specific data needed to meet the demands of the end user can be challenging. This requires the identification of the needs and visualization requirements of this specific user group (Kasireddy et al., 2015). So, it is important to first evaluate the visualization requirements of the windstorm inspectors to design decision aids to meet their needs.

1.4. Related work

The design of technologies to support SA has been investigated extensively in aviation and healthcare. Additionally, the SA theory proposed by Endsley (1995) was applied to evaluate the effect of various types of display strategies, specifically tactical vs. waterfall, for submarine track management in a simulated environment (Loft et al., 2015). Loft et al. (2015) studied the relationship between various SA measures such as Situation Present Assessment Method (SPAM) and Situation Awareness Global Assessment Technique (SAGAT) and performance, identifying a correlation among them. Another research reported a reduction in self-reported SA as a result of an increased amount of task relevant information, meaning increased task-relevant information, despite being accurate, might not help with decision making (Marusich et al., 2016). Researchers have also studied the effect of the nature of information presented on the SA of mobile crane operators; they identified a general trend in improvement in operator performance and SA with the use of a virtually reconstructed visualization of a lift scene (assistance system) over traditional systems (Fang et al., 2018). In addition to mobile crane monitoring and operations, studies have also been conducted investigating the effect of situation-augmented displays for UAV monitoring (Lu et al., 2013), the findings suggesting that situation-augmented displays may provide sufficient situation awareness to improve user performance (Lu et al., 2013).

The application of the SA framework to investigate various information presentation strategies can be seen in defense research as well. A recent study investigated the effect of presentation modality, auditory vs. visual and message presentation rate on the SA and the cognitive load of soldiers (Hollands et al., 2019). The findings revealed that visual messages and higher message presentation rate resulted in higher cognitive load and reduced SA. Similar studies have been conducted in the healthcare domain as well, for example, a study investigating the effect of head-worn display (HWD) providing continuous patient information on the SA of nursing students while responding to patient alarm. The researchers observed that the participants' responses to SA questions were more accurate when using HWD compared to the alarm only condition (Pascale et al., 2019). Researchers have also investigated the effect of other decision aids such as a checklist on SA. For example, one such study investigated if the use of a checklist improves SA during physician handoffs in a pediatric emergency department. Participants in this study reported an improvement in their SA following the use of a standardized checklist (Mullan et al., 2015).

However, none of the previous research developed decision aids for supporting the SA, performance and workload of infrastructure inspectors. More specifically, to date, no studies have been conducted with windstorm risk engineers. While researchers have investigated the potential of using Augmented Reality (AR)-based systems for flood visualization (Haynes et al., 2018), no studies have looked at the situation awareness requirements and performance of inspectors. In the study reported here, the researchers investigated how various visualization techniques could be designed to enhance and support the SA of risk engineers. The checklist and predictive display based decision aids used here were designed to meet the requirements identified from an exploratory research that followed an interview based approach to learn the sensemaking process and SA requirements of windstorm engineers. In addition, the principles proposed by Endsley for designing for situation awareness were also incorporated in the decision aids (Endsley, 2016). More specifically, this study designed and tested checklist-based and predictive display-based decision aids. While risk engineers currently use a high-level checklist, it is not standardized. The checklist used in this study was reviewed by a subject matter expert and the predictive display used in this research is a novel idea which has not yet been used for this application. To investigate the effectiveness of these decision aids, the following research questions were investigated:

RQ1. How is the SA of the participants affected by the context-based decision aids developed?

RQ2. How is the performance of the participants affected by the context-based decision aids developed?

RQ3. How does the nature of context-based visual decisions aids affect the cognitive load imposed on the participants?

These hypotheses tested in this research were:

H1. Participants in the predictive display condition will have higher SA compared to participants in the checklist condition and control condition and participants in the checklist condition will have higher SA compared to participants in the control condition.

H2. Participants in the predictive display condition will have higher performance score compared to participants in the checklist condition and control condition and participants in the checklist condition will perform better compared to participants in the control condition.

H3. Participants in the predictive display condition will have lower cognitive load compared to participants in the checklist condition and control condition and participants in the checklist condition will have lower cognitive load compared to participants in the control condition.

2. Method

2.1. Study sample

For this research study, undergraduate civil engineering or construction science and managements students in their 3rd or 4th year (Junior or Senior year) or graduate students with the same background were recruited. These students have taken roof inspection and construction management related course and have basic background in this topic. Sixty-five participants, (min = 20, max = 41 years old, M = 23.35, SD = 3.37) were recruited for this study. Table 1 illustrates more details about the study participants.

2.2. Apparatus

This study used a Dell desktop computer with an Intel(R) Xeon(R) CPU E5-1620 v4 processor and a Quadro FX 5800 GPU to run the simulations of a windstorm risk survey. An LG ultrawide monitor with a diagonal dimension of 38.8 inches was used as the display. The simulations were developed using the Unity game engine (Unity, 2005). A laptop computer was used to administer the questionnaires prior to, during and after the study through Qualtrics Research Suite (Qualtrics, 2005). Fig. 1 illustrates the lab setup used in this study.

2.3. Simulation

The participants completed this study in a simulated environment. An academic building located within a 10-miles radius of the Atlantic Coast was used as the simulated scenario. The exposure category used in this study was Category C with generally open terrain with limited obstructions (Windexpo, 2019). This location and exposure category were chosen to simulate moderate wind exposure and related damages. The location has only two buildings. The front yard of the main academic building had a pond and the backyard had a lake. The building had a mechanically fastened thermoplastic olefin (TPO) roof. This roof type was used as this is one of the commonly inspected roof systems by

Table 1
Demographic characteristics of the participants.

Variable (N = 65)	N	%
Gender		
Female	13	20
Male	52	80
Race		
White	39	60
Asian	18	28
Black/African American	5	8
Other	3	4
Major		
Civil Engineering	55	85
Construction Science	10	15
Degree Pursuing		
Undergraduate	37	57
Graduate	17	26
Doctorate	11	17

Fig. 1. Experimental setup.

windstorm risk engineers. The building had a number of pieces of rooftop equipment ranging from antennas to duct work. The rooftop also had certain issues such as ponding, missing fasteners, a flashing issue, a membrane fissure and clogged drains. Fig. 2 illustrates four example images of the simulation used in this study.

2.4. Visualization stimuli development

Contextual visual aids can be developed following SA design principles (Endsley, 2016) to enhance the SA of novice as well as experienced users. The requirements supporting SA in this domain were identified from a previous exploratory research investigating the needs and requirements of risk engineers (Agnisarman et al., 2018). This study identified a few factors that affect the decision making process of windstorm risk engineers. Their experience level, site conditions, wind speed in the event of a hurricane and building code requirements are a few examples of the factors identified. All these factors impact the mental models of the engineers and, hence, their perception of information. Furthermore, novice engineers fail to consider alternate frames to analyze various reasons for possible damages when making sense of cues available in the environment. This will affect their Level 1 and Level 2 SA. Additionally, in complex situations they may have a difficult time predicting potential damages and future state of the building. This will affect their Level 3 SA. The decision aids developed and evaluated in this study are expected to support their Level 1, Level 2 and Level 3 SA.

The context based visual aids developed here were expected to assist the risk professionals when completing the inspection tasks.

2.5. Scenarios and tasks completed

To develop the study scenarios, the various components of a building as defined by Unanwa (1997): the roof covering, the roof sheathing and roof frame, the building envelope, the building occupancy and the structural system were considered. These building components were then used to develop the simulation for this study. The tasks that needed to be completed in the risk assessment of the building were designed based on the insights gained from the previous exploratory research (Agnisarman et al., 2018). The participants completed the following tasks validated by the subject matter expert:

- Investigating the surroundings to understand missile and flood exposure
- Observing roof underdeck, roof condition, flashing, roof deck, and attachments and obtaining building dimensions
- Investigating rooftop equipment to verify the adequacy of the securing method
- Investigating building envelope (windows, dock doors, External Insulation and Finishing System (EIFS))

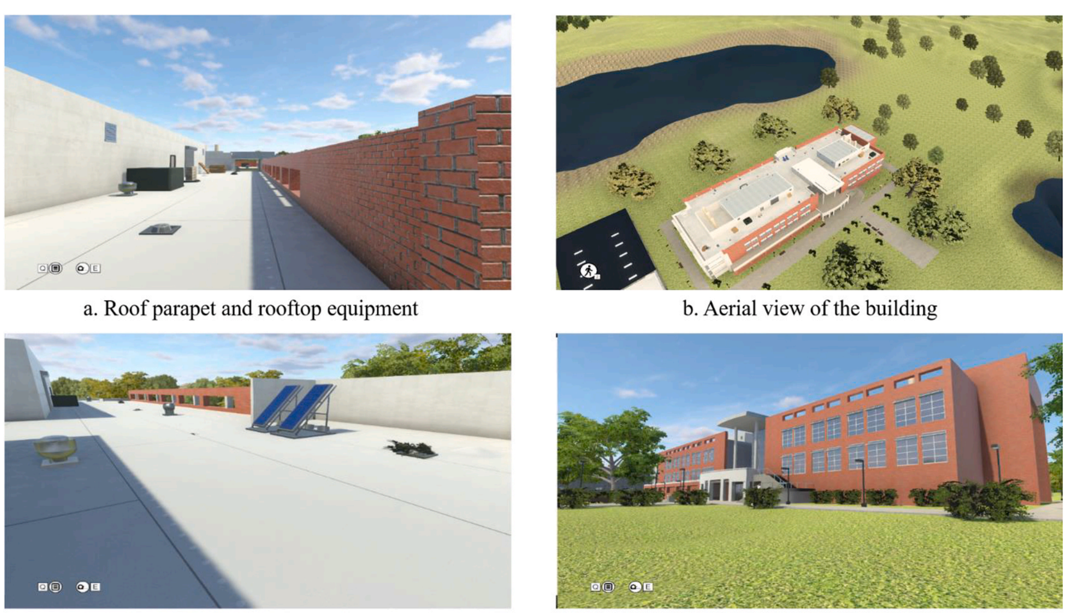
2.6. Experimental design

2.6.1. Independent variables

This following experimental conditions were tested:

Type of context-based visual decision aids presented (3 levels): The context-based visual aids supporting SA functioned as the between-subjects variable in the simulation at three levels:

No visual aid/control condition. In this condition, the participants were not provided any visual decision aids. They had to walk through the simulation and perform various inspection activities. They were



a. Roof parapet and rooftop equipment
b. Aerial view of the building
c. Rooftop equipment
d. Perspective view of the building

Fig. 2. Four screenshots from the simulation.

given a sheet of paper listing the tasks they needed to complete.

Checklist for aiding users and walking them through the inspection steps. This checklist-based visual aid used here provides participants context-based cues to help perceive relevant details about the infrastructure system and comprehend them to make sense of the information (Fig. 3 illustrates an example of this checklist-based display).

