Final Report

An Experimental Study of the Effect of Transparency on Pilot Trust in Emergency Landing Planner

Prepared for

Joseph Lyons, PhD

Human Trust and Interaction Branch

711 HPW/RHXS

Human Effectiveness Directorate

Air Force Research Laboratory

Contract Number: FA8650-04-D-6405

Submitted by

Nhut Ho, PhD

NVH Human Systems Integration

December 7, 2014

REPORT DOCUMENTATION PAGE	*Form Approved* **OMB No. 0704-0188**

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
December 24, 2014	Final	11/2013 – 11/2014

4. TITLE AND SUBTITLE	**5a. CONTRACT NUMBER**
An Experimental Study of the Effect of Transparency on Pilot Trust in Emergency Landing Planner	FA8650-04-D-6405
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Nhut Ho Kolina Koltai Walter Johnson	**5e. TASK NUMBER**
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	**8. PERFORMING ORGANIZATION REPORT NUMBER**
Subcontractor: NVH Human Systems Integration 8938 Rhea Ave, Northridge, CA 91324 Collaborator: NASA Ames Research Center Moffett Field, CA 94035	

Prime Contractor: SRA International, Inc. SRA C4ISR Center 5000 Springfield Street, Suite 200 Dayton, Ohio 45431	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory 711 Human Performance Wing Human Effectiveness Directorate 2255 H Street Wright-Patterson AFB, Ohio 45433	**10. SPONSOR/MONITOR'S ACRONYM(S)** AFRL/RHXS **11. SPONSORING/MONITORING AGENCY REPORT NUMBER**

12. DISTRIBUTION AVAILABILITY STATEMENT

Distribution Statement A: Approved for public release. Distribution is unlimited?

13. SUPPLEMENTARY NOTES

14. ABSTRACT This experimental study examined the effects of transparency (operationalized as increasing levels of explanation) on pilot trust of an automated emergency landing planner. A low-fidelity study was conducted where commercial pilots (N=12) interacted with simulated recommendations from NASA's Emergency Landing Planner (ELP). These recommendations varied in their associated levels of transparency. Results indicated that trust in the ELP was influenced by the level of transparency within the human-machine interface of the ELP.

15. KEYWORDS

Transparency, Trust in Automation, Reliance, Emergency Landing Planner, Risk, Commercial Aviation, Human-Machine Interface

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPONE NUMBER (*Include area code)*
UNCLASS	UNCLASS	UNCLASS	none	79	

This page left blank intentionally.

Table of Contents

Lists of figures and tables

List of figures

List of tables

1. Introduction

1.1 Trust and Transparency

Trust is believed to play a critical role in shaping how humans interact with, rely upon, and accept technology (Chen & Barnes, 2014; Lee & See, 2004; Lyons & Stokes, 2012). Designing for appropriate trust and reliance is a particularly important consideration because inappropriate reliance can lead to potential mishaps. Common examples of inappropriate reliance on automation includes misuse and disuse (Parasuraman & Riley, 1997). Misuse refers to using and relying upon automation under conditions, or for purposes other than those for which the automation was designed; while disuse refers to failing or refusing to use automation where that automation is appropriate and useful. Disuse undermines the potential strength and benefits of automation, while misuse can lead to risky and dangerous situations. In order to avoid misuse or disuse of automation, designers of automation should then seek methods that support appropriate trust to avoid misuse or disuse of automation; this is also known as trust calibration, where a person's trust in, and reliance upon, the automation corresponds to the automation's capabilities (Lee & See, 2004; Muir, 1987).

Designing for appropriate trust and reliance is a complex task. Extensive research has shown that reliance on automation depends on many factors, including the trust that the operator places on the automation and the capability and complexity of the automation (Lee & See, 2004). As automated systems continue to grow in complexity one design consideration in particular, the system's level of transparency, has increasing impact on operator reliance. Transparency broadly, represents a process for establishing shared awareness and shared intent between a human and machine (Lyons, 2013), and the importance of making a system transparent increases as the system increases in complexity. While a need for transparency can be addressed through training, this approach has limited efficacy for complex systems because of the time and effort that must be put into initial and recurrent training. A good example of this problem can be found with the flight management system (FMS) in modern day transports. Here pilot training has only partially solved the problem of transparency with pilots continuing to puzzle over the how and why of the FMS operation - "Why did it do that?" "What is it doing now?" and "What will it do next?" (Abbott, 1996). Clearly this issue of highly complex systems with low transparency can result in confusion and operator errors. In particular, there are cases of aviation mishaps that have been the result of low transparency due to the operator not understanding the system (Billings, 1996). Real life incidents, like the US Airways 1549 Hudson River landing and the Asiana 214 San Francisco crash, highlight the importance of transparency of automated systems in calibrating trust (NTSB, 2009, 2013). From what we know about creating trust in automated systems, a key factor is that an operator must understand and not be confused by the system (Lee & See, 2004). To be able to begin to trust the system, the operator must be able to effectively use and operate the technology. However, as humans we tend to be suboptimal with our reliance strategies involving technology (Lee & See, 2004; Lyons & Stokes, 2012), resulting in potential errors using a technology, choosing to use an error-prone tool, or by failing to use a potentially beneficial technology.

The current report will examine the role of transparency in mediating trust in automation through a low-fidelity study of transparency in the domain of automated tools for the commercial aviation. This low fidelity study is the starting point in a series of studies aimed at exploring this relationship between trust, reliance, and transparency that can later generate design guidelines that promote appropriate trust and reliance on automated systems.

1.2 The Emergency Landing Planner

The National Aeronautics and Space Administration (NASA) has developed tools to support emergency procedures in commercial aviation. One such automated tool is the Emergency Landing Planner (ELP) (Meuleau, Plaunt, & Smith, 2008). The ELP was designed to support real time analysis of complex situations (e.g., damage to the aircraft, adverse weather) and recommend a safe route, approach, and landing runway to pilots. Making a diversion decision, particularly in cases where the plane has been damaged, can be very difficult (Meuleau, Neukom, Plaunt, Smith, & Smith, 2013). Historical cases have shown that it can be difficult for a pilot to plan an effective and safe 3D path because the planning requires optimization of objectives that involve a large number of dynamic factors (Meuleau et al., 2008), and that pilot intuitions are prone to biases and thus are not always correct in cases of damaged aircraft. The ELP was developed to assist pilots in the task of choosing a diversion path and runway in the event of an emergency. The ELP uses a collection of information sources to find usable long range roadmap solutions in a 3D environment and generate prioritized recommendations. The ELP determines the priority of the recommendations by assessing the risk of various flight stages (enroute, approach, and landing) and conducting extensive analysis of many elements that would be challenging for a pilot to consider during an emergency situation. For instance, when an aircraft sustains damage or experiences equipment failure, the ELP estimates the changes in the aircraft flight dynamics and generate and recommend control actions (e.g., gentler turns) that stay within the climb/descent limitations of the aircraft. This helps reduce the workload of the pilot and increases her/his effectiveness in choosing a safe diversion.

However, like any automated tool, commercial pilots may not rely optimally on the ELP and its use may lead to other unintended consequences (Meuleau et al., 2013) as demonstrated in a study by Meuleau et al. (2013), which compared pilots' use of the ELP to other decision aids in several scenarios that varied in damage to the aircraft, severity of weather, and location. In this study, pilots were given Navigation displays (Figure 1) and a pair of keypads and control display units (CDUs) (Figure 2 & Figure 3) to access information generated by the ELP. In scenarios where the weather was very poor, the ELP generally led to quicker and better decisions. Pilots reported that they preferred having the ELP in all scenarios and that it reduced their workload. An observation that emerged from this study is that many pilots preferred long runways over short runways, even when the shorter runways had the better weather and winds. Pilots who chose suboptimal diversions that have longer runways tended to have fatal crashes in the simulator. This result suggested that the pilots may not have chosen recommendations with high risk of failure if they knew about the risk level, and that improvements can be made to the displays shown in Figures 2 and 3 to make the risk information more transparent. Because the ELP is a recommender system, it can be used to easily measure pilot trust in the recommendations and, thus, is a good platform to study effects of added transparency on user trust.

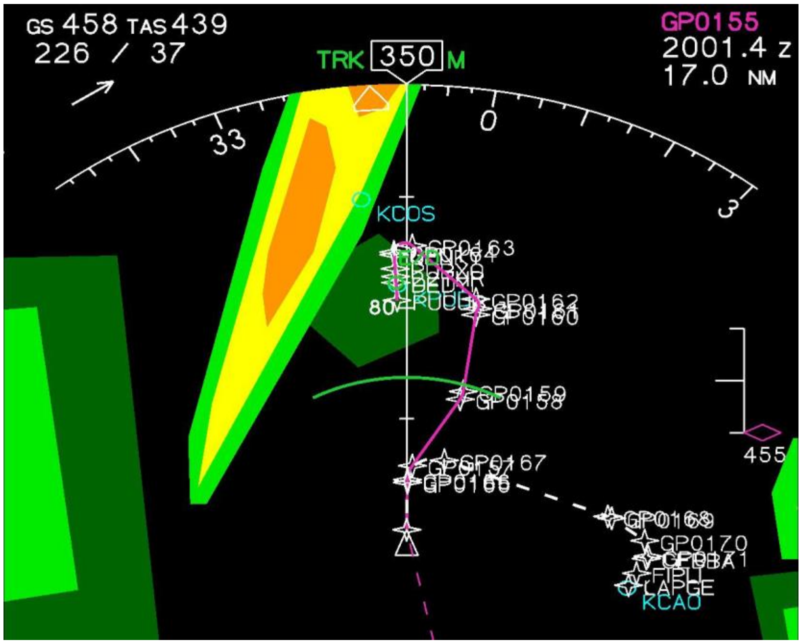


Figure 1: The navigation display showing both the current route (magenta) and the new route being considered (dashed white). Green, yellow, and orange areas indicate rain and thunderstorm activity

Figure 2: The first five diversion recommendations for a scenario displayed on the CDU

Figure 3: An airport information pages showing runways and current weather for KCAO

2. Analysis of Transparency Issues for Emergency Landing Procedures

In order to study the effects of transparency on pilot trust in the ELP, it was necessary to first understand pilot goals, information requirements, and the various factors that pilots typically take into consideration when managing a diversion. In order to obtain this information we conducted a task analysis and interviewed subject matter experts (SMEs) whose experience ranged across military and civil transport operations. Task analysis is the process of identifying and examining tasks that must be performed by the user(s) when they interact with systems (Kirwan & Ainsworth, 1992). The task analysis process is designed to help reveal the mental models of pilots, as defined by Ososky, Sanders, Jentsch, Hancock, & Chen (2014) where they state "Mental models represent the underlying, organizing framework for human understanding of robots, serving as the lens through which humans interpret and reason about a robot's actions, abilities, and usefulness. Thus, mental models provide for the activation of situation awareness in humans". To develop the task analysis, existing literature, consultations with SMEs, and interviews with current and retired commercial aircraft pilots (users) were all used to compile information about pilot mental models. Using Endsley's levels and requirements of situation awareness (Endsley & Jones, 2011) as a starting point, the cognitive tasks involved in diversion decision were mapped into the chart shown in Figure 4. Each subsequent task was then broken down into sub-tasks, and finally into SA requirements for each of those sub-tasks. Figure 5 shows an example of the SA requirements for each subtask under task 4.0 "Asses the quality of the alternative runways". SMEs were then asked to review and revise the task chart through a number of iterations until final iteration of the task chart was generated. The task chart was used as the basis to develop the experiment presented in the next section.

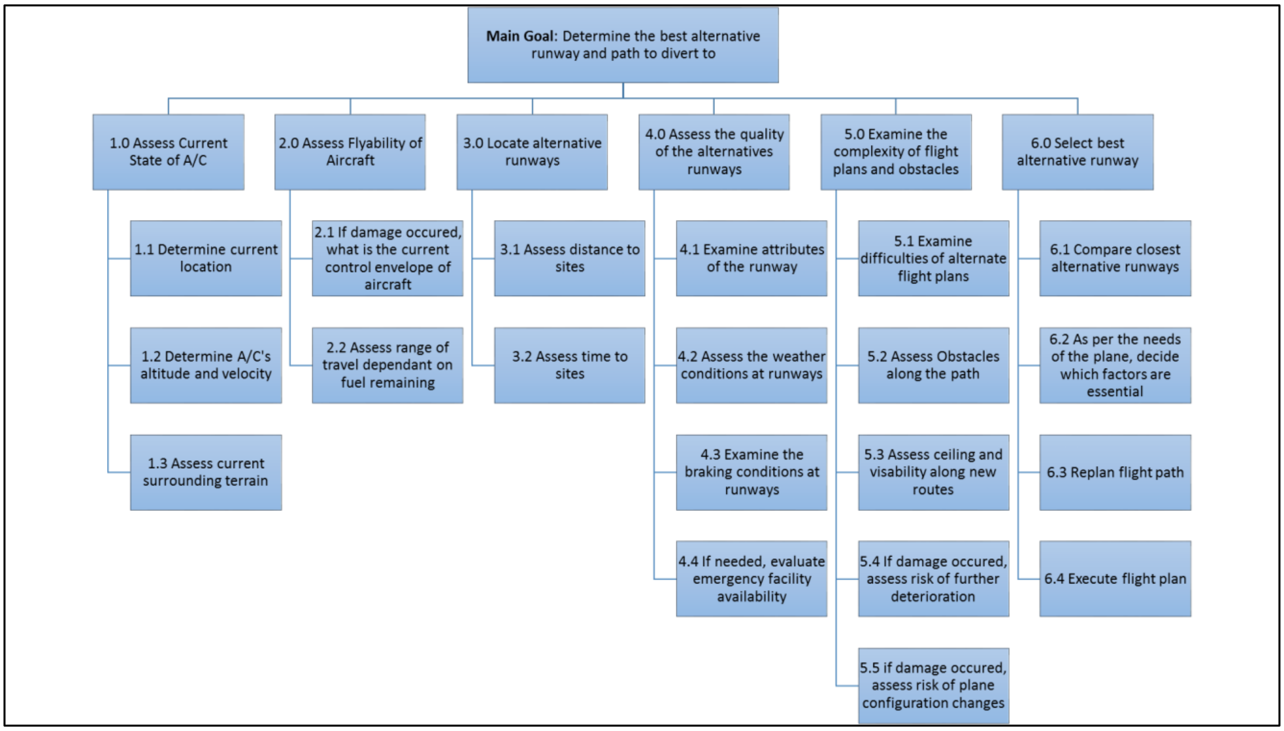


Figure 4: Task and subtask breakdown of a diversion decision

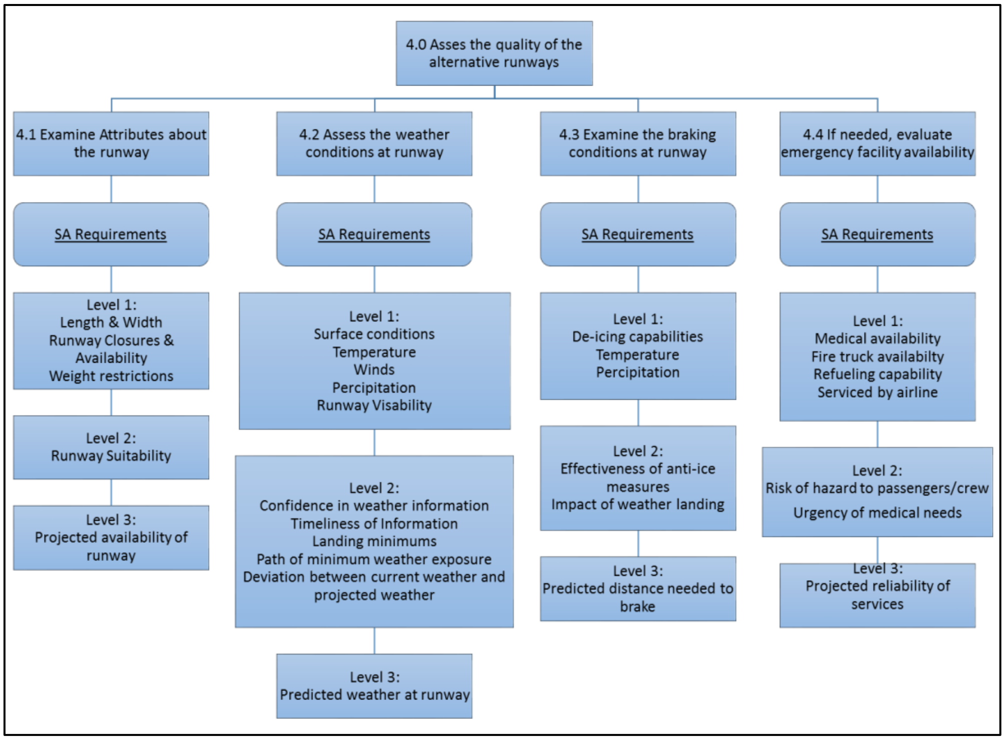


Figure 5: SA requirements for the subtasks of task 4.0, Assess the quality of the alternative runways

3. Methods, materials, and procedures

3.1 Objectives

The objective of this study was to explore how automation transparency mediates trust in automation. In the context of emergency landing operations involving the use of ELP, the following questions were of particular interest:

- How does the type of explanation (transparency) for ELP's recommendations affect trust calibration?
- How easy it is for participants to understand and evaluate the ELP's recommendations?
- To what extent did the alignment of the participant's decisions with the recommendations of the ELP vary as a function of transparency?
- How does the type of explanations affect participants' confidence in the ELP and their own decisions?
- How does the explanation type affect the workload of the participant?

3.2 Study design

3.2.1 IVs

The independent variables (IVs) for the experiment were **System Transparency** and **Scenario Information**. The study was a 3 x 2 fully within subject design. The three **System Transparency** conditions were: *Control*, *Value*, and *Logic*. The *Control* condition provided no feedback or evaluation from the ELP, but just provided a basic set of information which included all information that the ELP would normally use in its risk evaluations, as well as some other potentially pertinent information not utilized by the ELP. The *Control* condition was used as a baseline to compare to the other two conditions. The *Value* condition provided a risk evaluation number (aka a value) in addition to the same information received in the *Control* condition. The *Logic* condition provided the logic or reasoning behind the risk statement/value in addition the same information received in the *Value* conditions. The two **Scenario Information** conditions used were *Equal Information* (*EI*) or *Unequal Information* (*UI*). In *EI* scenarios, participants were given the same information that the ELP has in making a diversion decision, whereas in the *UI* scenarios, participants were given information that the ELP does not know in making a diversion decision. These conditions of the two IVs are further elaborated in the next two sections.

3.2.1.1 *System Transparency* Conditions

The *Control* condition served as the baseline for the amount of information given to the participant. In this condition, the participants received a list of diversion recommendations in a randomized order in order to prevent them from easily assessing what the ELP thought the best diversion was. An example of a diversion recommendation is shown in

Figure 6. The *Control* condition provided the same information the ELP would use to evaluate potential diversion paths and runways. This included the ATIS (Automatic Terminal Information Service) report for the airport, as well as information about the enroute weather and surrounding terrain. The ATIS report was comprised of decoded METAR (routine weather report provided at fixed intervals) information about an airport, runway number, type of approach, runway length, distance, and the bearing. Airport facility information was provided as well but the ELP did not take this factor into account when evaluating diversions. The airport facility information included the medical facilities, maintenance capability, airline support, passenger conveniences, and refueling resources available at the airport. Each airport facility was ranked and color coded based on the likelihood of that facility meeting the need of the diversion. Airport facilities were ranked as excellent (blue) if there was 100% likelihood of satisfying that need, good (green) represented 80% likelihood, fair (yellow) represented 50% likelihood, poor (orange) represented 25% likelihood, and unavailable (red) represented 0% likelihood. Medical facilities were based on the medical facilities in the surrounding area and at the airport. Maintenance capability was determined by the aircraft maintenance facilities available at the airport. Airline support was determined by the amount of support the pilot's airline company had at the airport. Passenger conveniences were based on the convenience for passengers such as their access to transportation, lodging, and services at the airport and in the area. Refueling resources were based on the availability of the required fuel & refueling support for their aircraft at the airport. Airport facilities are not taken into consideration when the ELP generates diversion options, but were additional information provided to the participants.

Participants also received a map that showed weather observances (e.g., thunderstorms, snow, etc.) with an ELP generated flight path for each recommended diversion. These were images created through Google Maps. Although the ELP generates a 3D flight path, participants were only shown a 2D image of the path from above.

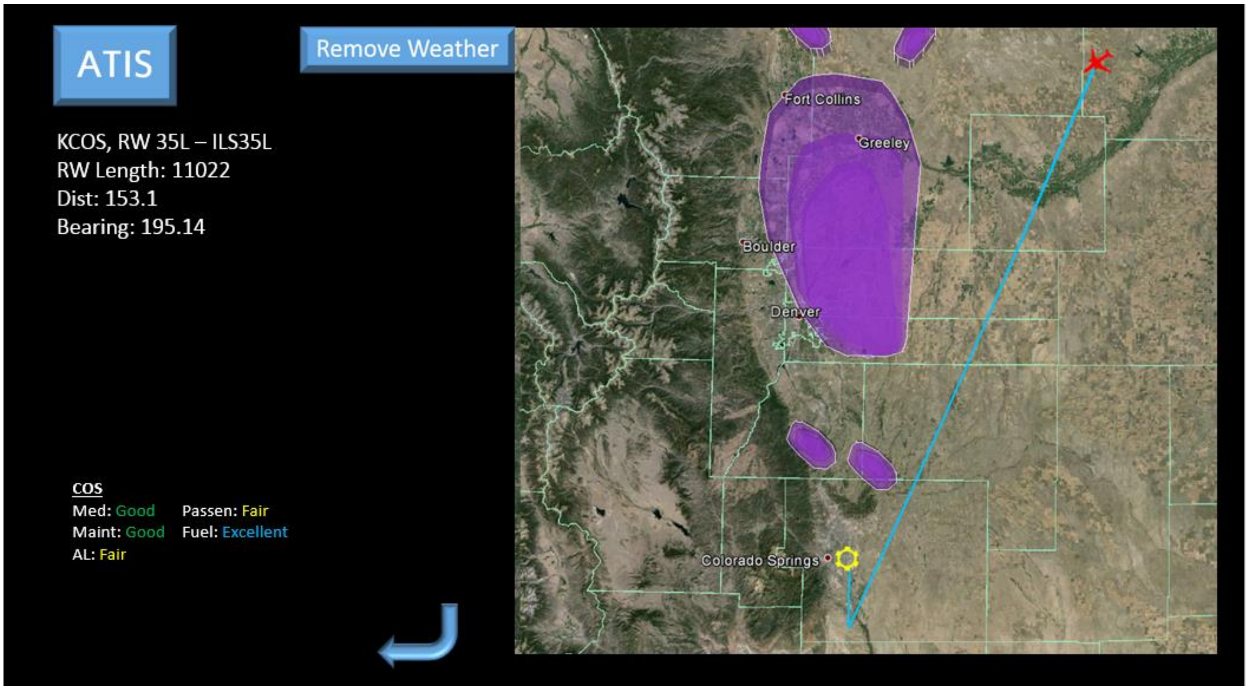


Figure 6: Example of a diversion recommendation in the Control condition

The *Value* condition included the same information provided in the *Control* condition, plus the addition of an overall risk statement and number, which was obtained by taking into account the enroute, approach, and runway segments for each diversion (

Figure 7). The ELP determined the risk number by first evaluating: 1) the enroute distance, turns in proposed diversion path, enroute weather, and the terrain, 2) the approach ceiling, approach weather, and visibility, and 3) the runway length, runway width, surface of the runway, required landing speed, and wind conditions. Internally, the ELP generates numerical risk percentages for the enroute, approach, and runway segments as well as for each individual factor. The ELP then takes all of the individual risk factors and their interactions into consideration when calculating the final overall risk value. The value in the risk statement represents the likelihood of successfully completing the approach and landing on the first attempt under the current conditions (e.g., "You have a 34% chance that you will be able to successfully complete the approach and landing under current conditions.") Note however, that in this study, a 34% success rate does not translate into a 66% chance of crashing, it just means that the probability of landing successfully on your first attempt is 34%. This was an important nuance that was explained to the pilots during their training with the ELP and its interface. The values were color coded based on the percentage of success and the percentage corresponded to a ranking: 100% to 90% was considered excellent (blue), 89%- 76% was considered good (green), 75%-56% was considered fair (yellow), 55%-41% was considered poor (orange), while 40% and below was unacceptable (red).