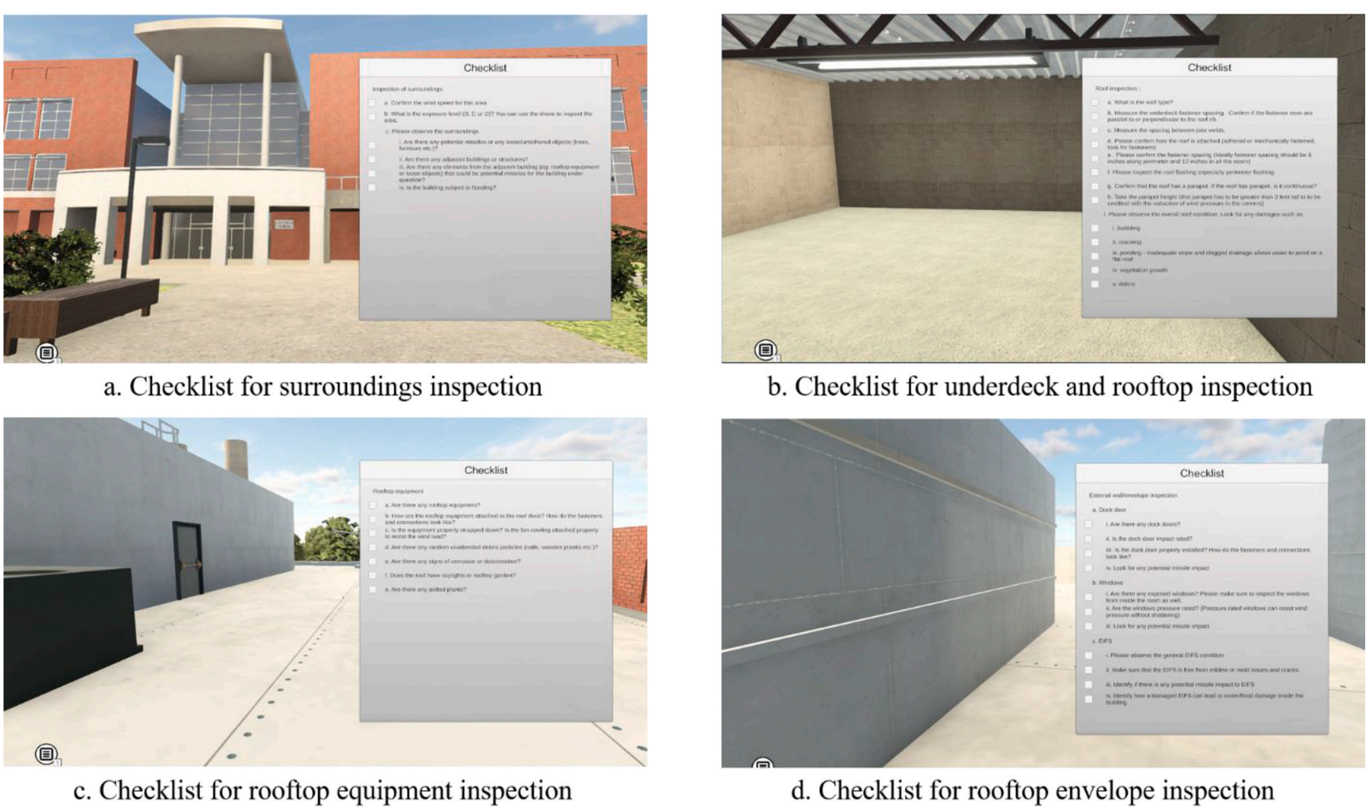
Interactive predictive visualization. This decision aid was developed by augmenting the checklist based visualization with an interactive display illustrating potential damages to the infrastructure system. Potential damages to the simulated infrastructure system (Damage State 4 as defined in HAZUZ) were shown in the interactive display as illustrated in Fig. 4. As per Hazus hurricane model user guide, severe damage involves major window damage or roof sheathing loss, major roof cover loss, and/or extensive damage to the interior from water (Hazus Hurricane Model User Guidance, 2018; Liao, 2007). However, this visualization shows only some possibilities of damages if there is a severe weather condition. What could actually happen will depend on several uncertain factors such as age of the infrastructure system, wind speed, location and materials. The interactive display prediction is hypothesized to guide participants through the inspection tasks by enhancing their Level 3 SA. The participants were not able to access both the predictive display and the checklist at the same time.

2.6.2. Dependent variables

2.6.2.1. Situation awareness. An adaptation of the Situation Awareness Global Assessment Technique (SAGAT) was used to assess the SA of the participants. This measure was originally developed to quantify the SA requirements of operators across all of its elements in the aviation domain (Endsley, 1995). The underlying assumption behind this global SA measure is the 3-level SA theory (Endsley, 1995). This technique is used for objectively calculating the SA requirements of operators at three different levels of SA using a freeze probe protocol. A higher level of accuracy in the operator's answer is attributed to higher levels of SA. The method requires the simulation to freeze at randomly selected times to administer SA queries. During the simulation freezes, a blank screen was shown to the participants.

Since no standardized questions querying SA requirements for risk inspection task exist, the SAGAT queries used in this study were developed based on the insights gained from detailed one on one interviews with 10 risk engineers (Agnisarman et al., 2018). In addition, in this study, these queries were not administered at randomly selected times; rather they were administered at predefined times as was done in a previous study investigating the SA of medical trainees (Gardner et al., 2017). The questions were presented at five pre-selected intervals during the simulation. However, the participants were told that the simulation would freeze at randomly selected times. They didn't know when the simulation was going to freeze. Each set of questions was presented following the completion of each task except for the second task (inspection of roof underdeck, roof condition, flashing, roof deck, attachments and obtaining building dimensions). As this task involved more steps than the other tasks, the simulation froze once during the task and after task completion. Questions representing perception, comprehension and prediction phases of SA were included in each freeze. Questions probing level 1 SA required participants to respond to questions about the elements in the environment. Level 2 SA questions probed participants' understanding of the current state of the environment. These questions tested participants' ability to comprehend the cues perceived. Level 3 SA questions tested participants' ability to predict the future state of the building in the event of a high speed wind condition.

2.6.2.2. Workload. Uncertainty or ambiguity in information leads to increased cognitive load while making sense of such information (Block, 2013; Zuk and Carpendale, 2006). Visualizing these uncertainties will facilitate decision making. However, adding additional elements about uncertainties in the visualization can, in turn, increase the cognitive load on users (Block, 2013). Ideally, the integrated visualization design proposed in this study should result in decreased cognitive load. Though



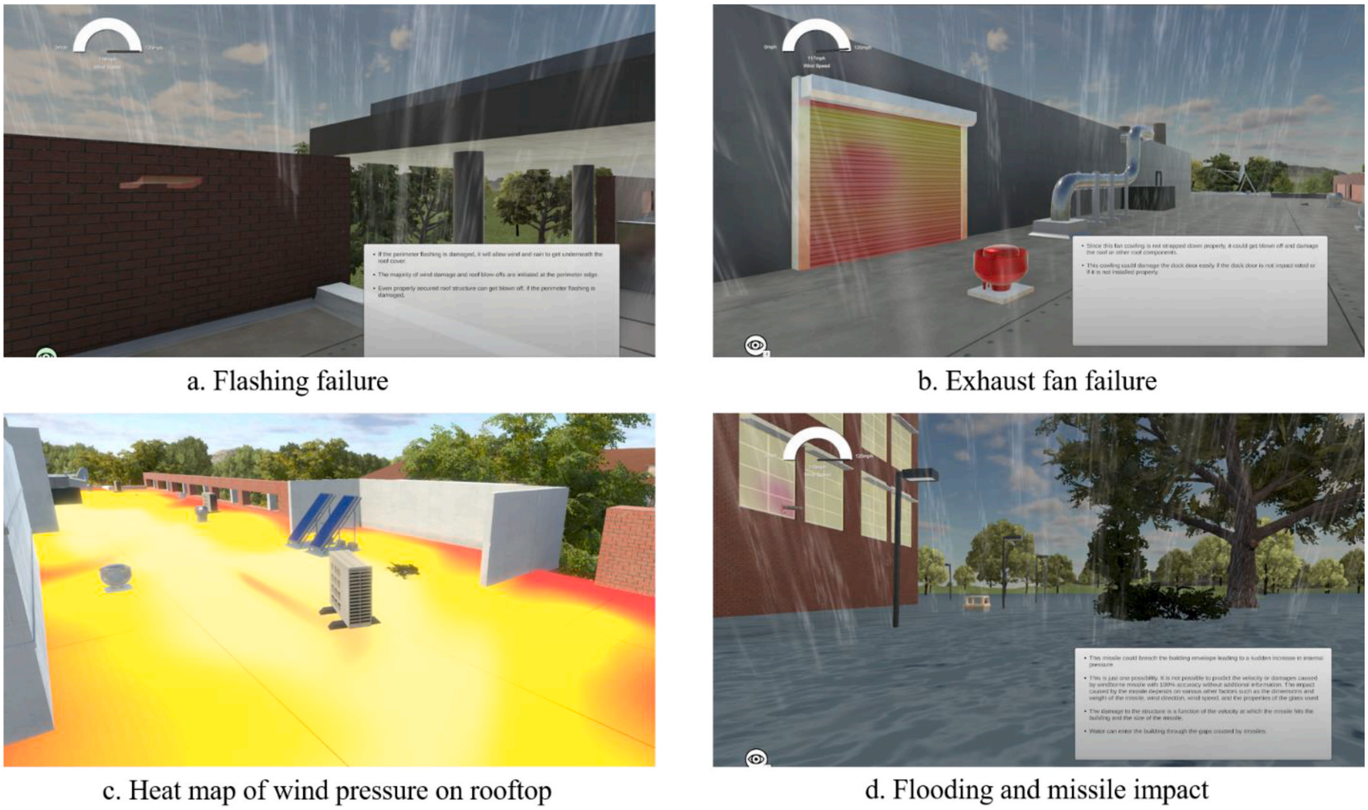
a. Checklist for surroundings inspection

b. Checklist for underdeck and rooftop inspection

c. Checklist for rooftop equipment inspection

d. Checklist for rooftop envelope inspection

Fig. 3. Examples of the checklist used in the study.



a. Flashing failure
b. Exhaust fan failure
c. Heat map of wind pressure on rooftop
d. Flooding and missile impact

Fig. 4. Examples of the predictive display used in the study.

measuring cognitive load directly can be challenging, this study used workload as an indirect measure of it (Block, 2013). The workload was subjectively measured using The National Aeronautics and Space Administration Task Load Index (NASA-TLX) questionnaire (Hart, 2006; Hart and Staveland, 1988). As measuring workload after each task can be time consuming, NASA TLX was administered upon completion of all inspection tasks.

2.6.2.3. Performance. Higher SA does not guarantee improved performance. According to Endsley and Garland (2000), there is only a probabilistic relationship between SA and performance. In this research, the performance of participants was measured to study the improvement, if any, as a result of using context-based visual decision aids using a multidimensional approach. A performance questionnaire was designed using the format of a typical school exam, with each correct response contributing to the overall score determined as the sum of correct responses. This performance test was designed based on the tasks assigned to the participants, and the survey asked questions about the tasks completed in the simulation. Though the difference between the SAGAT questionnaire and the performance questionnaire is subtle, the former does not include procedural questions. The performance test was validated by a subject matter expert.

2.6.2.4. Time. Time taken to complete the inspection task was tracked in the simulation. In real-world, time taken to complete the inspection task depends on factors such as complexity and size of building, type of roof and number and type of rooftop equipment. However, in this controlled study, all participants exposed to any experimental conditions carried out the inspection tasks using the same simulation. So, any difference in their time taken to complete the task can be attributed to the experimental condition and their individual differences.

2.7. Procedure

To examine the context-based visual decision aids, the entire inspection scenario was simulated using the Unity game engine. The complexity of the inspection tasks was simplified significantly for novice participants. Following a between-subjects experimental design, each participant was randomly assigned to one study condition. The study began with the researcher explaining the study procedure. This step was followed by the participants signing the consent form and then completing a demographic questionnaire. A video was then presented to the participants to explain the various steps involved in the windstorm risk inspection process. More specifically, the video explained and exemplified the types of issues observed in the real-world as well as the tasks the participants were expected to complete. Next, following the random assignment, the participants completed a training scenario in a simulated environment, which used the simulation of a warehouse building with various pieces of rooftop equipment. Through this simulation, participants became familiar with the navigation controls and decision aids (only for the participants in the decision aid condition).