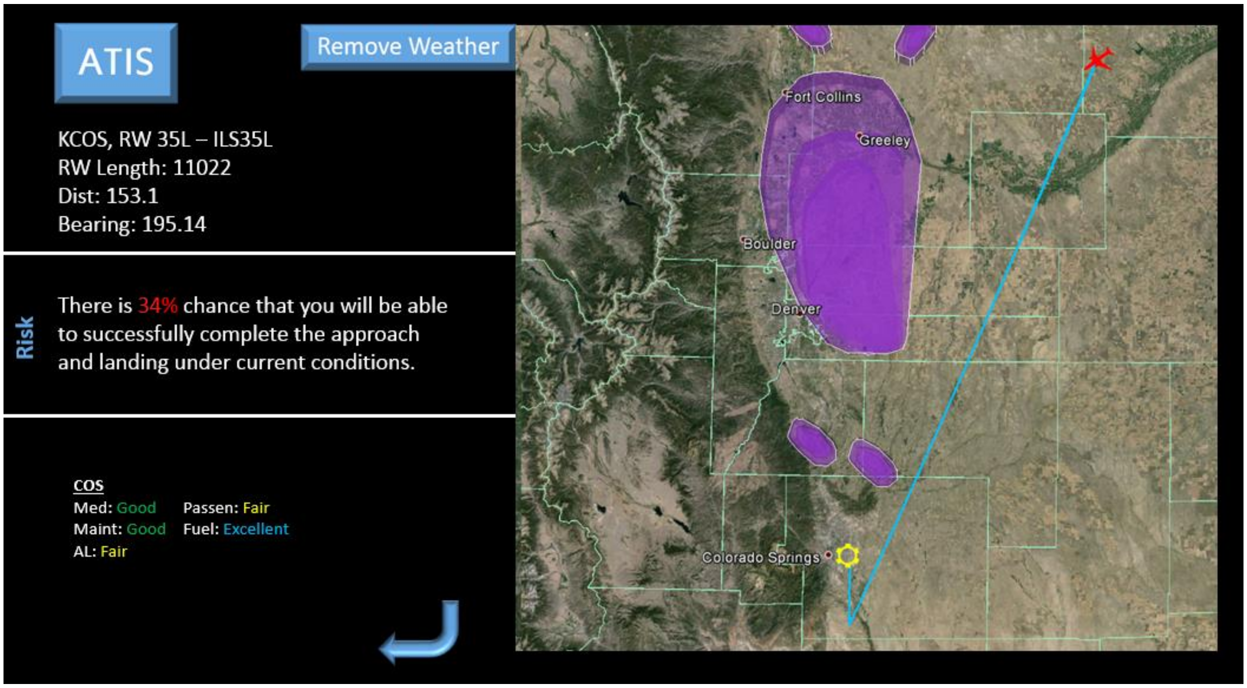


Figure 7: Example of a diversion recommendation in the Value condition

Finally, the *Logic* condition included all the information in the *Value* condition in addition to summary statements of the logic, or reasoning, behind the risk statement for each diversion option (Figure 8). The *Logic* condition displayed the component risks to enroute, approach and landing segments, and provided a summary logic statement for each of these. The logic statements were generated by examining the enroute, approach, and runway risks for each diversion option. Using those numbers, a logic was provided to the participants (e.g., Runway: Unacceptable – The landing crosswind is too high for a safe landing). In the *Logic* condition, the participants were given the categorical values (excellent, good, fair, poor, and unacceptable) for the enroute, approach, and runway and a statement for why the automation made that assessment. The values were color coded using the same scale as the risk statement (see above).

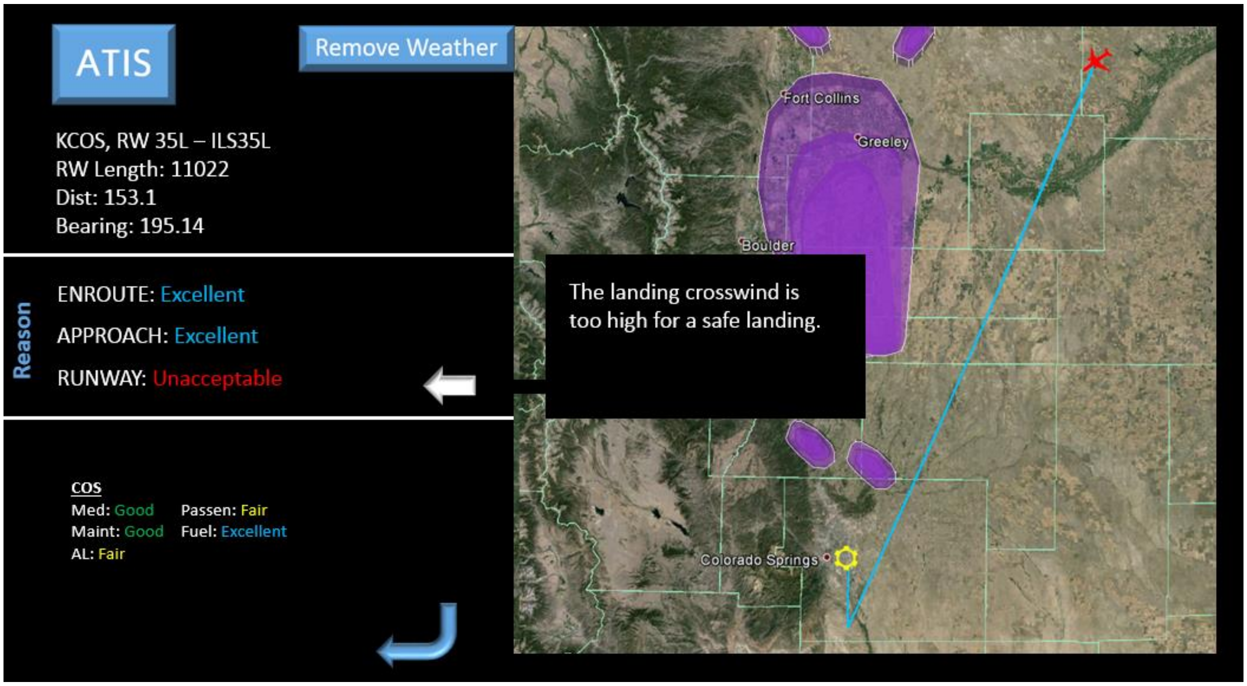


Figure 8: Example of a diversion recommendation in the Logic condition

3.2.1.2 *Scenario Information*

Six different scenarios were adapted from a previous study conducted by the Flight Deck Display Research Laboratory (FDDRL) of NASA Ames Research Center (Lachter, Battiste, et al., 2014; Lachter, Brandt, et al., 2014). These scenarios required the pilot to divert from the current flight plan (due to either weather, medical, or maintenance issues) and select an alternative airport to land. Flights were either arriving or departing from the Denver area. There was a total of six airports used: Denver, Cheyenne, Colorado Springs, Eagle, Pueblo, and Grand Junction. These scenarios and simulated weather were designed to be different enough from each other and provide a range of diversion situations. For this study, these scenarios were slightly modified to decrease the number of diversion decisions a pilot would have to make in a single scenario from 2-3 to 1, and the initial position and bearing of the plane. These scenarios were created by SMEs of the FDDRL and have already been validated to consistently have a high difficulty while not being too similar to one other.

There were two **Scenario Information** conditions: equal information (*EI*) and unequal information (*UI*). In *EI* scenarios, participants and the ELP have access to the same, or equal, relevant information necessary for evaluating a diversion option. For example, in a case of a weather diversion, the pilot has access to the exact same weather information that the ELP is using to evaluate the six diversion options. In *UI* scenarios, the pilot has access to information that the ELP does not have. For example, in one *UI* scenario, the reason for diversion is a medical emergency with a passenger. The participant is given this information but the ELP is unaware of the state of the passengers. In this scenario, the ELP does not provide the best recommendation because it does not have all the information that the pilot has. This kind of scenario facilitates the investigation of appropriate trust because it involves situations where the

ELP is not able to generate the best diversion. Three *EI* and three *UI* scenarios used for this study are shown in
Table 1 below.

Flight Plan	Description of scenario	Scenario Type
ORD to DEN	DEN closed for microburst (NTD - Weather)	*EI*
PHX to COS	COS ATIS below mins (NTD - Weather)	*EI*
SFO to EGE	EGE closed due to disabled aircraft (NTD - Obstacle)	*EI*
DEN to SFO	Medical Emergency, a NO LAND 3 at nearest airport DEN (NTD - Medical)	*UI*
DEN to SEA	Shattered non-structural windshield, nearest airport DEN closed for snow (NTD - Maintenance)	*UI*
DEN to DFW	Dispatched w/inop left pack. Right pack fails, depressurize. Nearest airport DEN closed for snow (NTD - Maintenance)	*UI*

Table 1: Description and types of scenarios used. NTD = Need to Divert.

3.2.2 DVs and Instruments

Multiple dependent variables were collected in this study. The main research question focused on the effect of transparency on trust; however, several other dependent variables of interest were also collected in the study, including agreement with the ELP, confidence, workload, opinions on the ELP, and strategies for decision making. Agreement with the ELP covered two separate aspects: agreement with the evaluation of the diversion (e.g., how safe is the diversion), and agreement with the ranking of the diversions (e.g., this is the best, second best, etc. diversion). In regards to confidence, both the subject's confidence in their choices and their confidence in the ELP were measured. Workload was measured to determine whether transparency affected the pilot's mental working capacity when making a diversion decision. Opinions and feedback on the tools and information provided to the participants were also collected in the different **System Transparency** conditions. Finally, participants' strategies for decision making were captured to understand pilot preferences and reasoning for choosing a certain diversion option, and to verify and substantiate the task analysis results described in section 2.1.

3.2.2.1 Task Worksheet

Participants were required to complete a total of six task worksheets, one after each scenario. Task worksheets were in paper format and were distributed to participants prior to the start of each scenario. Each worksheet (an example is given in Appendix A) gave participants two specific tasks. First, participants had to evaluate the safety of each diversion. When evaluating diversion safety, participants were repeatedly told not to take into account the reason for diversion or the facilities available at the airport. They were told to evaluate the diversion on safety alone, which referred to the safety of the enroute, arrival and landing segments. Thus, they were instructed to use the same information that the ELP had in evaluating the diversions. They were instructed to categorize the safety of the diversion as excellent, good, fair, poor, or unacceptable and also indicate if they would be willing to fly the diversion. Participants were provided a scale as a guide to categorize the diversions. The scale was the same one used for the risk statement and logic categorization (see section 3.2.1). There was no requirement for how many times each rating was used. More than one division path could be marked as excellent, or no diversion path could be marked as excellent.

The second portion of the worksheet required participants to rank the diversions from their 1st choice to their 5th choice. For this task, they were instructed to use all of the information provided to them when making their decisions, including the reason for diversion. It was made clear to the participants that this ranking differs from the previous ranking task because the safest diversion may not be the best option. For example, if the diversion was for a medical emergency, the safest diversion option may not have adequate medical facilities to address medical needs, whereas another diversion option may be both safe and have medical facilities.

3.2.2.2 Surveys

Four different surveys were administered in the study: demographics, post-trial, post condition, and debriefing. All surveys were administered using Qualtrics as the online host. Each participant completed the demographics survey (Appendix E) before starting the study. The demographics survey collects the participants' current employment status as a pilot, their experience with different aircraft, and past flying experiences.

Post-trial surveys (Appendix B) included questions on workload, confidence, ratings of usefulness, and participant feedback and opinions after each scenario. A modified overall subjective workload scale ranging from 1 to 10 was used (Hill, et. al, 1992). Self-reported confidence was measured on a Likert scale ranging from 1 (no confidence) to – 7 (extremely confident). Four separate post-trial confidence measures were gathered from the participants. These were confidence that: 1) the ratings they gave the diversions were accurate, 2) their choice of the best (top-ranked) diversion was accurate, 3) their choice of the worst (lowest-ranked) diversion was accurate, and 4) the diversions that the ELP gave them were appropriate. The pilots were also asked to rate (on a 1-5 scale, with 1 being very important, and 5 being very unimportant) the importance of the information on the weather at the runway, the approach weather, enroute weather, ceiling and visibility, runway characteristics, approach plates, the risk statement, the reasoning (logic), flight path depiction, distance, fuel, difficulty of the route and landing, and airport facilities. Each participant completed a total of six post-trial surveys (Appendix B).

Once participants completed the two trials in a condition, they were issued the post-condition survey (Appendix C). The post-condition survey was designed to measure participants' trust in the ELP and self-reported workload. Trust was measured by having participants rate 8 statements on a Likert scale ranging from 1 (strongly disagree) to – 7 (strongly agree). Example statements include, "I would feel comfortable relying on the recommendations of the ELP in the future.", "When the task was hard, I felt like I could depend on the ELP.", and "I would be comfortable allowing the system to make a diversion decision for me." . The sum of the ratings was used as the measure of trust. Workload was measured in the same manner as in the post-trial survey.

The debriefing survey (Appendix D) was administered after participants had completed all the scenarios. The debriefing survey included a series of questions on how helpful the tools were, opinions of the ELP, and the ranking of the importance of factors in the decision making process. The pilots were also asked to rate the helpfulness of the ELP's information on the ATIS information (i.e., real-time weather information), approach information, the logic and value statements provided in the *Logic* and *Value* conditions, respectively, airport ratings, the list of diversions, individual diversion or runway information, airport facilities descriptions (e.g., emergency vehicle information), and the diversion path.

3.2.2.3 Interview, observation, and notes

Observational data was also collected. Camtasia screen and audio recording software was used to record the participant's audio, movements, and their screen activity. A webcam was set-up next to the computer monitor to capture what the participant was doing while maneuvering through the trial. As part of the study, participants were asked to speak out loud during their decision making process. Participants were instructed to voice what they thought about during the trial, particularly what they liked or didn't like about a diversion. A researcher was continually present in the room with the participant in order to document this thought process and any comments and the rationales participants made while completing the task worksheet. After the completion of all the scenarios and surveys, an interview was conducted to pose follow-up questions regarding the participants' answers on the task worksheets. A researcher would review the scenarios and task worksheet with the participant to get further clarification about their reasoning for their evaluations and rankings of diversions. These addressed what their thought process was when they evaluated and/or ranked diversion paths differently than the ELP. Furthermore, the researcher questioned each participant in order to determine any heuristic, or rule of thumb, they might have used when making a diversion decision.

3.2.2.4 Technical development

Instead of using a dynamic aircraft simulator, PowerPoint slides were created to present the scenarios and diversion recommendations to the participants. Participants were able to navigate through the PowerPoints using hyperlink buttons in a manner similar to how they would interact with the ELP through the CDU. To create the slides, the scenarios described in section 3.2.1.2 were input into the ELP to generate recommendations. Using the outputs generated by the ELP, a unique set of PowerPoint slides were created for each scenario in each **System Transparency** condition.

3.3 Procedures

After the introduction and consent forms were completed, participants went through training. A training manual was given to the participant to use throughout the experiment (Appendix F). Training included an introduction to, and familiarization with, the ELP, including the information and factors the ELP takes into consideration when generating the list of diversion options. The participant was also trained on how the ELP would be used for the current study. For example, normally, the ELP generates and lists the diversion options for a damaged aircraft in a ranked order and the pilot can update the ELP to account for the changing location, altitude, and velocity of the aircraft. However, for the current study, the participant was informed that 1) the airworthiness of the aircraft was not affected, 2) the list of diversions would be randomized, and 3) there would be no opportunity, or need, to update the ELP since they were examining the evaluations for a discrete moment in time – not during a dynamically unfolding flight.

Training also included a summary of the tasks and objectives of the study. Because participants may have been familiar with the airports being used, they were instructed to only use the information provided to them and disregard their knowledge of the airports. They were also instructed that the provided information about the available airport facilities would differ between scenarios. The participants were not given a time limit, but were asked to complete each task to the best of their ability in a reasonable amount of time.

After the participant received his or her orientation, they were escorted to a computer station where they completed a demographics survey and then completed six scenarios. The six scenarios were divided into three sets of two scenarios each, with each set utilizing one of the three **System Transparency** conditions (*Control*, *Value*, and *Logic*). Within each set the participant might receive two *UI*, two *EI*, or a *UI* an *EI* scenario. The order of the scenarios and the conditions were counterbalanced to mitigate learning and/or ordering effects. At the beginning of each of the sets, participants received training on the tools they would be using for that particular **System Transparency** condition. After training was concluded, the researcher asked the participant if they had any questions or needed any additional clarification. Once any additional questions and clarifications were addressed, the researcher provided the necessary paperwork for the scenario and the participant begin the first scenario of that System Transparency condition. Participants were given a task worksheet (plus scratch paper for notes) for each scenario. Participants were supplied with a binder of approach plates for all potential diversion options for reference. For an overview on the procedure of this study, refer to Figure 9.

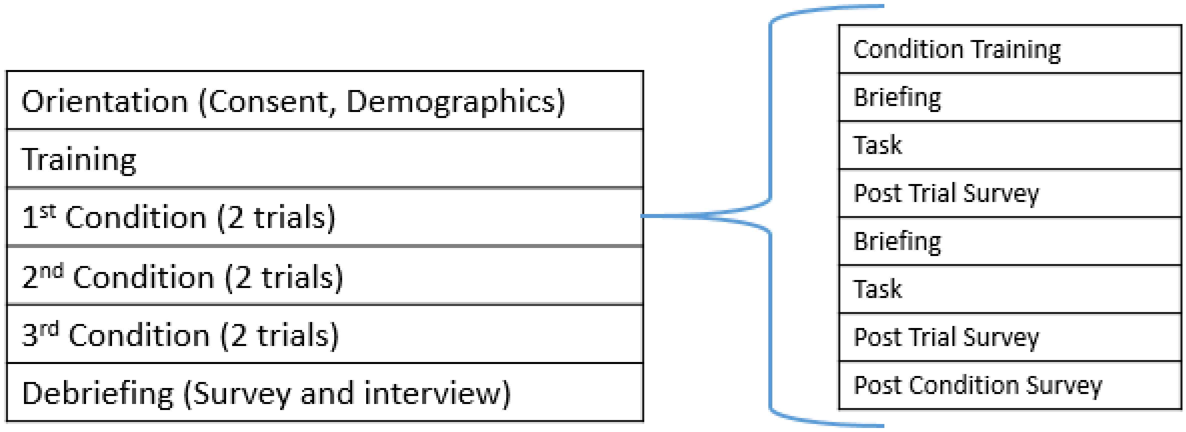


Figure 9: Overview of the procedure of the study

Each trial began with a briefing of the scenario to help bring up the participant's situational awareness. The briefing included information on the aircraft type, the departing or destination airport, the distance to or from the destination, the ATIS report of the departing/arrival airport, and indication of sufficient fuel. A map was also provided to help the participant orientate the location (Figure 10) and included information on the aircraft's current position and heading, the current weather in the area (depicted in purple), and the airports (indicated as yellow cogs) in the surrounding area. Participants were also given a hard copy of the briefing in case they needed to refer to it later on in the scenario.

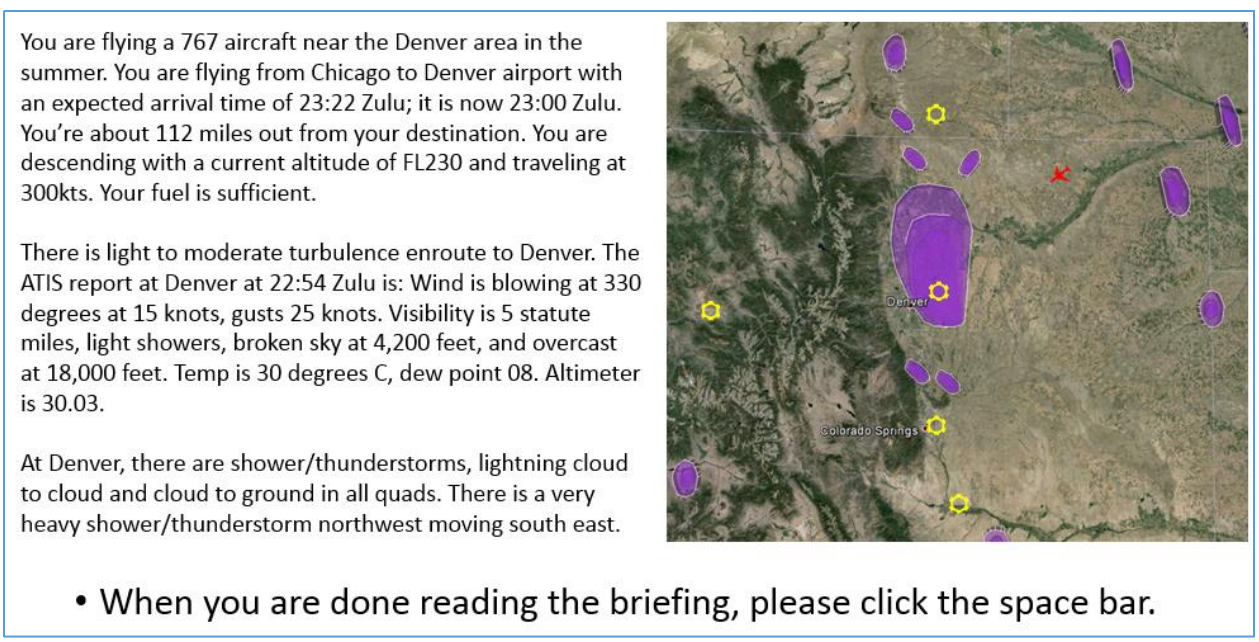


Figure 10: Example of the briefing page in scenario

After the briefing, the participant would click on the space bar and receive an alert for diversion as shown in Figure 11. As shown, underneath the alert and reason for diversion was an icon labeled, "Start ELP." The page also included a map that was identical to the map on the briefing page. The participant was required to click the icon to move to the next page.

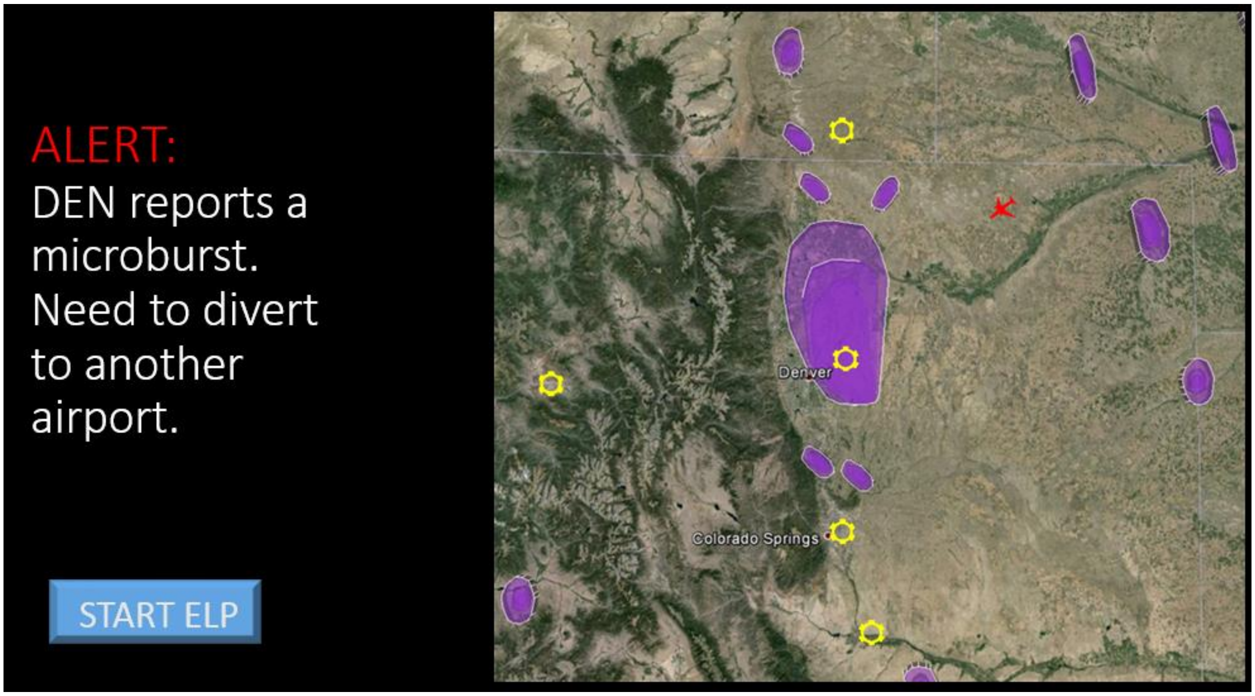


Figure 11: Example of a diversion alert in a scenario

The next page was the main summary page which included a list of diversion options (randomly ordered), a map, and an end scenario button (Figure 12). The map displayed the weather observed by ground radar (with an option to hide or show the weather obstacles), current location and heading, potential diversion airports, and suggested diversion paths. A blue line would indicate a proposed diversion path. Under normal circumstances the ELP would have rank-ordered the diversions on the left side of the screen. However, since one of the participant's tasks was to rank-order the diversions, the diversions were shown in a randomized order on the left hand side of the screen. Under the list of diversions was a red button that would end the trial when clicked.