The participants were then exposed to the study condition and the tasks they were assigned to complete in the simulation. They were able to take notes during the inspection process using a pen and paper provided. After each task, the participants were asked to complete the SAGAT questions; however, they were not allowed to consult their notes while completing the questionnaire. At the end of the final task, participants were given the performance and NASA-TLX questionnaires; while completing the performance questionnaire, participants were able to use their notes. They then participated in a retrospective think aloud session where they were asked to reflect on their performance. This procedure is illustrated in Fig. 5.

2.8. Data analysis

R language for statistical computing (R Core Team, 2019) was used

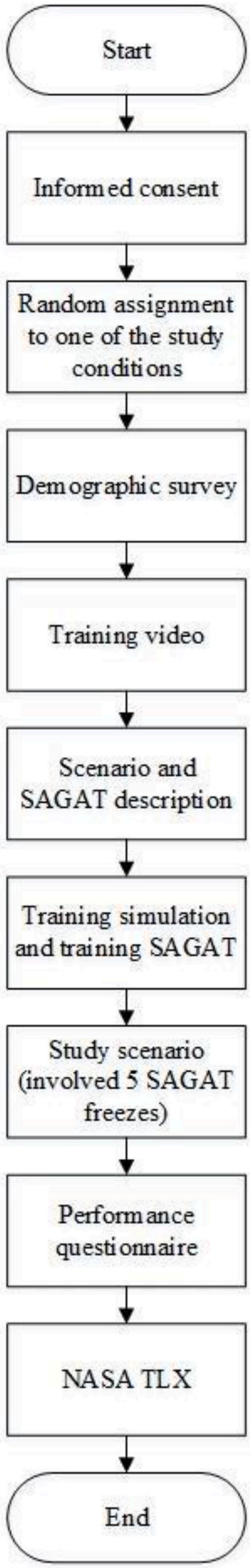


Fig. 5. Flow chart outlining experiment procedure.

for data analysis. Cook's Distance was used to identify influential cases. Standardized deviance residuals and standardized residuals were used to detect outlier values. The SAGAT responses were analyzed using multilevel binary logistic regression with a logit link function. For this variable, an additional independent variable indicating the SA level was also considered in the analysis. The SAGAT questions were categorized into three levels based on the SA level each represented. Questions related to the perception phase were categorized under Level 1 SA, questions related to the comprehension phase under Level 2 SA and questions related to the prediction phase under Level 3 SA. This variable was included in the analysis to identify the specific effects of the decision aids on the perception, comprehension and prediction phases of SA.

Workload data collected using the NASA-TLX and the performance data were analyzed using one-way between-subjects ANOVA. These dependent variables were tested for normality using the Shapiro-Wilk's test, and standardized residual values were calculated to identify extreme outliers (values greater/less than ± 3). Homogeneity of variances was tested using Levene's test of homogeneity of variances. In addition, Cook's Distance was used to identify any influential cases.

3. Results

3.1. SAGAT

SAGAT responses were coded as 1 (if the response is correct) and 0 (if the response is wrong). Each SAGAT query was analyzed individually to allow for comparisons to be made among the different conditions (Stanton et al., 2004). Separate multilevel logistic regression analyses were conducted to analyze the five sets of SAGAT responses recorded following the simulation freeze. The lme4 package available in R was used for analyzing SAGT responses (Bates et al., 2015). The multilevel logistic regression model for the SAGAT queries was built iteratively, with the intercept only model being used as the baseline and the final model including the experimental conditions presented and the SA levels and/or the interaction between the types of visual aids and the SA level. No extreme data points were identified as assessed by deviance residuals and Cook's Distance.

3.1.1. Inspection of surroundings (SAGAT 1)

The first set of SAGAT responses was recorded following the completion of the first task, which involved the inspection of building surroundings to identify the exposure level and to evaluate missile impact to the building. Following this task, the first SAGAT questionnaire containing 10 questions was administered. The multilevel logistic regression model was built iteratively. Table 2 illustrates the details of the iterative model building.

A test of the full model with two independent variable and one 2-way interaction effect was significant, χ^2 (9, N = 65) = 111.87, $p < 0.001$, $R^2_L = 0.13$. The main effects of type of visual decision aid ($\Delta\chi^2 = 37.53$, $p < 0.001$) and SA level are significant ($\Delta\chi^2 = 36.66$, $p < 0.001$). The analysis revealed an interaction between experimental condition and the SA level, $\Delta\chi^2 = 17.42$, $p = 0.002$. Post-hoc analysis was carried out to investigate the nature of this interaction. This interaction effect is illustrated in Table 3 and Fig. 6.

As illustrated in Fig. 6, participants exposed to the experimental conditions had higher SA compared to participants exposed to the control condition. However, this difference was moderated by the SA level. More specifically, there was no significant difference in the SA among participants exposed to the control, checklist and predictive conditions when they were questioned on their perception state of SA. Participants in the checklist condition (b = 1.625, p = 0.02, OR = 5.08, (95% CI: 1.10, 23.36)) and predictive display condition (b = 2.98, p = 0.0001, OR = 19.59, (95% CI [2.71, 141.35])) had significantly higher SA than participants in the control condition when they were questioned on their Level 2 SA. There was no significant difference between the SA of participants exposed to the checklist condition and the

Table 2
Model summary for multilevel logistic regression analysis for inspection of surroundings (SAGAT 1).

Variable	Model1				Model2 ($\Delta\chi^2 = 139.00$, df = 1, $p < 0.001$), $R^2_L = 0.14$				Model3 ($\Delta\chi^2 = 37.70$, df = 2, $p < 0.001$), $\Delta R^2_L = 0.05$, $R^2_L = 0.08$				Model4 ($\Delta\chi^2 = 82.96$, df = 2, $p < 0.001$), $\Delta R^2_L = 0.02$, $R^2_L = 0.10$				Model5 ($\Delta\chi^2 = 9.78$, df = 4, $p = 0.04$), $\Delta R^2_L = 0.03$, $R^2_L = 0.13$			
	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper
Constant	1.25 (0.09)	3.49	2.90	4.20	1.42 (0.15)	4.14	3.13	5.79	0.41 (0.16)	1.51	1.10	2.11	0.48 (0.18)	1.62	1.13	2.37	0.79 (0.21)	2.21	1.46	3.44
Experimental Condition (type of visualization)																				
Checklist									1.25 (0.25)	3.49	2.14	5.94	1.29 (0.26)	3.63	2.19	6.29	0.81 (0.33)	2.26	1.20	4.39
Predictive Display									1.68 (0.28)	5.36	3.16	9.61	1.73 (0.29)	5.62	3.27	10.27	1.05 (0.34)	2.85	1.48	5.73
SA level																				
Level 2													−0.70 (0.24)	0.49	0.31	0.79	−1.40 (0.38)	0.25	0.11	0.51
Level 3													0.42 (0.29)	1.53	0.88	2.74	−0.39 (0.37)	0.68	0.32	1.42
Interaction between Condition and SA Level																				
Checklist: SALevel2																	0.81 (0.56)	2.25	0.76	6.94
Predictive display: SALevel2																	1.93 (0.70)	6.88	1.87	30.20
Checklist: SALevel3																	2.62 (1.10)	13.73	2.26	267.11
Predictive display: SALevel3																	1.67 (0.85)	5.31	1.17	38.49

Table 3
Mean probability of correctly answering SAGAT questions for inspection of surroundings task (SAGAT 1).

		SA level		
		Level 1	Level 2	Level 3
Type of visualization	Control	0.69	0.35	0.60
	Checklist	0.83	0.73	0.98
	Predictive display	0.86	0.91	0.96

predictive display condition when questioned on their Level 2 SA ($b = 1.35$, $p = 0.47$, $OR = 3.86$, (095% CI: 0.53, 28.11)). Furthermore, participants exposed to the checklist ($b = 3.43$, $p = 0.03$, $OR = 30.97$, (95% CI: 1.12, 850.39)) and the predictive display condition ($b = 2.71$, $p = 0.02$, $OR = 15.11$, (95% CI [1.25, 182.24])) had significantly higher SA than participants in the control condition when they were probed on their Level 3 SA. However, the SA of participants exposed to the predictive display condition and the checklist condition when probed on their Level 3 SA ($b = -\ 0.49$, $p = 0.57$, $OR = 0.49$, (95% CI [0.01, 23.46])) were not significantly different.

3.1.2. Inspection of underdeck and rooftop (SAGAT 2)

The second set of SAGAT responses was recorded during the second task, which involved underdeck inspection and rooftop inspection. More specifically, the participants measured the underdeck and rooftop fastener spacing and the distance between joist welded connections and inspected the general condition of the roof deck. In the middle of this task, the second SAGAT questionnaire containing eight questions was administered, and the multilevel model was again built iteratively. Table 4 illustrates the details of iterative model building.

A test of the full model with two independent variables and one 2-way interaction effect was significant, χ^2 (9, $N = 65$) $= 237.02$, $p < 0.001$, $R^2_L = 0.25$. The main effects of type of visual decision aid ($\Delta\chi^2 = 17.42$, $p = 0.002$) and SA level were significant ($\Delta\chi^2 = 82.96$, $p < 0.001$). The analysis revealed an interaction between experimental condition and the SA level, $\Delta\chi^2 = 9.78$, $p = 0.04$. Table 5 and Fig. 7 explain this interaction effect.

As illustrated in Fig. 7, participants exposed to both the experimental conditions had higher situation awareness compared to participants exposed to the control condition. However, this difference is moderated by the SA level. More specifically, the SA of participants exposed to the control, checklist and predictive display condition did not differ significantly when they were asked questions about the perception phase of SA. However, participants in the checklist condition ($b = 1.73$, $p = 0.005$, $OR = 5.66$, (95% CI, 1.36 to 23.62)) and predictive display condition ($b = 2.61$, $p < 0.001$, $OR = 13.62$, (95% CI [3.11, 59.68])) had significantly higher SA than participants in the control condition when

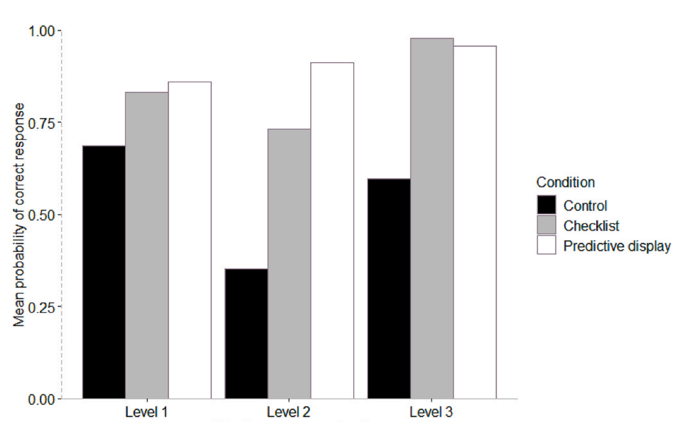


Fig. 6. Interaction effect of types of SA level on the relationship between SA and the decision aids presented (inspection of surroundings - SAGAT 1).

Table 4
Model summary for multilevel logistic regression analysis for inspection of underdeck and rooftop (SAGAT 2).