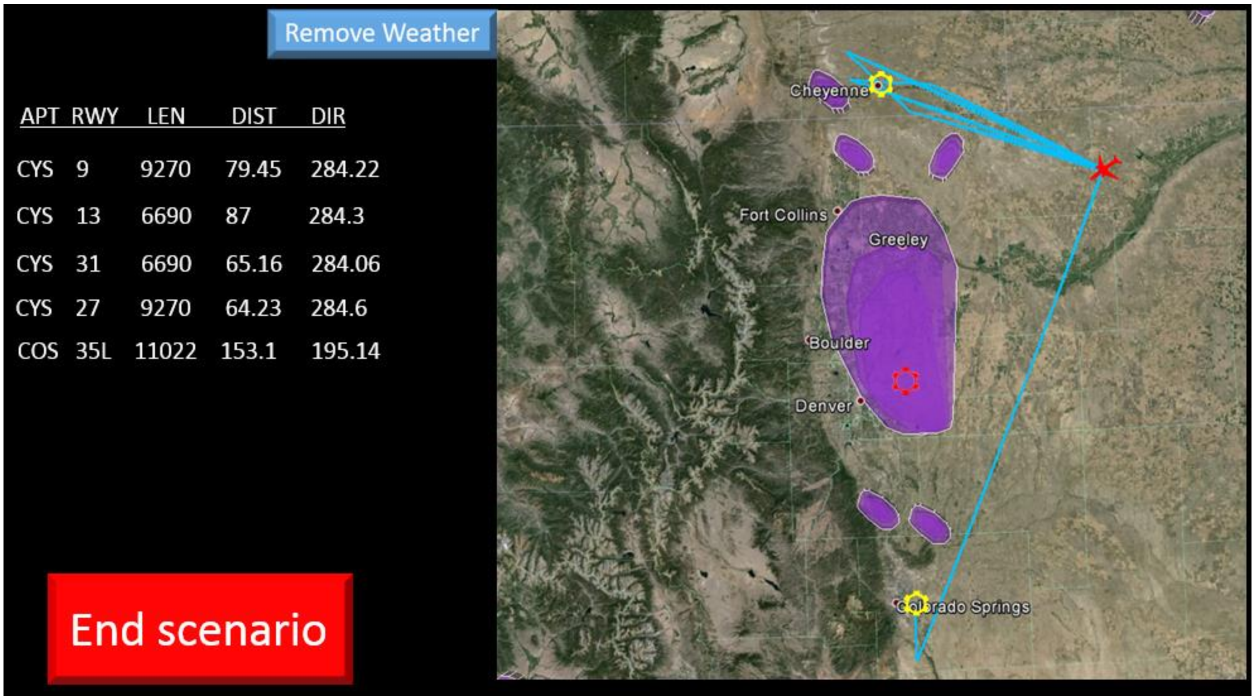


Figure 12: Example of a main summary page in a scenario

Participants could click on any of the diversion options to be directed to an individual diversion page that included additional information regarding that diversion option. The amount and type of information provided depend on the current System Transparency condition. For the *Control* condition, participants were given the ATIS report, runway information, and airport facility rankings (

Figure 6). Each individual summary page also included a map with the recommended flight path to that airport and runway. Participants always had the option to show or remove the weather obstacles from the map. The individual summary page for the *Value* condition includes the same information as in the *Control* condition with the addition of the risk statement (

Figure 7). The individual summary page for the *Logic* condition included the same information as in the *Value* condition, but with an additional option to view the logic behind the risk statement (Figure 8).

Participants were given a task worksheet to complete during the scenario. They were encouraged to review each diversion option and use only the information provided to them (e.g., not to rely on previous knowledge about the airports or runways) while completing the task worksheet. Participants were also asked to speak aloud while they went through the diversions and explain their decision making process. They were instructed to indicate what diversion they liked/disliked and the rationale. Their screen activity, audio, and movement were recorded using Camtasia screen and audio recording software. A researcher was also present in the room to answer any questions, document observations, and to verify that participants were following instructions.

After participants completed their task worksheet, they clicked the end scenario button, which lead them to a screen that notifies them that they had reached the end of the trial. They were instructed to notify the experimenter and click on a link that would direct them to the post-trial survey. Recording continued while the participant completed the survey and was suspended after all the required surveys were completed. After every two scenarios, participants were required to also complete a post-condition survey. To reduce fatigue, participants were asked to take a break after every two scenarios.

After completing all the scenario, a debriefing survey was administered. Finally, an interview with the participant was conducted. The participants were asked questions regarding their evaluations and rankings of the diversion options for each scenario. If participant’s rankings were drastically different from the ELP’s recommendations, they were asked to explain their reasoning and give clarification for why they evaluated or ranked a diversion as such. To unveil the decision making process, all participants were also asked if they had a heuristic or rule of thumb when making a diversion decision in an off-nominal situation in real life.

3.4 Participants

The participants were recruited by the San Jose State University Research Foundation (SJSURF), and consisted of 12 commercial pilots (N = 12). The study required participants to be either commercial airline pilots or student commercial pilots. To be eligible for recruitment, pilots were required to be at least 18 years old and have commercial airline experience as either a captain (83.3%) or first officer (16.6%). There were no requirements for which airline company participants were employed by, and pilots could be cargo or passenger airline pilots. It was also requested that pilots have glass cockpit experience with flight management systems (FMS). Pilots could be either active (83.3%) or retired (16.6%) as long as their retirement does not exceed three years. Pilots’ experience in their current position ranged from 1-5 years (25%), 5-10 years (16.6%), and over 10 years (58.3%). Participants’ total hours flown varied from 5001-10,000 hours (41.6%) to over 10,000 hours (58.3%). The distribution of total hours in a Boeing glass cockpit was: 1-1000 hrs. (16.6%), 1001-3000 hrs. (16.6%), 3001-5000 hrs. (8.3%), 5001-10,000 hrs. (25%), and over 10,000 hrs. (16.6%). Two pilots did not have glass cockpit experience (16.6%). For those that did have experience, their total hours for the previous six months ranged from 0-300 hours (41.6%), 301-400 hours (25%), and 401-500 hours (16.6%). Participants’ total hours in an Airbus Glass cockpit ranged from 1-1000 hours (25%) to 1001-3000 hours (25%); six participants had no experience (50%). There was no gender requirement to participate, although all participants were male. There was also no preference for hand dominance or vision as long as the pilot had normal or corrected-to-normal vision. Military experience was not required, although half (50%) participants had previous military experience. Because the scenarios were in the Denver area, pilots were asked to indicate their familiarity with the area. 35% of pilots indicated that they were very familiar, 33.33% said they were familiar, 33.3% said somewhat familiar, 8.3% said a little familiar, and no pilots were said they were not familiar at all with the Denver area.

Pilots were not allowed to participate in the present study if they had participated in a previous study conducted at FDDRL (Flight Deck Display Research Laboratory) that used highly similar scenarios to the ones used in this study. However, if a participant participated in any other study by FDDRL or any other lab at NASA, they were still eligible. A total of 9 (75%) of the participants had participated in other studies by FDDRL and all 12 (100%) participants had previously participated in past studies in the FDDRL or in the NASA Airspace Operations Lab. Travel accommodations such as hotel, rental car, and per diems were not provided and prior preparations such as readings were not required for participation. Participants were compensated for four hours of their time at a rate of $33.37 per hour.

4. Results and discussion

4.1 Trust and transparency

The trust scale used evidenced acceptable reliability (alpha > .85). As shown in Figure 13 trust scores rose with increasing **System Transparency**. A within-subjects ANOVA found this effect to be significant ($F(2, 22) = 4.92$. $p < 0.05$). Although the overall trend is as expected, follow-up pairwise contrasts only confirmed that trust in the *Control* conditions was significantly lower than the trust in the *Logic* condition ($p < 0.05$).

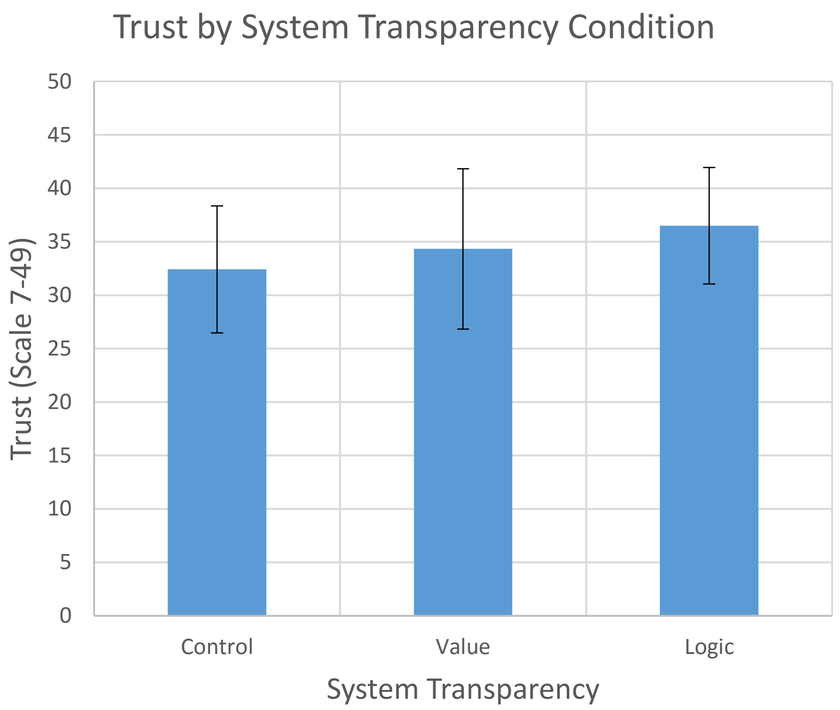


Figure 13: Trust by ***System Transparency*** *condition*

When evaluating each individual item in the trust scale, only two of the items were found to differ as a function of System Transparency. "I would feel comfortable relying on the recommendations of the ELP in the future", and "If I were facing a very hard task in the future, I would want to have the ELP with me", were both significantly higher in the *Logic* condition than in the *Control* condition ($p < .05$); and one statement, "If the ELP gave me a top recommendation, I would rely on the top recommendation of the ELP without hesitation", showed a similar, but only marginally significant effect ($p = 0.068$).

One concern with the design of this study was that, regardless of the counterbalancing of the order in which participants completed conditions, participants would gradually gain trust in the ELP over time as they were using the same ELP algorithm in each condition. There was no significant effect of trust by the order in which they completed the conditions ($F(2,22) = 0.15$, $p > 0.05$).

Participants were asked in the post condition surveys to rate the statement "The ELP was transparent" from strongly agree to strongly disagree. While the trend across the System Transparency conditions was as expected (Figure 14), it was not statistically significant ($F(2,22) = 1.34$, $p > 0.05$). This lack of statistical significance could, potentially, be attributed to the design of the study. Participants were asked to evaluate the transparency after each condition instead of at the end of the study. A participant's ability to evaluate transparency of the conditions would evolve and change after gaining experience with the various **System Transparency** conditions. In hindsight, the question should have been asked at the end of the study.