Variable	Model1				Model2 ($\Delta\chi^2 = 139.00$, df = 1, $p < 0.001$), $R^2_L = 0.06$				Model3 ($\Delta\chi^2 = 37.70$, df = 2, $p < 0.001$), $\Delta R^2_L = 0.06$, $R^2_L = 0.11$				Model4 ($\Delta\chi^2 = 82.96$, df = 2, $p < 0.001$), $\Delta R^2_L = 0.14$, $R^2_L = 0.23$				Model5 ($\Delta\chi^2 = 9.78$, df = 4, $p = 0.04$), $\Delta R^2_L = 0.02$, $R^2_L = 0.25$			
	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper
Constant	0.63 (0.09)	1.87	1.57	2.25	0.79 (0.17)	2.19	1.58	3.16	−0.31 (0.20)	0.74	0.48	1.10	1.05 (0.31)	2.86	1.58	5.43	1.33 (0.37)	3.80	1.89	8.24
Experimental Condition (type of visualization)																				
Checklist									1.06 (0.30)	2.88	1.62	5.36	1.35 (0.38)	3.86	1.86	8.47	0.88 (0.58)	2.41	0.79	7.86
Predictive Display									2.10 (0.32)	8.17	4.43	16.26	2.59 (0.41)	13.31	6.21	31.65	1.79 (0.68)	5.97	1.67	25.51
Situation awareness level																				
Level 2													−2.23 (0.29)	0.11	0.06	0.19	−2.67 (0.44)	0.07	0.03	0.16
Level 3													−2.36 (0.39)	0.09	0.04	0.20	−2.98 (0.68)	0.05	0.01	0.18
Interaction between Condition and SA Level																				
Checklist: SALevel2																	2.52 (1.15)	12.45	1.42	143.51
Predictive display: SALevel2																	0.85 (0.63)	2.35	0.67	8.06
Checklist: SALevel3																	0.83 (0.73)	2.28	0.50	9.09
Predictive display: SALevel3																	0.19 (0.92)	1.21	0.20	7.78

Table 5
Mean probability of correctly answering SAGAT questions for underdeck and rooftop inspection task (SAGAT 2).

SA level				
		Level 1	Level 2	Level 3
Type of visualization	Control	0.78	0.22	0.18
	Checklist	0.88	0.59	0.37
	Predictive display	0.95	0.76	0.92

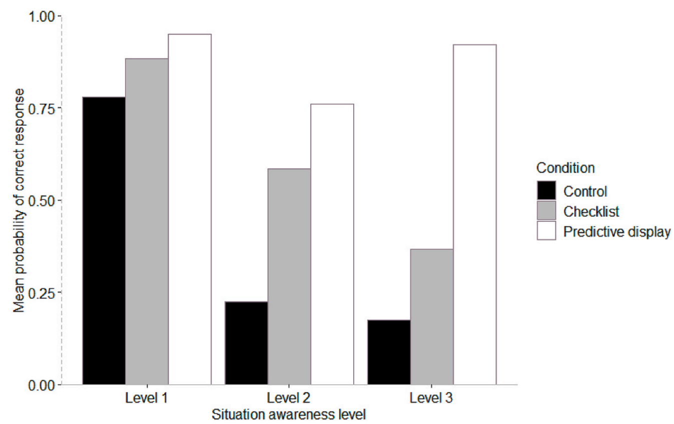


Fig. 7. Interaction effect of types of SA level on the relationship between SA and the decision aids presented (inspection of underdeck and rooftop - SAGAT 2).

they were asked questions related to the comprehension stage of SA. However, the SA of participants exposed to the two experimental conditions when probed on their Level 2 SA ($b = 0.88$, $p = 0.56$, OR = 2.41, (95% CI, 0.61 9.57)) did not differ significantly. Similarly, participants exposed to the predictive display condition had significantly higher SA than participants in both control condition ($b = 4.31$, $p < 0.001$, OR = 74.31, (95% CI [3.17, 1740.11])) and checklist condition ($b = 3.24$, $p = 0.002$, OR = 25.41, (95% CI: 1.30, 496.28)) when they were probed on their Level 3 SA. However, the SA of participants exposed to the checklist decision aid and no decision aid ($b = 1.07$, $p = 0.92$, OR = 2.92, (095% CI [0.24, 36.07])) did not differ significantly when probed on their Level 3 SA.

3.1.3. Inspection of underdeck and rooftop continuation (SAGAT 3)

The third set of SAGAT responses was recorded following the completion of the second task. This questionnaire contained eight questions, and the multilevel model was built iteratively. Table 6 illustrates the details of this iterative model building. As this table shows, the model including only the main effect of SA level, and the model including the main effects of SA level and types of visualization and the interaction effect of these two variables are not significantly different from the model containing only the main effect of type of visualization. Thus, the main effect of SA level and the interaction effect between the type of visualization and SA level were removed from the model. Model 3 is used as the final model.

A test of the model with type of visualization against the baseline model is significant χ^2 (3, N = 65) = 127.62, $p < 0.001$, $R^2_L = 0.09$, indicating that the predictor reliably distinguished participants who correctly answered the SAGAT questionnaire from those who did not. As illustrated in Fig. 8, participants exposed to the checklist ($b = 1.24$, $p = 0.0001$, OR = 3.45, (95% CI [1.70, 6.98])) and the predictive display ($b = 1.85$, $p < 0.001$, OR = 6.33, (95% CI [2.95, 13.59])) conditions had higher SA than participants who were not exposed to any decision aids. However, the SA of participants assigned to the experimental conditions ($b = 0.61$, $p = 0.16$, OR = 1.83, (95% CI [0.84, 4.02]) did not differ

Table 6
Model summary for multilevel logistic regression analysis for the second part of inspection of underdeck and rooftop (SAGAT 3).

Variable	Model1				Model2 ($\Delta\chi^2 = 143.61$, df = 1, p < 0.001), $R^2_L = 0.05$				Model3 ($\Delta\chi^2 = 30.22$, df = 2, p < 0.001), $\Delta R^2_L = 0.04$, $R^2_L = 0.09$				Model4 ($\Delta\chi^2 = 3.11$, df = 2, p = 0.211), $\Delta R^2_L = 0.004$, $R^2_L = 0.10$				Model5 ($\Delta\chi^2 = 2.81$, df = 4, p = 0.59), $\Delta R^2_L < 0.001$, $R^2_L = 0.10$			
	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper
Constant	1.06 (0.09)	2.89	2.43	3.46	1.30 (0.17)	3.68	2.66	5.37	0.23 (0.20)	1.26	0.85	1.92	0.40 (0.24)	1.49	0.94	2.37	0.32 (0.27)	1.38	0.81	2.35
Experimental Condition (type of visualization)																				
Checklist									1.24 (0.30)	3.45	1.92	6.46	1.25 (0.30)	3.48	1.92	11.60	1.58 (0.44)	4.83	2.06	11.33
Predictive Display									1.85 (0.33)	6.33	3.40	12.63	1.86 (0.33)	6.41	3.37	12.18	1.79 (0.45)	5.97	2.46	14.48
Situation awareness level																				
Level 2													−0.41 (0.23)	0.66	0.42	1.05	−0.42 (0.35)	0.66	0.33	1.31
Level 3													−0.13 (0.24)	0.88	0.55	1.40	0.14 (0.35)	1.15	0.58	2.30
Interaction between Condition and SA Level																				
Checklist: SALevel2																	−0.32 (0.54)	0.73	0.25	2.13
Predictive display: SALevel2																	0.45 (0.61)	1.58	0.48	5.24
Checklist: SALevel3																	−0.71 (0.55)	0.49	0.17	1.47
Predictive display: SALevel3																	−0.25 (0.60)	0.78	0.24	2.55

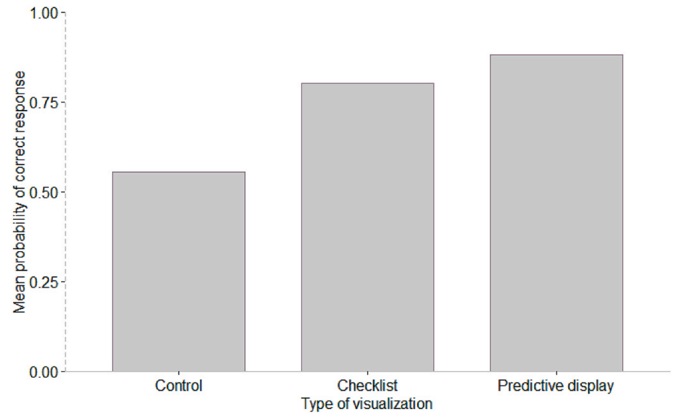


Fig. 8. Main effect of the decision aids presented (inspection of underdeck and rooftop continuation - SAGAT 3).

significantly. The mean probability value can be found in Table 7.

3.1.4. Inspection of rooftop equipment (SAGAT 4)

The fourth set of SAGAT responses was recorded following the completion of the third task, which involved the inspection of rooftop equipment. Participants had to inspect how the equipment on rooftop is fastened to the roof in addition to how equipment and other components on the roof will be affected in the event of extreme weather conditions. The SAGAT questionnaire contained eight questions, and the multilevel model was built iteratively. Table 8 illustrates the details of the iterative model building. As shown in this table, the model containing the interaction effect of the type of visualization and the SA level is not significantly different from the model containing only the main effects of the independent variables. Thus, the interaction effect was removed from the model. Model 4 is used as the final model.

A test of model with the main effect of independent variables was significant χ^2 (5, N = 65) = 135.06, p < 0.001, $R^2_L = 0.15$. The main effects of type of visual decision aid ($\Delta\chi^2 = 37.75$, p < 0.001) and SA level were significant ($\Delta\chi^2 = 33.53$, p < 0.001). Participants assigned to the predictive display conditions had higher SA than participants in the checklist condition (b = 1.45, p = 0.001, OR = 4.26, (95% CI [1.43, 12.75])) and the control condition (b = 2.23, p < 0.001, OR = 9.26, (95% CI [3.04, 28.21])). This is shown in Fig. 9. However, there was no significant difference between SA of participants in the checklist condition and control condition (b = 0.78, p = 0.18, OR = 2.17, (95% CI [0.86, 5.47])). The mean probability value can be found in Table 9. As illustrated in Fig. 10, the participants' Level 2 SA was significantly lower than their Level 1 SA (b = −1.56, p < 0.001, OR = 0.21, (95% CI [0.09, 0.50])) and Level 3 SA (b = −1.04, p = 0.003, OR = 0.353, (95% CI [0.15, 0.81])). However, Level 1 SA was not significantly different from Level 3 SA (b = 0.53, p = 0.51, OR = 1.70, (95% CI [0.76, 3.79])). The mean probability value can be found in Table 9.

3.1.5. Inspection of envelope (SAGAT 5)

The fifth set of SAGAT responses was recorded following the completion of the fourth and final task, which involved the inspection of the envelope. The envelope included windows, doors/dock doors, and EIFS. To make the inspection task less complex, the tasks only included

Table 7
Mean probability of correctly answering SAGAT questions for the second part of underdeck and rooftop inspection task (SAGAT 3).

Type of visualization	
Control	0.55
Checklist	0.80
Predictive display	0.88

Table 8
Model summary for multilevel logistic regression analysis for inspection of rooftop equipment (SAGAT 4).

Variable	Model1				Model2 ($\Delta\chi^2 = 109.76$, df = 1, $p < 0.001$), $R^2_L = 0.04$				Model3 ($\Delta\chi^2 = 37.73$ df = 2, $p < 0.001$), $\Delta R^2_L = 0.06$, $R^2_L = 0.09$				Model4 ($\Delta\chi^2 = 33.53$, df = 2, $p < 0.001$), $\Delta R^2_L = 0.06$, $R^2_L = 0.15$				Model5 ($\Delta\chi^2 = 4.91$, df = 4, $p = 0.30$), $\Delta R^2_L < 0.001$, $R^2_L = 0.16$			
	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper
Constant	0.85 (0.09)	2.33	1.94	2.82	1.00 (0.16)	2.74	2.02	3.90	0.10 (0.19)	1.10	0.76	1.61	0.69 (0.25)	1.99	1.21	3.35	0.66 (0.30)	1.93	1.08	3.58
Experimental Condition (type of visualization)																				
Checklist									0.70 (0.27)	2.02	1.19	3.54	0.78 (0.30)	2.17	1.21	4.05	0.85 (0.46)	2.33	0.97	5.91
Predictive Display									2.04 (0.33)	7.69	4.16	15.47	2.22 (0.36)	9.26	4.73	20.04	2.24 (0.62)	9.38	3.04	36.85
Situation awareness level																				
Level 2													−1.56 (0.28)	0.21	0.12	0.36	−1.40 (0.44)	0.25	0.10	0.57
Level 3													−0.53 (0.26)	0.59	0.35	0.98	−0.56 (0.38)	0.57	0.27	1.19
Interaction between Condition and SA Level																				
Checklist: SALevel2																	−0.01 (0.62)	0.99	0.29	3.37
Predictive display: SALevel2																	−0.56 (0.76)	0.57	0.12	2.43
Checklist: SALevel3																	−0.20 (0.56)	0.82	0.27	2.47
Predictive display: SALevel3																	0.87 (0.87)	2.39	0.43	14.68

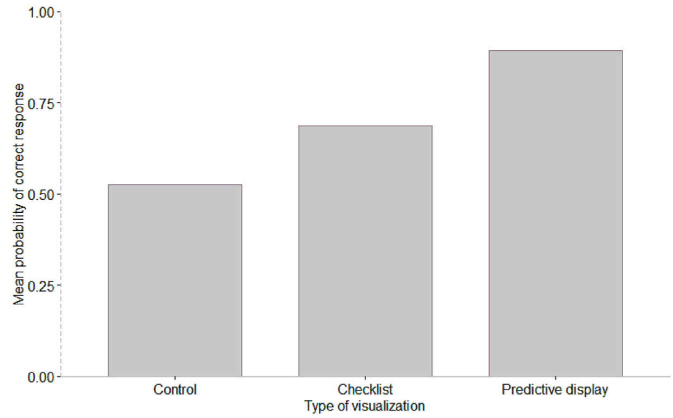


Fig. 9. Main effect of the decision aids presented (inspection of rooftop equipment - SAGAT 4).

Table 9
Mean probability of correctly answering SAGAT questions for inspection of rooftop equipment (SAGAT 4).

Types of visualization		SA level	
Control	0.52	Level 1	0.81
Checklist	0.69	Level 2	0.52
Predictive display	0.89	Level 3	0.72

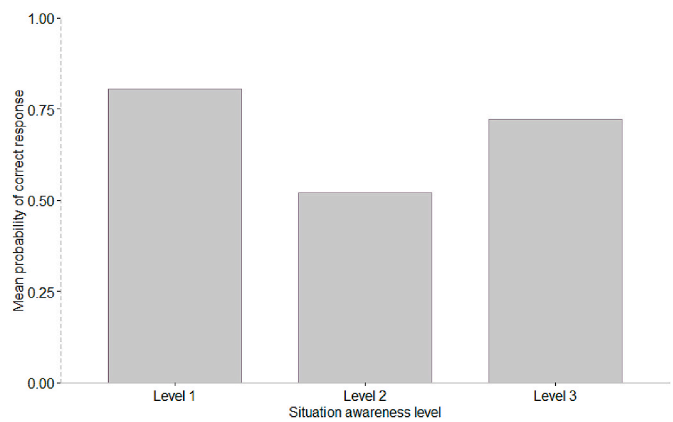


Fig. 10. Main effect of situation awareness level (inspection of rooftop equipment - SAGAT 4).

the inspection of the envelope of the rooms on the rooftop. The SAGAT questionnaire contained eight questions, and the multilevel model was built iteratively. Table 10 illustrates the details of the iterative model building. As shown in the table, the model containing the interaction effect of the type of visualization and SA level is not significantly different from the model containing only the main effects. Thus, the interaction effect was removed from the model. Model 4 is used as the final model.

A full model with the main effect of independent variables was significant χ^2 (5, N = 65) = 240.04, $p < 0.001$, $R^2_L = 0.23$. The main effect of type of visual decision aid ($\Delta\chi^2 = 28.33$, $p < 0.001$) and SA level is significant ($\Delta\chi^2 = 85.93$, $p < 0.001$). As illustrated in Fig. 11, SA of the participants in the predictive display condition significantly differed from that of participants in the control condition (b = 2.55, $p < 0.001$, OR = 12.80, (95% CI [2.90, 56.38])). Participants exposed to the checklist conditions had marginally significantly higher SA than participants in the control condition (b = 1.31, p = 0.06, OR = 3.71, (95% CI [0.98, 14.06])). However, the SA of participants exposed to the predictive display condition and the checklist condition (b = 1.24, p = 0.15,

Table 10
Model summary for multilevel logistic regression analysis for inspection of envelope (SAGAT 5).

Variable	Model1				Model2 ($\Delta\chi2 = 141.82$, df = 1, $p < 0.001$), R2L = 0.06				Model3 ($\Delta\chi2 = 28.08$, df = 2, $p < 0.001$), ΔR2L = 0.05, R2L = 0.10				Model4 ($\Delta\chi2 = 85.93$, df = 2, $p < 0.001$), ΔR2L = 0.15, R2L = 0.23				Model5 ($\Delta\chi2 = 2.12$, df = 4, $p = 0.71$), ΔR2L = 0.004, R2L = 0.24			
	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper	B (SE)	OR	CI Lower	CI Upper
Constant	0.75 (0.09)	2.11	1.76	2.55	0.94 (0.18)	2.55	1.82	3.72	−0.08 (0.23)	0.92	0.58	1.45	0.89 (0.34)	2.45	1.26	4.92	0.94 (0.38)	2.56	1.22	5.61
Experimental Condition (type of visualization)																				
Checklist									1.02 (0.34)	2.78	1.44	5.35	1.31 (0.43)	3.71	1.60	9.04	1.42 (0.61)	4.12	1.29	14.33
Predictive Display									2.00 (0.37)	7.42	3.40	15.32	2.55 (0.48)	12.79	5.22	35.37	2.11 (0.69)	8.21	2.28	35.58
Situation awareness level																				
Level 2													−2.39 (0.31)	0.09	0.05	0.16	−2.36 (0.46)	0.09	0.04	0.23
Level 3													−0.51 (0.33)	0.60	0.32	1.14	−0.70 (0.46)	0.49	0.20	1.21
Interaction between Condition and SA Level																				
Predictive display: SALevel2																	−0.23 (0.67)	0.79	0.21	2.93
Predictive display: SALevel2																	0.35 (0.74)	1.42	0.31	5.91
Checklist: SALevel3																	−0.01 (0.73)	0.99	0.24	4.16
Predictive display: SALevel3																	1.31 (1.02)	3.70	0.55	33.84

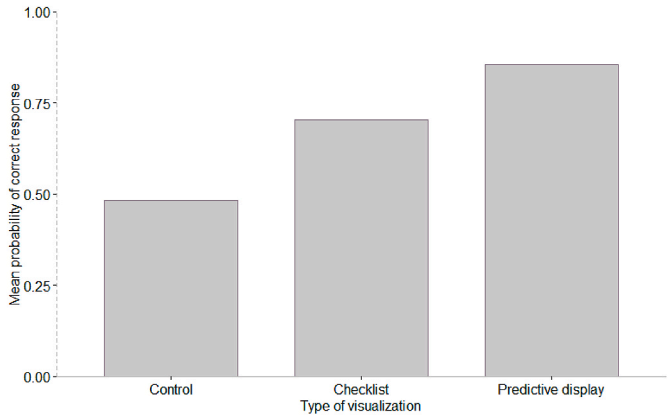


Fig. 11. Main effect of decision aids presented (inspection of envelope - SAGAT 5).

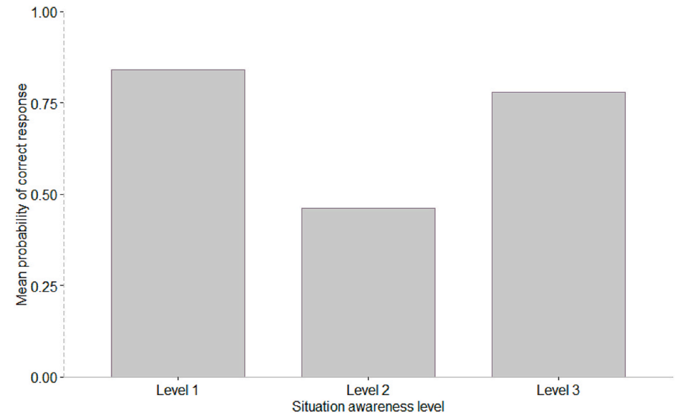


Fig. 12. Main effect of situation awareness level (inspection of envelope - SAGAT 5).

OR = 3.45, (95% CI [0.82, 14.49])) did not differ significantly. The mean probability value can be found in Table 11.

As illustrated in Fig. 12, the participants' Level 2 SA was significantly lower than their Level 1 SA ($b = -2.39$, $p < 0.001$, OR = 0.09, (95% CI [0.035, 0.24])) and Level 3 SA ($b = -1.88$, $p < 0.001$, OR = 0.152, (95% CI [0.057, 0.41])). However, Level 1 SA did not differ significantly from Level 3 SA ($b = 0.51$, $p = 0.82$, OR = 1.66, (95% CI [0.61, 4.56])). The mean probability value can be found in Table 11.

3.2. Performance

Sum of the scores for correct responses to the performance questionnaire was obtained. The maximum possible score was 56, and the individual scores were converted to percentages. This score for one participant was missing completely at random (MCAR). Thus, this data point was imputed using the MICE package available in R (van Buuren and Groothuis-Oudshoorn, 2011).

A between-subjects ANOVA was carried out to analyze the impact of type of visualization on the performance score. A significant difference in performance was observed among participants exposed to different conditions ($F(2, 62) = 17.47$, $p < 0.001$, $\omega^2 = 0.34$). The performance score increased from the control condition ($M = 54.38$, $SD = 12.35$) to the checklist condition ($M = 65.83$, $SD = 14.80$) to the predictive display condition ($M = 76.70$, $SD = 9.38$). A post-hoc analysis with Bonferroni correction revealed that the mean increase in performance from the control condition to the checklist condition (11.45, 95% CI [2.16, 20.7]) was statistically significant ($p = 0.011$). Additionally, the differences in the performance scores between the control condition and the predictive display condition (22.32, 95% CI [13.03, 31.6], $p < 0.001$), and the checklist condition and the predictive display condition (10.87, 95% CI [1.69, 20.1], $p = 0.015$) were significantly different. This effect of type of visualization is illustrated in Fig. 13.

3.3. Time

The simulation tracked the time taken to complete the inspection task. One missing data point was imputed using the MICE package. A between-subjects ANOVA was carried out to analyze this variable. A

Table 11
Mean probability of correctly answering SAGAT questions for inspection of envelope (SAGAT 5).