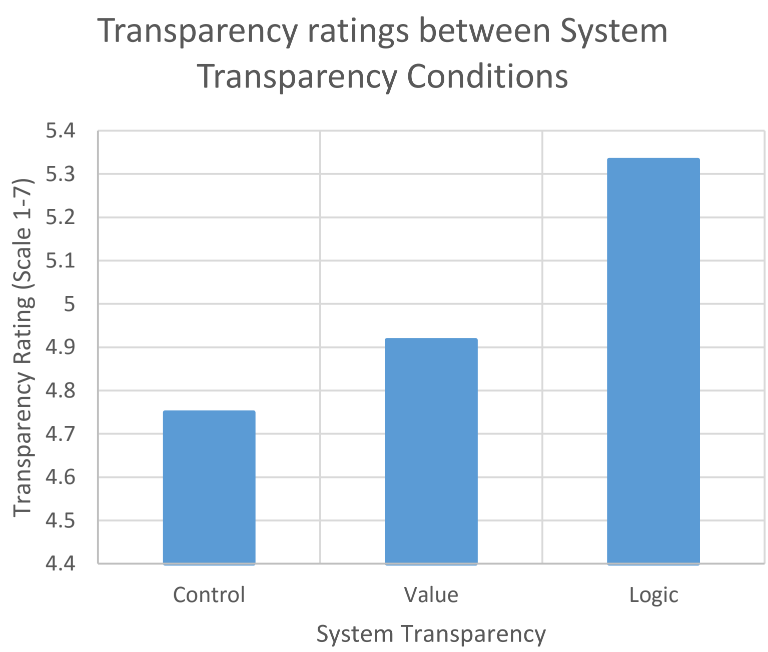


Figure 14: Transparency rating by condition

However, in contrast to this, the correlation between a participant's trust score and rating of transparency, was statistically significant, $r = .480$, $p < 0.01$ (Figure 15), confirming the hypothesized relationship between these two variables.

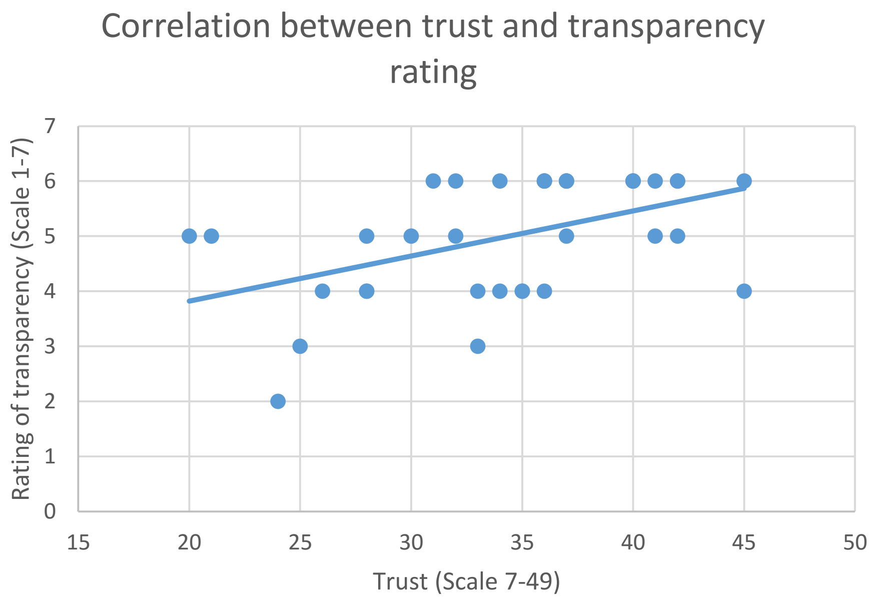


Figure 15: Correlation between trust and transparency

4.2 Agreement (Evaluations and Ranking)

There were two ways of measuring the agreement of the participants with the ELP: the safety of the diversions and the rank ordering of the diversions.

4.2.1. Evaluations of the safety of the diversions

The agreement between participants' categorical ratings of the diversion safety (excellent, good, fair, poor, unacceptable) and the corresponding values given by the ELP (see section 3.2.2.1) was measured with the inter-rater reliability kappa score based on the Landis and Koch's magnitude guidelines (Landis & Koch, 1977). Figure 16 shows the expected trend of agreement rising with increasing level of **System Transparency**. This effect was statistically significant ($F(2,22) = 4.05$, $p < 0.05$). Follow-up pairwise comparisons revealed a statistically significant difference between the *Control* condition and the *Logic* condition ($p < 0.01$).

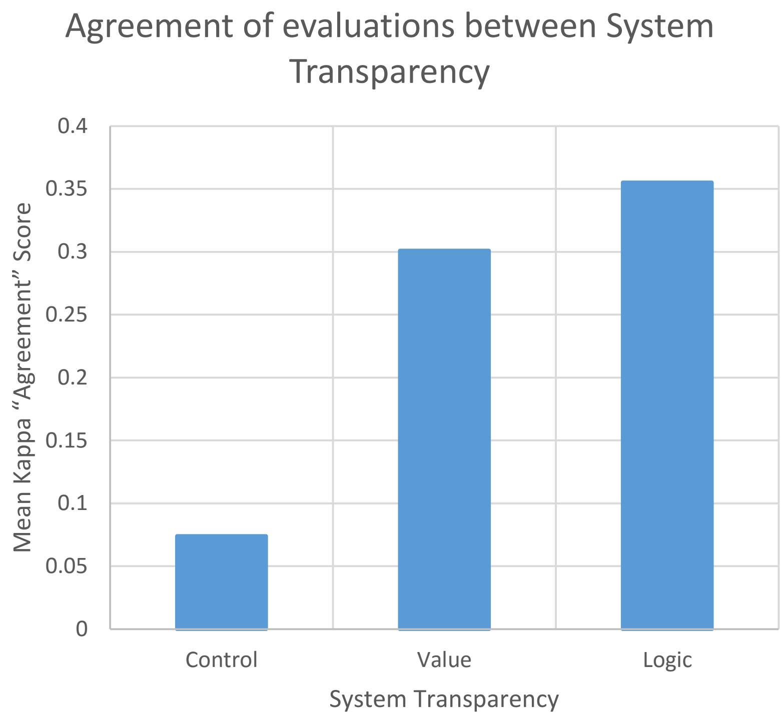


Figure 16: Agreement scores by condition

The overall agreement scores were lower than expected, and can be attributed to a number of reasons. First, participants stated that the task of evaluating the diversions was difficult because it was hard for them to be objective (e.g., without taking prior and extraneous information into consideration) and to remove their own biases and preferences (e.g., preference for a certain airport). They also tended to use their own scale when evaluating the diversions instead of the provided scale. They often based their assessment of safety on their own ability (e.g., "Oh, I know I can fly this diversion no problem."). Some participants also had a difficult time following directions. Some participants insisted on evaluating the diversion using different metrics than instructed. Finally, participants tended to be less conservative than the ELP in evaluating the safety of the diversions.

4.2.2. Rank order

The ELP's and participants' rank orderings as the 1st, 2nd, 3rd, 4th, and 5th diversion choice were also compared.

also compared.

Table 2 shows the expected relationship between these correlations and System Transparency (i.e., rising agreement between participants and the ELP as **System Transparency** increases), but this effect of System Transparency was not statistically significant, $F(2,22) = 1.12$, $p > 0.05$). One reason is that one would expect lower correlations when participants had additional relevant information that the ELP did not have in the UI conditions. Indeed, when comparing the correlations between the two **Scenario Information** conditions, correlations were significantly lower in UI scenarios than in EI scenarios ($F(1,35) = 10.45$, $p < 0.01$) as shown in Table 3.

Control	*Value*	*Logic*
0.338	0.471	0.583

Table 2: Rank correlation by condition; ELP & participants

Equal Information	*Unequal Information*
0.669	0.258

Table 3: Rank correlation by scenario; ELP & participants

	Control	*Value*	*Logic*
Equal Information	0.5	0.783	0.75
Unequal Information	0.2	0.158	0.417

Table 4: Rank correlation by condition and scenario; ELP & participants

Control	Value	Logic

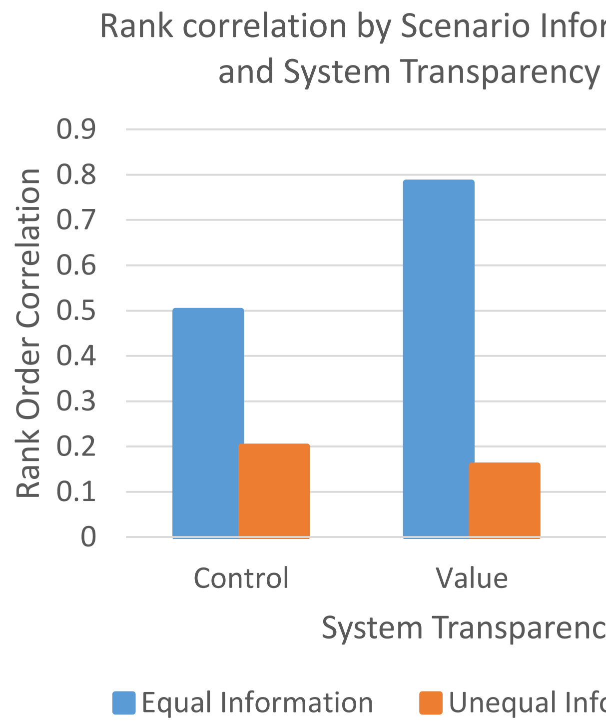


Figure 17: Rank correlations by condition and scenario; ELP & participants

When these rank correlations are expanded by both condition and scenario, detailed interactions can be examined (Table 4, Figure 17). As transparency rises across the **System Transparency** conditions (*Control*, *Value*, *Logic*), pilots' agreement in ranking increases within *EI* scenarios (i.e., where the ELP is making appropriate recommendations). A test of within subjects shows that this too is trending toward significance ($F(2,22) = 2.99$, $p > 0.05$), and the lack of statistical significance maybe due to insufficient variance between the *Value* and the *Logic* conditions. Conversely, across these same conditions, pilots' agreement in ranking is low and had no significant change ($F(2,22) = 0.73$, $p > 0.05$) within the *UI* scenarios (i.e., where the ELP is making inappropriate recommendations). Examining the figures, there is a jump in correlation in the *Logic* condition. This jump could be the result of the different types of scenarios used in the *UI* scenarios that caused the pilots to respond non-uniformly. Another possibility is that the transparency in these scenarios may have caused over-trust in the ELP. The transparency of the system may have been too convincing for the participants and influenced their agreement when they should have been disagreeing with the ELP. This issue deserves further investigation in future studies.

In addition to comparing the participants' rankings to the ELP, their responses were also compared to an SME response. To generate an SME response, all participants' responses were aggregated to create an SME ranking of diversions for each scenario. In other words, the SME rankings were created by examining what diversions participants typical ranked as 1^{st}, 2^{nd}, etc. for each scenario. Rank order correlation between the SME ranking and the participants rankings showed higher correlations across the board than that of the ELP and the participants.

0.704	0.637	0.662

Table 5: Rank correlation by condition; SME & participants

Equal Information	*Unequal Information*
0.727	0.608

Table 6: Rank correlation by scenario; SME & participants

	Control	*Value*	*Logic*
Equal Information	0.708	0.624	0.850
Unequal Information	0.715	0.650	0.475

Table 7: Rank correlation by condition and scenario; SME & participants

Similar analysis was conducted to compare SME and participants' correlation across **System Transparency** conditions, no statistically significant difference was found; however, when this comparison was done between the *EI* and *UI* scenarios, there was a slightly lower correlation in the *UI* scenarios (**Error! Reference source not found.**, Table 6). This seems to indicate that there is less agreement in the ranking of diversions amongst the participants in the scenarios where they are given additional information. This gives reason to suspect that the participants responded less uniformly to the additional information. When these correlations were expanded to both condition and scenario, this was an uncharacteristic dip in correlation in the *Logic* condition (Table 7, Figure 18). The dip again gives evidence that the participants have even less agreement in the highest **System Transparency** condition when given additional information that the ELP does not have. The reduced agreement could be attributed to the nature of the scenarios and the fact that participants may not all place the same value on the additional information. This was a sentiment expressed by participants in their debriefing interview. This interaction also warrants further research.