Types of visualization		SA level	
Control	0.48	Level 1	0.84
Checklist	0.70	Level 2	0.46
Predictive display	0.86	Level 3	0.78

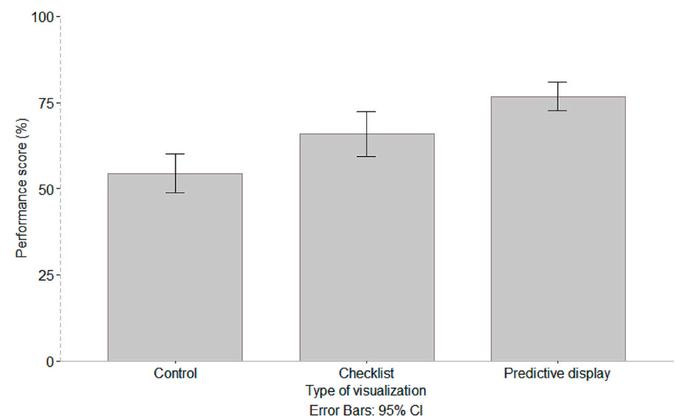


Fig. 13. Effect of decision aids presented on performance.

significant difference in time taken was observed among participants exposed to the different conditions ($F(2, 62) = 34.40$, $p < 0.001$, $\omega^2 = 0.51$). As illustrated in Fig. 14, time taken in seconds to complete the inspection tasks increased from the control condition ($M = 961.64$, $SD = 47.03$) to the checklist condition ($M = 1623.24$, $SD = 64.22$) and the predictive display condition ($M = 1713.61$, $SD = 88.26$). A post-hoc analysis with Bonferroni correction revealed that the mean increase in time taken from the control to the checklist condition (661.60, 95% CI [419, 904], $p < 0.001$]) and predictive display condition (752.00, 95% CI [509, 995], $p < 0.001$]) is statistically significant. However, the time

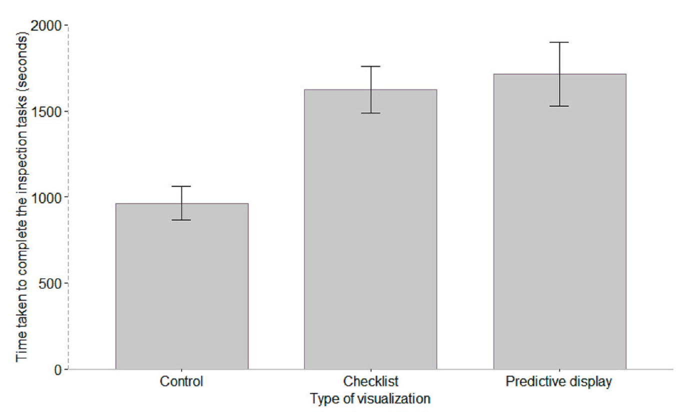


Fig. 14. Effect of decision aids presented on time taken to complete inspection tasks.

taken by the participants in the checklist condition did not differ significantly from the time taken in the predictive display condition (90.4, 95% CI [-149, 330], p = 0.99]).

3.4. Workload

3.4.1. Total workload

NASA TLX tool was used for assessing workload. A between-subjects ANOVA was carried out to analyze this variable. As illustrated in Fig. 15, total workload decreased from the control condition (M = 52.51, SD = 16.81) to the checklist condition (M = 49.56, SD = 17.46) to the predictive display condition (M = 45.92, SD = 13.74). However, these workload values were not statistically different across various test conditions ($F(2, 62) = 0.906$, $p = 0.41$, $\omega^2 = -0.003$).

3.4.2. Mental demand

Mental demand data was assessed using the NASA TLX tool. A between-subjects ANOVA was used to analyze this variable. As illustrated in Fig. 15, perceived mental demand decreased from the control condition (M = 18.23, SD = 9.29) to the checklist condition (M = 17.42, SD = 8.91) to the predictive display condition (M = 15.61, SD = 6.54). However, no significant difference in the mental demand experienced was observed among participants exposed to the different conditions ($F(2, 62) = 0.567$, $p = 0.57$, $\omega^2 = -0.013$).

3.4.3. Temporal demand

The perceived temporal demand was measured subjectively using the NASA TLX tool. A between-subjects ANOVA was carried out to analyze this variable. As illustrated in Fig. 15, perceived temporal demand increased from the control condition (M = 8.19, SD = 6.57) to the checklist condition (M = 8.39, SD = 7.99) to the predictive display condition (M = 8.73, SD = 9.70). However, no significant difference in the temporal demand experienced was observed among participants exposed to the different conditions ($F(2, 62) = 0.024$, $p = 0.98$, $\omega^2 = -0.031$).

3.4.4. Subjective performance

The subjective performance was measured using the NASA TLX tool. Higher values of performance rating indicate lower perceived performance, and lower values of performance rating indicate higher perceived performance.

A between-subjects ANOVA was conducted to investigate how the type of visualization affected the perceived performance reported by the participants. The perceived performance differed significantly among the participants exposed to the different conditions ($F(2, 62) = 4.71$, $p = 0.01$, $\omega^2 = 0.102$). As illustrated in Fig. 15, the perceived performance rating increased from the control condition (M = 11.65, SD = 6.19) to the predictive display condition (M = 8.64, SD = 5.40) to the checklist condition (M = 6.91, SD = 3.40). A post-hoc analysis with Bonferroni correction revealed that the mean increase in perceived performance from the control to the checklist condition (−4.74, 95% CI [-8.58, −0.899, p = 0.01]) was statistically significant. However, predictive display condition and the control condition (−3.01, 95% CI [-6.86, 0.828, p = 0.17]), and the checklist condition and the predictive display condition (−1.73, 95% CI [-5.53, 2.07, p = 0.80]) were not statistically different.

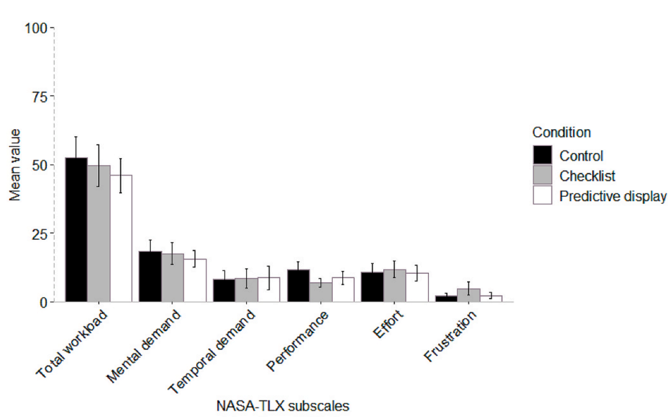


Fig. 15. Effect of decision aids presented on NASA TLX subscales.

3.4.5. Effort

The subjective effort rating was measured using the NASA TLX tool. A between-subjects ANOVA was carried to analyze this variable. As illustrated in Fig. 15, perceived effort increased from the predictive display condition (M = 10.35, SD = 6.36) to the control condition (M = 10.59, SD = 6.89) to the checklist (M = 11.80, SD = 6.81). However, these differences were not significantly different from each other ($F(2, 62) = 0.299$, $p = 0.74$, $\omega^2 = -0.022$).

3.4.6. Frustration

The subjective frustration was measured using the NASA TLX tool. The homogeneity of variance assumption was violated as assessed by Levene's test (p = 0.03); as a result, Welch's F test was used to test the hypothesis.

A one-way analysis of means not assuming equal variances using Welch's test was carried out. As illustrated in Fig. 15, perceived frustration increased from the predictive display condition (M = 2.17, SD = 2.57) to the control condition (M = 2.06, SD = 2.27) to the checklist condition (M = 4.76, SD = 5.52). However, no significant difference in the perceived effort reported was observed among participants exposed to different conditions ($F(2, 38.91) = 2.28$, $p = 0.12$).

4. Discussion

This article investigated how context-enabled visual decision aids can be used to enhance the situation awareness and performance of risk inspectors. Sixty-five civil engineering and construction science students were recruited for this study. The dependent variables measured were SAGAT, performance, NASA TLX and task time.

The visual decision aids used in this study were designed based on the user-centered design approach proposed by Endsley (2016). A checklist based decision aid and an interactive predictive visual aid were tested in this study. In general, participants exposed to these test conditions reported higher SA compared to those who were not exposed to any decision aids, suggesting that these decision aids developed for risk inspection were effective in supporting the SA requirements of the participants. Additionally, participants had higher Level 1 and Level 3 SA, a result that appears counterintuitive as the latter is more complex and difficult to achieve. However, accurate prediction of potential damages to the building led to significantly higher Level 3 SA than Level 2 SA.

For tasks requiring the participants to inspect the building surroundings and assess potential missile impact and water damage, those received the decision aids exhibited a higher Level 2 SA. Past studies have suggested that using procedural checklists could improve the SA of participants. For example, a longitudinal descriptive study investigating the effectiveness of a checklist in improving SA during physician handoffs in a pediatric emergency department reported that the users experienced improved SA with the help of a standardized checklist (Mullan et al., 2015). For the task that involves the inspection of building surroundings, participants in the predictive display condition achieved a higher Level 3 SA compared to other participants. Interactive predictive visualizations showed participants what if scenarios in the event of a Category 4 hurricane. This knowledge may have contributed to the significantly higher Level 3 SA for those participants and helped them better predict the building's future state. The cues presented in the predictive visualization situated around their SA requirements and translated the captured data into a meaningful prediction, resulting in higher SA (Endsley and Connors, 2008). A study investigating the effect of a situation-augmented display on an unmanned aerial vehicle monitoring task suggested that use of such displays may improve the SA of participants. However, this study used time to detect abnormalities as a measure of SA (Lu et al., 2013). Use of measures like SAGAT or SART may be more useful in identifying the actual effect of such visualizations on SA.

A similar trend was observed for tasks requiring the participants to inspect the general condition of a roof underdeck and a rooftop. Participants in all three conditions had the same Level 1 SA. Both experienced as well as novice personnel can have the same Level 1 SA. However, integrating this information to comprehend the situation can be challenging for novice engineers (Endsley, 2016). Though we recruited novice participants for this study, those exposed to the experimental condition achieved higher Level 2 and Level 3 SA. Participants also had to take several measurements including fastener spacing and parapet dimensions. A previous study investigating the needs and requirements of windstorm engineers revealed that taking dimensions is one of the tasks they frequently forget (Agnisarman et al., 2018). Thus, providing context-based decision aids to support this SA requirement through a checklist resulted in improved SA. Endsley (2016) suggested that providing assistance for Level 2 SA and Level 3 SA will positively influence situation awareness. The checklist helped participants thoroughly investigate the surroundings through cues and reminders. Additionally, the predictive display processed the Level 1 information and presented details supporting their Level 2 SA and provided assistance to predict potential damages to the infrastructure, leading to higher Level 2 and Level 3 SA. For example, all participants were asked to identify the areas experiencing higher wind pressure based on the presence of parapet and fastener spacing. The predictive display used a heat map to directly show this information as illustrated in Fig. 4c, leading to higher SA.

The second task additionally required the participants to inspect other roof issues including roof drainage, parapet and the general condition of the roof membrane. Most of the tasks they were asked to complete were related to such obvious issues as the identification of a clogged drain, stagnant water on the rooftop and a membrane tear. However, the interactive predictive visualization and the checklist assisted them better inspect the building, resulting in higher SA. The checklist explicitly asked them to look for these issues, leading to higher probability in correctly answering the SAGAT questions. The predictive display did not have any additional value compared to the checklist condition. Though the predictive display showed the participants the potential damages to the building under severe weather condition, participants found it easier to predict the consequence of some obvious issues like a clogged drain and discontinuous parapet.

For tasks requiring the inspection of the condition of rooftop equipment, participants in the predictive display condition had higher SA compared to participants in the control condition and the checklist condition. Building rooftop housed several improperly attached pieces of equipment. Predicting the specific behavior of some of them and some of their potential impacts was not a straightforward task. For this reason, the checklist alone was not useful enough to complete this task. However, the predictive display helped them develop a better mental model of the potential interaction among different building components. For example, as illustrated in Fig. 4d, the dislodged exhaust fan could impact the dock door and damage it. Additionally, the dock door was not impact rated or pressure rated, both of which could exacerbate the damage. Participants in the predictive display condition were given sufficient information for integrating the available cues to create an accurate mental model, leading to higher SA.