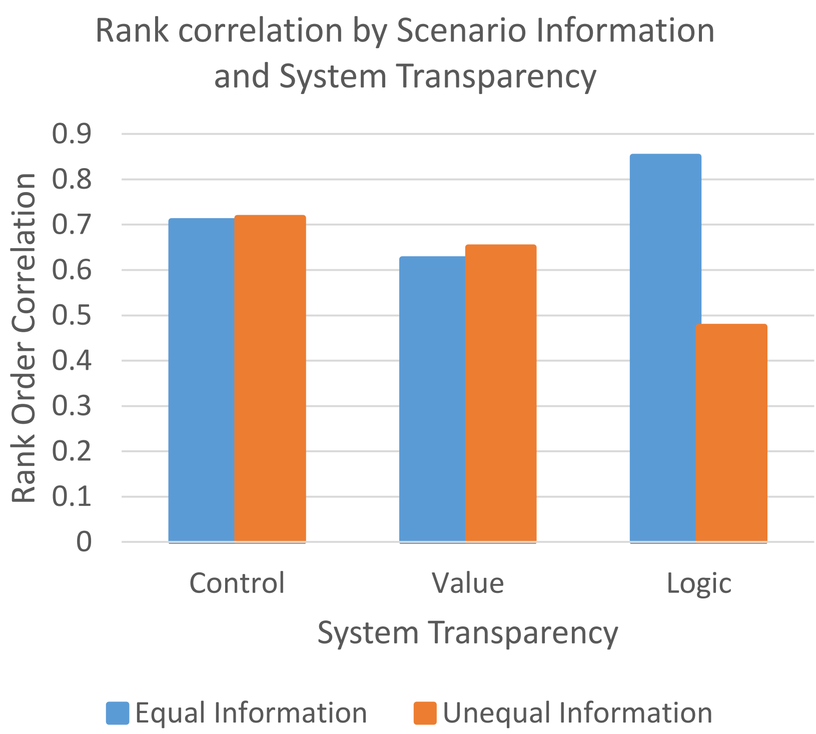


Figure 18: Rank correlations by condition and scenario; SME & participants

4.3 Confidence

In the post-trial survey, participants were asked to rank four statements about confidence with 1 being no confidence and 7 being extremely confident (refer to Appendix B). For each of the confidence statements, there was no significant difference as a function of **System Transparency** or **Scenario Information** conditions. For "confidence that the ratings I gave the diversions were accurate", the tests yielded $F(2,22) = 0.33$, $p > 0.05$ for **System Transparency**, and $F(1,11) = 0.89$, $p > 0.05$ for **Scenario Information**. For "confidence that the route I choose as the best route, was the best route", the test yielded $F(2,22) = 0.87$, $p > 0.05$ for **System Transparency** , and $F(1,11) = 0.0$, $p > 0.05$ for **Scenario Information**. For "confidence that the route I choose was the worst route, was the worst route", the tests yielded $F(2,22) = 1.36$, $p > 0.05$ for **System Transparency** , and $F(1,11) = 0.07$, $p > 0.05$ for **Scenario Information**. And finally, for "confidence that the diversions that I was given were appropriate", the test yielded $F(2,22) = 0.72$, $p > 0.05$ for **System Transparency** , and $F(1,11) = 0.10$, $p > 0.05$ for **Scenario Information**. Overall, participants had consistently high confidence for all four statements. In retrospect it was not very surprising to see no effect of the manipulations on participants' rated confidence in their choices. The participants are a skilled group who are likely to have high confidence in their ability to delineate diversion options.

4.4 Workload

An overall workload question was asked in both the post-trial and post-condition surveys (refer to appendices B & C). A slightly modified version of the overall workload scale was used (Hill et al., 1992). There was no significant effect on workload of **System Transparency** ($F(2,22) = 1.63$, $p > 0.05$), although there is a marginally significant effect of **Scenario Information** ($F(1,11) = 3.76$ $p = 0.079$). As shown in Figure 19, the workload in the *Value* condition is lower than in both the *Control* and the *Logic* conditions. This results suggests that when presented with only the value, the participants may have relied on the value and consequently spent less cognitive effort to examine other information. In addition, the workload is higher in the *UI* scenarios as expected because these scenarios requiring more cognitive effort to process additional information that the automation did not have.

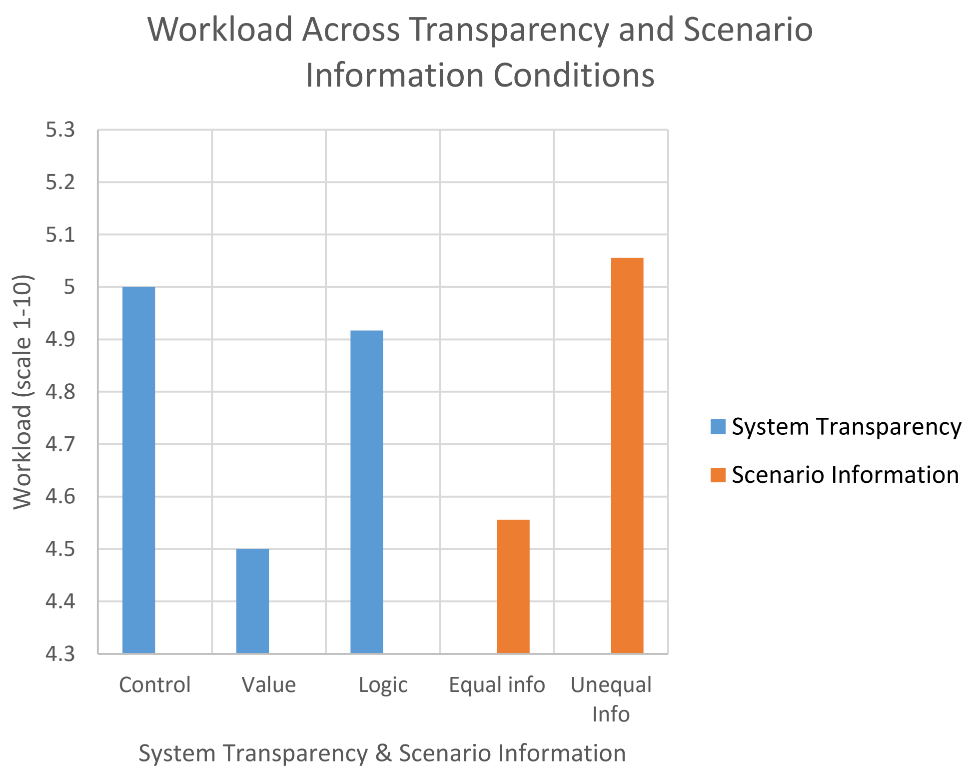


Figure 19: Workload by condition and scenario

4.5 Ratings, feedback, opinions, and other qualitative analysis

Participants were asked after each trial to rate the importance of the tools and information presented in the *Value* and *Logic* conditions from 1 (very important) to 5 (very unimportant). The additional information in the *Value* and *Logic* conditions (risk statement, enroute/approach/runway ratings, and the reasoning of the enroute/approach/runway ratings) were all rated highly (Figure 20). Reasoning and risk were not rated any differently from each other, and were just as important as enroute weather and the difficulty of the route and landing.

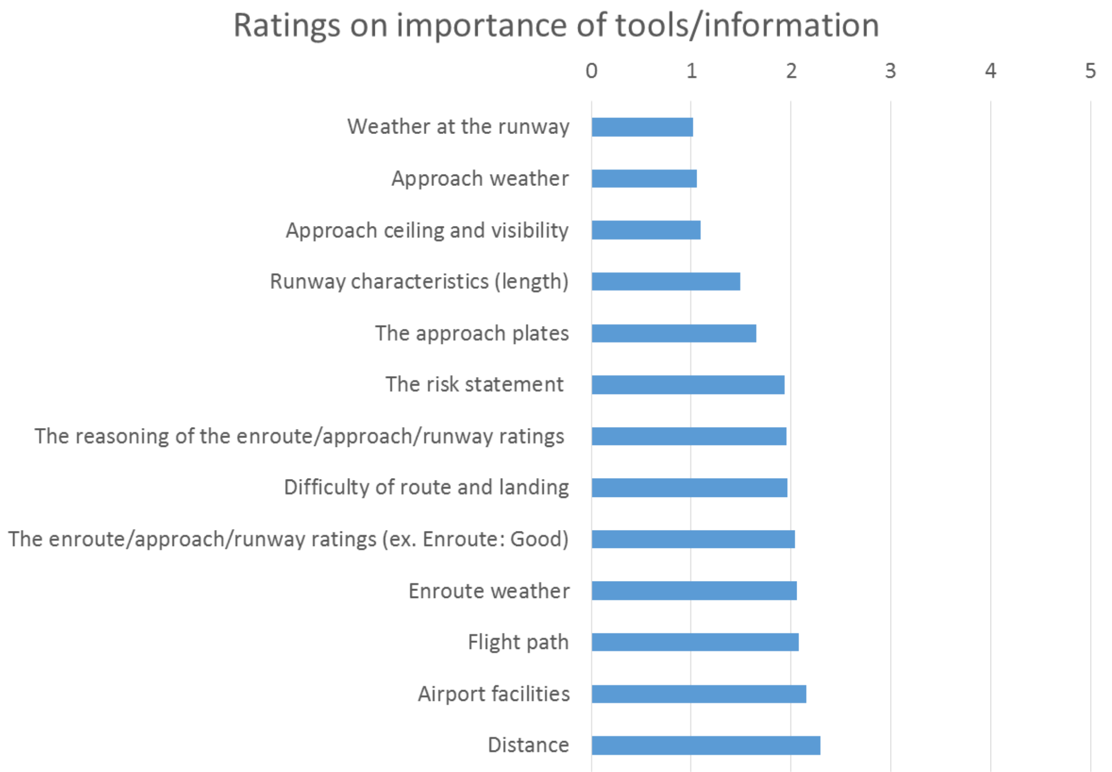


Figure 20: Ratings on importance of tools/information

Participants were also asked at the end of the study to rate the helpfulness of the tools and information from 1 (extremely helpful) and 6 (a distraction). The risk statement, enroute/approach/runway ratings, and the reasoning of the enroute/approach/runway ratings were all rated as helpful to the participant (Figure 21). The lowest rated information/tool was the computed ELP diversion path; however, this should be taken with a grain of salt. The 2D path display (overlaid on top of a Google earth image) used in the study is different than what the participants are familiar with and is not representative of how the ELP currently displays its recommended 3D flight path.