The final task required the participants to inspect the building envelope. For simplicity, participants had to inspect only the envelope of the rooms on the rooftop. Participants who used the decision aids exhibited higher SA compared to participants who did not receive any decision aids. Participants who completed the inspection tasks without any decision aids failed to identify if the windows and dock doors in the rooftop were impact rated or pressure rated. Additionally, they failed to inspect the condition of the EIFS. As the checklist and interactive visualization guided the participants through these steps and given cues to look for these details, they achieved a higher SA. The participants in the predictive display condition, nonetheless, did not have better SA than that of those in the checklist condition. As some participants suggested, predicting what could happen to a dock door that was not impact rated is pretty straightforward, suggesting that predictive visualization did not add any additional value beyond the value of checklist.

According to Endsley and Garland (2000), higher SA might lead to better performance. In this study, participants in the checklist condition performed better than those receiving no decision aids. Participants mentioned that the step by step instructions helped them keep track of all the tasks they had to complete. Additionally, it avoided the need to remember the inspection steps in their working memory. Checklists have been extensively used in commercial aviation and past research suggests they provide retrieval cues that help pilots activate the sequence of activities they must perform (Degani and Wiener, 1990; Reason, 1990; Wickens et al., 2015). Though in the domain of infrastructure risk inspection, errors of omission may not always result in a catastrophe, it could lead to building owners having to pay for a loss that could have been avoided if the inspector had detected the issue in advance. Use of a checklist reduces the chance of an omission error by limiting the reliance on memory (Rosenfield and Chang, 2009), resulting in higher performance. There is sufficient evidence in the literature suggesting improved performance with the use of checklists. For instance, a past study investigated the application of a checklist for controlling severe local anesthetic systemic toxicity situation reported improved performance for the group exposed to the checklist in a simulated environment (Neal et al., 2012). In addition to the healthcare domain, checklists are considered one of the simplest tools for reducing human error across different disciplines including aviation and product manufacturing (Hales and Pronovost, 2006). However, their effectiveness in infrastructure inspection still needs to be comprehensively investigated. To improve the effectiveness of digital checklists, as suggested by some participants, it can be augmented with pictures of issues to help users identify them in the building.

The participants in the predictive display condition exhibited higher performance than those in the checklist condition as well as those who did not receive any test conditions. For tasks involving the assessment of complex interactions like the one illustrated in Fig. 4d, the predictive display was particularly useful. These participants were aware of various direct as well as indirect consequences of a loosely attached exhaust hood. They saw how the fan hood could damage the non-impact rated dock door and the EIFS. However, for much less complicated tasks, checklists alone are sufficient. The predictive display can train novice engineers to probe the scene thoroughly to identify various interactions among different components in the building. Thus, providing an option to activate the predictive display, if necessary, will help the novice engineers. Most participants appreciated the predictive display; nonetheless, they suggested that its usefulness is limited to the training phase. However, their significant benefit to expert engineers may be limited as their experience helps them achieve Level 3 SA and predict potential damages to the building.

Though participants in both test conditions exhibited higher SAGAT and performance values, the NASA TLX workload measure was not affected by these decision aids. Despite the lack of significance in the workload score, the score was lower for the checklist and lowest for the predictive display condition in the sample. Though the use of the checklist did not result in significant reduction in workload, this finding is promising as it did not place any additional workload on participants. This research is in agreement with the findings from past studies investigating the use of a checklist for pediatric trauma resuscitation (Parsons et al., 2014). Higher workload can have a detrimental effect on

SA because of users' inability to comprehend and synthesize the cues available in the environment and by requiring the use of already limited working memory (Endsley, 2016; Mahadevan, 2009). Decision aids that reduce the demands on working memory can, in turn, eliminate excessive workload and improve SA. One example of such a decision aid is automation, which has been found to reduce mental demand and thereby improve SA (Endsley, 2016). The predictive display reduced users' mental demand by providing additional support for analyzing and interpreting the data available. It helped the participants integrate seemingly disparate cues and comprehend the data. Furthermore, participants exposed to the experimental conditions spent more time in the field completing the inspection task, a finding that was not unexpected as those participants completed more required steps than the participants in the control condition. For example, participants in the checklist and predictive display condition measured fastener spacing, welded connection spacing and parapet height. Additionally, the checklist prompted the participant to look for the general condition of rooftop equipment and roof such as tear ponding and vegetation. However, most participants in control condition failed to inspect these aspects. Furthermore, participants in the predictive display condition interacted with the predictive visualization during the inspection task. This resulted in increased task completion time for these participants.

Though the application of the decision aids had significant positive effects on performance and SA, it is important to discuss some of the behaviors observed during the study. Some participants failed to use the checklist effectively. They forgot to open it and had to be reminded to use it from time to time. Participants activated the checklist whenever they wanted. However, keeping them static in the device would eliminate the need for them to remember to activate the checklist. Further, using the checklist can lead to errors of omission if it is not comprehensive. The checklist used in this study was designed specifically for the building used in the simulation. In the real world, risk engineers encounter facilities with different roof systems, components and occupancy. Thus, there is a need to develop checklists that can be adapted to the specific condition the engineers will be investigating. It can also be augmented with representative images from real-world situations to improve cue saliency. In addition, using a predictive display can have several consequences as a result of an increased reliability on the system, leading to automation complacency (Wickens et al., 2015); because of increased clue reliance, participants failed to observe other areas despite the fact they may have issues that the predictive display failed to highlight.

This phenomenon associated with automation complacency is known as attentional narrowing or tunneling (Wickens et al., 2015). For example, the predictive display showed the potential damage for building flashing under severe weather condition. Subsequently, the participants based their conclusion about the flashing solely on the predictive visualization, failing to look for flashing issues in the other locations. Though these did not create any significant issues for the participants' SA or performance for the simplified inspection task used in this study, in a real-world application with complicated inspection tasks, these issues might affect inspectors' performance. Thus, it is important to study attentional tunneling in detail when designing AI-based decision aids for risk engineers. Multimodal cues based on AI-based algorithms can be developed to provide different types of cues such as visual, auditory and haptic to reduce the information processing demands on users (Burke et al., 2006). Multimodal displays exemplify the framework of multiple resources theory by utilizing our capability to process compatible resources at the same time (Burke et al., 2006; Wickens, 2008). We need to further investigate the performance of risk engineers while controlling automation enabled technologies such as drones to collect inspection data. Multimodal displays can be used to provide feedback on inspection tasks as well as controlling tasks. Furthermore, the usability of these systems needs to be evaluated to improve acceptance by its users (Agnisarman et al., 2017; Narasimha et al., 2018).

Furthermore, this cross-sectional study investigated how context-based decision aids influenced participants' SA and performance immediately after watching the training video and completing the training scenario. The retention effect or the training value of these decision aids is still unknown. Further follow-up studies need to be conducted without these decision aids to investigate the retention effect of these aids on user performance and SA.

This study is not without limitations. The use of convenient sampling is a major limitation of this study. This study recruited university students with relevant academic background. Furthermore, the performance questionnaire used in this study is not a validated questionnaire. The questionnaire tested participants' awareness about the scenario presented. An experienced windstorm engineer validated the questionnaire.

5. Conclusions

This experimental study investigated how checklist and predictive display based decision aids influenced the performance and situation awareness of participants using a simulated environment. The findings suggest that the participants exposed to the decision aids had higher performance and SA compared to the participants who did not receive any decision aids. The application of decision aids had a positive effect by reducing the reliance on memory. Additionally, the decision aids helped users integrate the cues available to make sense of the environment. More specifically, the checklist alone was sufficient for some tasks including the inspection of obvious issues like roof ponding, cracking and clogged drainage. However, for other tasks involving the identification of the interaction among different components in the building, the predictive display provided additional benefits. This finding is important to consider when selecting decision aids for infrastructure inspection. By providing predictive visualization for only complicated tasks, the computational demands may also be reduced.

The results suggest that the use of checklist and predictive display might result in reduced workload. However, the decision aids need to be tested with the actual windstorm risk engineers in real inspection scenarios to learn the effect of these aids on their SA and performance in a real-world situation. In addition, we noticed that use of these decision aids can lead to attentional tunneling. The potential of using additional decision aids such as haptic cues based on AI algorithms need to be investigated in detail in future research endeavors. Finally, the potential of these decision aids on training risk engineers needs to be investigated further to learn how they can be used to impart procedural knowledge as well as to improve SA. There is a need to investigate the long-term effect of these decision aids on the SA requirements of participants.

Funding

This work was supported by the United States National Science Foundation under Grant #1700621.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Mr. Andrew Brooks for the support extended to conduct the study and Dr. Bart Knijnenburg for helping us with the data analysis. This work was supported by the United States National Science Foundation under Grant #1700621.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ergon.2021.103108.

References

Agnisarman, S.O., Chalil Madathil, K., Smith, K., Ashok, A., Welch, B., McElligott, J.T., 2017. Lessons learned from the usability assessment of home-based telemedicine systems. Appl. Ergon. 58, 424–434. https://doi.org/10.1016/j.apergo.2016.08.003.

Agnisarman, S., Khasawneh, A., Ponathil, A., Lopes, S., Madathil, K.C., 2018. A qualitative study investigating the sensemaking process of engineers performing windstorm risk surveys. In: Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting Human Factors and Ergonomics Society. Meeting, 62. https://doi.org/10.1177/1541931218621402 (1), 1776–1776.

Agnisarman, S., Lopes, S., Chalil Madathil, K., Piratla, K., Gramopadhye, A., 2019a. A survey of automation-enabled human-in-the-loop systems for infrastructure visual inspection. Autom. ConStruct. 97, 52–76. https://doi.org/10.1016/j.autcon.2018.10.019.

Agnisarman, S., Madathil, K.C., Bertrand, J., 2019b. Context-based visual aids to support the situation awareness of field engineers conducting windstorm risk surveys. In: Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting Human Factors and Ergonomics Society. Meeting, 63, pp. 1898–1902. https://doi.org/10.1177/1071181319631138 (1).

Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67 (1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Block, G., 2013. Reducing Cognitive Load using Adaptive Uncertainty Visualization. [Nova Southeastern University]. https://nsuworks.nova.edu/gscis_etd/95/.

Burke, J.L., Prewett, M.S., Gray, A.A., Yang, L., 2006. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In: Proceedings of the 8th. https://dl.acm.org/citation.cfm?id=1181017.

Degani, A., Wiener, E.L., 1990. Human Factors of Flight-Deck Checklists: the Normal Checklist (NASA CR-177549). NASA, Washington, DC.

Endsley, M.R., 1995. Toward a theory of situation awareness in dynamic systems. Hum. Factors 37 (1), 32–64. https://doi.org/10.1518/001872095779049543.

Endsley, M., 1999. Situation Awareness in Aviation Systems. Handbook of Aviation Human Factors, 257–276. http://www.pacdeff.com/pdfs/Situation%20Awareness%20in%20Aviation%20Endsley%201999.pdf.

Endsley, M.R., 2016. Designing for Situation Awareness: an Approach to User-Centered Design. CRC press. https://www.taylorfrancis.com/books/9781420063585.

Endsley, M.R., Connors, E.S., 2008. In: Situation Awareness: State of the Art. 2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 1–4. https://doi.org/10.1109/PES.2008.4596937.

Endsley, M.R., Garland, D.J., 2000. Theoretical underpinnings of situation awareness: a critical review. Situation Awareness Anal. Meas. 1, 24. https://books.google.com/books?hl=en&lr=&id=WrJGDsjJakcC&oi=fnd&pg=PA3&dq=THEORETICAL+UNDERPINNINGS+OF+SITUATION+AWARENESS:+A+CRITICAL+REVIEW&ots=XI8ksyYmFP&sig=xwwNNn9iAS9nx6dCfLubbcsBJXI.

Endsley, M., Kiris, E.O., 1995. The out-of-the-loop performance problem and level of control in automation. Hum. Factors 37 (2), 381–394. https://doi.org/10.1518/001872095779064555.

Endsley, M.R., Robertson, M.M., 2000. Situation awareness in aircraft maintenance teams. Int. J. Ind. Ergon. 26 (2), 301–325. https://doi.org/10.1016/S0169-8141(99)00073-6.

Fang, Y., Cho, Y.K., Durso, F., Seo, J., 2018. Assessment of operator's situation awareness for smart operation of mobile cranes. Autom. ConStruct. 85, 65–75. https://doi.org/10.1016/j.autcon.2017.10.007.

Fernández, C.C., 2001. Hurricane Wind Damage Simulation Using GIS (Geographic Information System). Texas Tech University. https://ttu-ir.tdl.org/ttu-ir/handle/2346/21797.

Fioratou, E., Flin, R., Glavin, R., Patey, R., 2010. Beyond monitoring: distributed situation awareness in anaesthesia. Br. J. Anaesth. 105 (1), 83–90. https://doi.org/10.1093/bja/aeq137.

Gardner, A.K., Kosemund, M., Martinez, J., 2017. Examining the feasibility and predictive validity of the SAGAT tool to assess situation awareness among medical trainees. Simulat. Healthc. J. Soc. Med. Simulat.: J. Soc. Simulat. Healthcare 12 (1), 17–21. https://doi.org/10.1097/SIH.0000000000000181.

Guszcza, J., 2018, January 22. Smarter Together: Why Artificial Intelligence Needs Human-Centered Design. Deloitte Insights. https://www2.deloitte.com/insights/us/en/deloitte-review/issue-22/artificial-intelligence-human-centric-design.html.

Hales, B.M., Pronovost, P.J., 2006. The checklist—a tool for error management and performance improvement. J. Crit. Care 21 (3), 231–235. https://doi.org/10.1016/j.jcrc.2006.06.002.

Hart, S.G., 2006. Nasa-task load Index (NASA-TLX); 20 Years later. In: Proceedings of the Human Factors and Ergonomics Society ... Annual Meeting Human Factors and Ergonomics Society. Meeting, vol. 50, pp. 904–908. https://doi.org/10.1177/154193120605000909, 9.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Advances in Psychology, vol. 52. North-Holland, pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Haynes, P., Hehl-Lange, S., Lange, E., 2018. Mobile augmented reality for flood visualisation. Environ. Model. Software 109, 380–389. https://doi.org/10.1016/j.envsoft.2018.05.012.

Hazus Hurricane Model User Guidance (Version Hazus 4.2). (2018). https://www.fema.gov/media-library-data/1523994427931-0e664abd3b58607fc87b194c236e8506/Hazus_4-2_Hurricane_User_Manual_April_2018.pdf.

Hollands, J.G., Spivak, T., Kramkowski, E.W., 2019. Cognitive load and situation awareness for soldiers: effects of message presentation rate and sensory modality. Hum. Factors 61 (5), 763–773. https://doi.org/10.1177/0018720819825803.

Hurricane Costs, 2019. National Oceanic and Atmospheric Administration. https://www.coast.noaa.gov/states/fast-facts/hurricane-costs.html.

Jones, D.G., Endsley, M.R., 1996. Sources of situation awareness errors in aviation. Aviat Space Environ. Med. 67 (6), 507–512. https://www.ncbi.nlm.nih.gov/pubmed/8827130.

Kapler, T., Wright, W., 2005. GeoTime information visualization. Inf. Visual. 4 (2), 136–146. https://doi.org/10.1057/palgrave.ivs.9500097.

Kasireddy, V., Ergan, S., Akinci, B., Gulgec, N.S., 2015. Visualization requirements of engineers for risk assessment of embankment dams. Visual. Eng. 3 (1), 1. https://doi.org/10.1186/s40327-014-0014-y.

Khasawneh, A., Rogers, H., Bertrand, J., 2019. Human adaptation to latency in teleoperated multi-robot human-agent search and rescue teams. Automat. Construct. https://www.sciencedirect.com/science/article/pii/S0926580517310890.

Khazai, B., Merz, M., Schulz, C., Borst, D., 2013. An integrated indicator framework for spatial assessment of industrial and social vulnerability to indirect disaster losses. Nat. Hazards 67 (2), 145–167. https://doi.org/10.1007/s11069-013-0551-z.

Liao, D., 2007. Physically-based Visualization of Residential Building Damage Process In Hurricane. University of Central Florida. http://stars.library.ucf.edu/etd/3241/.

Loft, S., Bowden, V., Braithwaite, J., Morrell, D.B., Huf, S., Durso, F.T., 2015. Situation awareness measures for simulated submarine track management. Hum. Factors 57 (2), 298–310. https://doi.org/10.1177/0018720814545515.

Lu, J.-L., Horng, R.-Y., Chao, C.-J., 2013. Design and test of a situation-augmented display for an unmanned aerial vehicle monitoring task. Percept. Mot. Skills 117 (1), 1187–1207. https://www.ncbi.nlm.nih.gov/pubmed/24422345.

Mahadevan, S., 2009. Visualization Methods and User Interface Design Guidelines for Rapid Decision Making in Complex Multi-Task Time-Critical Environments [Wright State University. http://rave.ohiolink.edu/etdc/view?acc_num=wright1236970008.

Marusich, L.R., Bakdash, J.Z., Onal, E., Yu, M.S., Schaffer, J., O'Donovan, J., Höllerer, T., Buchler, N., Gonzalez, C., 2016. Effects of information availability on command-and-control decision making: performance, trust, and situation awareness. Hum. Factors 58 (2), 301–321. https://doi.org/10.1177/0018720815619515.

Mehta, K.C., Smith, D.A., McDonald, J.R., 1981. Procedure for predicting wind damage to buildings. J. Struct. Div. 107 (11), 2089–2096. http://cedb.asce.org/CEDBsearch/record.jsp?dockey=0010544.

Mullan, P.C., Macias, C.G., Hsu, D., Alam, S., Patel, B., 2015. A novel briefing checklist at shift handoff in an emergency department improves situational awareness and safety event identification. Pediatr. Emerg. Care 31 (4), 231–238. https://doi.org/10.1097/PEC.0000000000000194.

Narasimha, S., Agnisarman, S., Chalil Madathil, K., Gramopadhye, A., McElligott, J.T., 2018. Designing home-based telemedicine systems for the geriatric population: an empirical study. Telemed. J. e Health: Off. J. Am. Telemed. Assoc. 24 (2), 94–110. https://doi.org/10.1089/tmj.2017.0047.

Neal, J.M., Hsiung, R.L., Mulroy, M.F., Halpern, B.B., Dragnich, A.D., Slee, A.E., 2012. ASRA checklist improves trainee performance during a simulated episode of local anesthetic systemic toxicity. Reg. Anesth. Pain Med. 37 (1), 8–15. https://doi.org/10.1097/AAP.0b013e31823d825a.

NWS Analyze, Forecast and Support Office, 2018. National Weather Service. https://www.nws.noaa.gov/om/hazstats.shtml.

Parsons, S.E., Carter, E.A., Waterhouse, L.J., Fritzeen, J., Kelleher, D.C., O'connell, K.J., Sarcevic, A., Baker, K.M., Nelson, E., Werner, N.E., Boehm-Davis, D.A., Burd, R.S., 2014. Improving ATLS performance in simulated pediatric trauma resuscitation using a checklist. Ann. Surg. 259 (4), 807–813. https://doi.org/10.1097/SLA.0000000000000259.

Pascale, M.T., Sanderson, P., Liu, D., Mohamed, I., Brecknell, B., Loeb, R.G., 2019. The impact of head-worn displays on strategic alarm management and situation awareness. Hum. Factors 61 (4), 537–563. https://doi.org/10.1177/0018720818814969.

Qualtrics (Version March, 2019). (2005). [Computer software]. https://www.qualtrics.com.

R Core Team, 2019. R: A Language and Environment for Statistical Computing (Version 3.6.1). R Foundation for Statistical Computing [Computer software]. https://www.R-project.org/.

Reason, J., 1990. Human Error. Cambridge University Press. https://play.google.com/store/books/details?id=WJL8NZc8lZ8C.

Rosenfield, L.K., Chang, D.S., 2009. The error of omission: a simple checklist approach for improving operating room safety. Plast. Reconstr. Surg. 123 (1), 399–402. https://doi.org/10.1097/PRS.0b013e318193472f.

Schlesinger, H., Venezian, E., 1986. Insurance markets with loss-prevention activity: profits, market structure, and consumer welfare. Rand J. Econ. 17 (2), 227–238. https://doi.org/10.2307/2555386.

Smith Thomas, L., 2011. Wind vulnerability assessment of roof systems and rooftop equipment for critical facilities: a preliminary protocol for design professionals. Struct. Congr. 2011, 1594–1605. https://doi.org/10.1061/41171(401)139.

Stanton, N.A., Hedge, A., Brookhuis, K., Salas, E., Hendrick, H.W., 2004. Situation awareness measurement and the situation awareness global assessment technique. In: Handbook of Human Factors and Ergonomics Methods. CRC Press, pp. 445–453.

https://www.taylorfrancis.com/books/e/9780429205712/chapters/10.1201/9780203489925-53.
Tokgoz, B.E., 2012. In: Gheorghe, A.V. (Ed.), Probabilistic Resilience Quantification and Visualization Building Performance to Hurricane Wind Speeds, [Phd, Old Dominion University. http://search.proquest.com/openview/7c3aa7738d30a6d692c887bef4626e98/1?pq-origsite=gscholar&cbl=18750&diss=y.
Unanwa, C.O., 1997. A Model for Probable Maximum Loss in Hurricanes [Texas Tech University. https://ttu-ir.tdl.org/ttu-ir/handle/2346/18641.
Unity (Version 2019), 2005. Windows. https://unity.com/.
Väisänen, T., Noponen, S., Latvala, O.-M., Kuusijärvi, J., 2018. Combining real-time risk visualization and anomaly detection. In: Proceedings of the 12th European Conference on Software Architecture: Companion Proceedings, 55. https://doi.org/10.1145/3241403.3241460, 1–55:7.
van Buuren, S., Groothuis-Oudshoorn, K., 2011. Mice: multivariate imputation by chained equations in R. J. Stat. Software 45 (3), 1–67. https://www.jstatsoft.org/v45/i03/.
Wickens, C.D., 2008. Multiple resources and mental workload. Hum. Factors 50 (3), 449–455. https://doi.org/10.1518/001872008X288394.
Wickens, C.D., Hollands, J.G., Banbury, S., Parasuraman, R., 2015. Engineering Psychology and Human Performance. Psychology Press. https://www.taylorfrancis.com/books/9781315665177.
Windexpo (n.d.). Retrieved October 18, 2019, from. http://www.cityofsitka.com/government/departments/building/Windexpo.html.
Zuk, T., Carpendale, S., 2006. Theoretical analysis of uncertainty visualizations. In Visualization and data analysis 2006 (Vol. 6060, p. 606007). International Society for Optics and Photonics (2006, January).