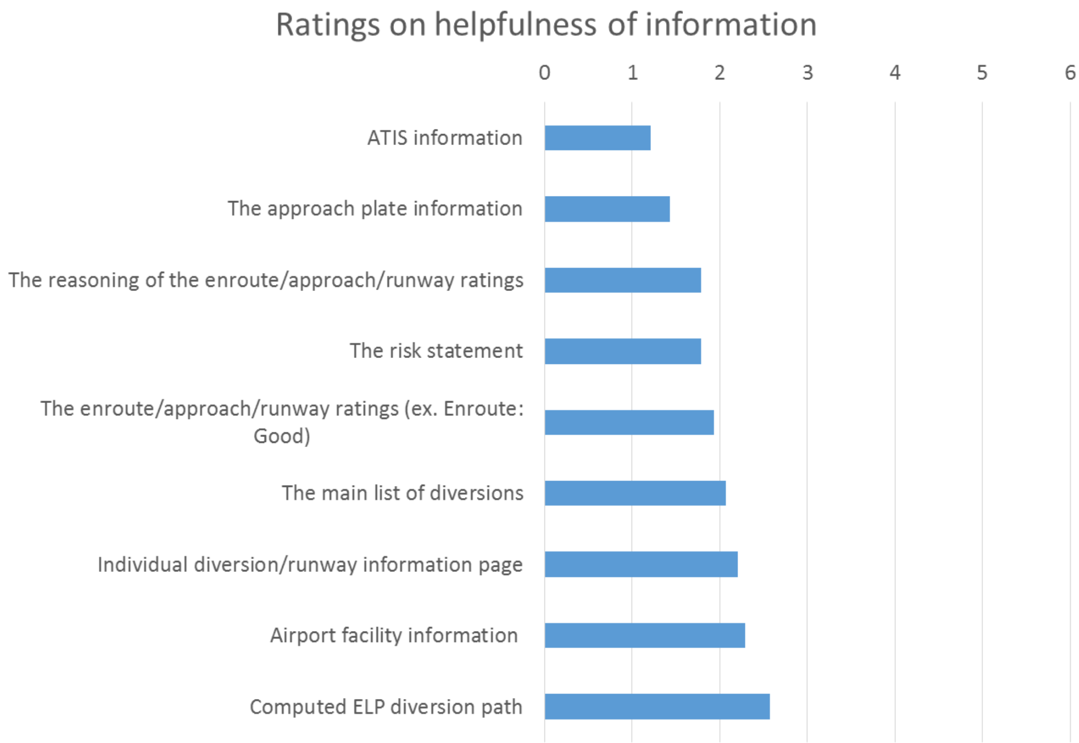


Figure 21: Ratings on helpfulness of tools and information

Overall, 100% of the participants preferred the *Logic* condition. Their feedback was very positive. For example, some participants stated that, "It gave me the most useful information at a glance", "I realized I missed the reasoning when I didn't have it", and "it helped me catch something that I would have otherwise missed." Participants also stated that for the most part, they understood the recommendations and evaluations the ELP made. Participants also provided a number of suggestions for what they would like to see in the future. Their feedback falls into two categories: 1) additional information that they would like to see, and 2) different display/information configurations. It is important to note that some of the suggested changes made by participants are current elements of the ELP. In this study, a modified version of the ELP was used and included randomizing the recommendations as well as providing evaluations on a secondary page instead of on the main list of diversions. In addition to the information in the *Logic* condition, participants would like to see or have: calculated tailwind/headwind component, modification to the way weather visibility is phrased (i.e., from "visibility being unsafe" to "weather was at minimums"), runway lighting configuration, more enroute weather information, air traffic control (ATC) inputs and opinions, dispatch inputs and opinions, runway braking action reports, approximate landing fuel and weight for each diversion, terminal aerodrome forecast (TAF)'s and weather trending reports, the ability to designate the type of diversion (e.g., weather, medical, fuel, etc.), and terrain. The different display configurations participants requested were: display risk percentages on the main page to minimize clicking, showing diversions in a ranked (instead of randomized) order, clearer display of main reason for ELP's choice of a diversion, weather shown in METAR format instead of being written out in words, only offer runways in use, the addition of a higher fidelity map with weather and flight path (to include altitude/3D path), and a better way to identify facility offerings by carrier at the airport. Participants were also asked if there were factors that should have been weighed more heavily by the ELP. Participants indicated a desire to have thunderstorm activity, distance from current position to airport, maintenance issues, weather and approach minimums, aircraft performance data, and enroute weather to be weighed heavier.

5. Conclusion

The current study corroborates the wisdom that sharing intent and awareness can help the operator place appropriate trust in complex automated systems. Specifically, the results suggest that for automated systems that offer recommendations that these recommendations are accompanied by the logic which they were derived from. Our design and implementation of the value the risk and the logic behind each diversion recommendation helped pilots understand the ELP's evaluation process and recommendations by giving pilots useful information and the rationale of the recommendations.

There were several limitations to this study. The first of which is that it was a low fidelity environment. Participants were not put in a dynamic situation involved in a diversion decision. They were also in a limited environment as they did not have access to dispatch or other sources of information that they typically use in a diversion situation. The map and information displays were also in formats that were different to what pilots are familiar with. The study also used a small sample size (n=12). The experimental tasks given to pilots also pushed them out of their comfort zone and may have been too challenging to do because a number of participants failed or had a difficult time following instructions on how to complete the tasks. It was difficult for pilots to make unbiased and objective evaluations.

6. List of symbols, abbreviations, and acronyms

AGL	Above Ground Level
AL	Airline
ALT	Altitude
APT	Airport
APPCH	Approach
ATIS	Automatic Terminal Information Service
ATC	Air Traffic Control
CDU	Computer Display Unit
COS	Colorado Springs Airport
CYS	Cheyenne Regional Airport
DEN	Denver International Airport
DIR	Direction
Dist.	Distance
DFW	Dallas/Fort Worth International Airport
EGE	Eagle County Regional Airport
EI	Equal Information
ELP	Emergency Landing Planner
FDDRL	Flight Deck Display Research Laboratory
Ft.	Feet
FMS	Flight Management Systems
FNOC	Navy Fleet Numerical Oceanography Center
GCS	Ground Control Station
GJT	Grand Junction Regional Airport
ILS	Instrument Landing System
Kt. or kts.	Knot(s)
LDA/DME	Localizer Directional Aid/Distance Measuring Equipment
LEN	Length

LOC	Localizer
MINS	Minimums
Maint.	Maintenance
Med	Medical
METAR	Meteorological Aerodrome Report
NASA	National Aeronautics and Space Administration
NTD	Need to Divert
NWS	National Weather Service
ORD	Chicago O’Hare International Airport
Passen.	Passenger
PHX	Phoenix Sky Harbor International Airport
PUB	Pueblo Memorial Airport
RMK	Remark
RNV (RNAV)	Area navigation approach
RW or RWY	Runway
SA	Situational Awareness
SEA	Seattle–Tacoma International Airport
SFO	San Francisco International Airport
SME	Subject Matter Expert
SJSURF	San Jose State University Research Foundation
TAF	Terminal Area Forecast
T/TD	Temperature and Dew point
UI	Unequal Information
VIS	Visibility
WX	Weather
Z	Zulu

7. References

Abbott, K., Slotte, S., Stimson, D., Bollin, E., Hecht, S., Imrich, T. & Woods, D. (1996). The interfaces between flightcrews and modern flight deck systems. *Washington, DC: Federal Aviation Administration*.

Billings, C. E. (1996). *Aviation automation: The search for a human-centered approach*. Lawrence Erlbaum Associates Publishers.

Chen, J. Y. C., & Barnes, M. J. (2014). Human–Agent Teaming for Multirobot Control: A Review of Human Factors Issues. *IEEE Transactions on Human-Machine Systems*, *44*(1), 13–29. doi:10.1109/THMS.2013.2293535

Chen, J. Y. C., Barnes, M. J., & Harper-Sciarini, M. (2011). Supervisory control of multiple robots: Human-performance issues and user-interface design. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*.

Cummings, M. L., Buchin, M., Carrigan, G., & Donmez, B. (2010). Supporting intelligent and trustworthy maritime path planning decisions. *International Journal of Human Computer Studies*, *68*(10), 616–626.

Endsley, M., & Jones, D. (2011). *Designing for Situation Awareness: an approach to User-Centered Design*. Sound Parkway, NW: Taylor & Francis.

Hackos, J. T., & Redish, J. C. (1998). *User and Task Analysis for Interface Design*. Danvers, MA: Wiley Computer Publishing.

Hill, S. G., Iavecchia, H. P., Byers, J. C., Bittner, A. C., Zaklade, A. L., & Christ, R. E. (1992). Comparison of Four Subjective Workload Rating Scales. *The Journal of the Human Factors and Ergonomics Society*, *34*(4), 429–439. doi:10.1177/001872089203400405

Kirwan, B., & Ainsworth, L. K. (1992). *A Guide to Task Analysis*. Padstow, Cornwall: Taylor & Francis.

Lachter, J., Battiste, V., Matessa, M., Dao, Q. V, Koteskey, R., & Johnson, W. W. (2014). Toward Single Pilot Operations : The Impact of the Loss of Non-verbal Communication on the Flight Deck. In *HCI Aero*.

Lachter, J., Brandt, S. L., Battiste, V., Matessa, M., Ligda, S. V, & Johnson, W. W. (2014). Toward Single Pilot Operations : Developing a Ground Station. In *HCI Aero*.

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174.

Lee, J. D., & See, K. A. (2004). Trust in Automation: Designing for Appropriate Reliance. *The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80. doi:10.1518/hfes.46.1.50

Leite, J. C. S. do P., & Cappelli, C. (2010). Software Transparency. *Business & Information Systems Engineering*.

Lyons, J. B. (2013). Being Transparent about Transparency : A Model for Human-Robot Interaction. In *Trust and Autonomous Systems: Papers from the 2013 AAAI Spring Symposium* (pp. 48–53).

Lyons, J. B., & Stokes, C. K. (2012). Human-Human Reliance in the Context of Automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

Meuleau, N., Neukom, C., Plaunt, C., Smith, D. E., & Smith, T. (2013). The Emergency Landing Planner Experiment.

Meuleau, N., Plaunt, C., & Smith, D. E. (2008). *Emergency Landing Planning for Damaged Aircraft* (Vol. 1000).

Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*.

Ososky, S., Sanders, T., Jentsch, F., Hancock, P., & Chen, J. Y. C. (2014). Determinants of system transparency and its influence on trust in and reliance on unmanned robotic systems. In R. E. Karlsen, D. W. Gage, C. M. Shoemaker, & G. R. Gerhart (Eds.), *Unmanned Systems Technology XVI* (Vol. 9084, p. 90840E). doi:10.1117/12.2050622

Parasuraman, R., & Riley, V. (1997). Humans and Automation: Use, Misuse, Disuse, Abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*.

Sheridan, T. B. (1988). Trustworthiness of command and control systems. In *IFAC Man-Machine Systems Conference*. Oulu, Finland.

Sinha, R., & Swearingen, K. (2000). The Role of Transparency in Recommender Systems.

Wang, L., Jamieson, G. a., & Hollands, J. G. (2009). Trust and Reliance on an Automated Combat Identification System. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *51*(3), 281–291. doi:10.1177/0018720809338842

Appendix A: Task Worksheet

1) Please indicate your rating (**how safe**) for each of the diversions given. In other words, choose **how safe** each diversion is and if you would fly the diversion yourself if you were in the given scenario. **Please only take safety of the aircraft into mind when giving your ratings.** Below you will also find a guide that we recommend for rating the diversions.

100% - 90% success: Excellent
89% - 76% success: Good
75% - 56% success: Fair
55% - 41% success: Poor
40% and below: Unacceptable

	Rating					Would you fly this diversion?	
	Excellent	Good	Fair	Poor	Unacceptable	Yes	No
CYS 9							
CYS 13							
CYS 31							
CYS 27							
COS 35L							

2) Given **ALL** the circumstances of the scenario, please indicate which diversion option would be your 1st (best), 2nd, 3rd, 4th, and 5th (worst) choice to divert to.

CYS 9 _______

CYS 13 _______

CYS 31 _______

CYS 27 _______

COS 35L _______

3) For the diversion selected as your 1st choice in question 2, what factors contributed to your decision? Please check all that apply.

- ☐ Distance
- ☐ Weather
- ☐ Approach Plate Information
- ☐ Length of runway
- ☐ Approaches available to that airport
- ☐ Airport facilities available
- ☐ The provided risk statement (if applicable)
- ☐ The provided risk statement and reasoning (if applicable)
- ☐ Other: Please write below

4) If you were flying the suggested path you selected as your 1st choice in question 2, would you have SIGNIFICANTLY deviated from it?

- ☐ Yes (go to # 5)
- ☐ No (go to # 6)

5) If yes, which of the following factors would have contributed to your deviation?

- ☐ Weather
- ☐ Difficult to fly
- ☐ Shortcut
- ☐ Altitude issues
- ☐ Other: Please write below

6) For the diversion selected as your 5th choice in question 2, what factors contributed to your decision? Please check all that apply.

- ☐ Distance
- ☐ Approach Plate Information
- ☐ Approaches available to that airport
- ☐ The provided risk statement (if applicable)
- ☐ The provided risk statement and reasoning (if applicable)
- ☐ Other: Please write below
- ☐ Weather
- ☐ Length of runway
- ☐ Airport facilities available

Appendix B: Post Trial Survey

1) Please move the slider to the number which best corresponds to how you rate your overall workload for this scenario with 0 being "very low" and 10 being "very high. “For example:"2" would be enroute, cruise phase of flight in level flight (only monitoring aircraft state and occasionally talking with ATC) "8" would be descent and approach phase with poor weather at destination (ATC issues a major reroute then issues holding instructions requiring numerous CDU/FMS entries, then configuring the aircraft for holding, briefing for next steps, and assessing fuel load to determine how long to hold).

Workload	0	1	2	3	4	5	6	7	8	9	10

2) For the questions below, please indicate the amount of confidence you have on the following statements, with 1 being no confidence at all and 7 being extremely confident.

	1	2	3	4	5	6	7
Confidence that the ratings I gave the diversions are accurate							
Confidence that the route I choose as the best route, was the best route							
Confidence that the route I choose as the worst route, was the worst route							
Confidence that the diversions that I was given were appropriate							

3) Please indicate how much you disagree or agree about the following statements.

	Strongly Agree	Agree	Slightly Agree	Neutral	Slightly Disagree	Disagree	Strongly Disagree	N/A
I found the main summary page helpful.	❍	❍	❍	❍	❍	❍	❍	❍
I found the individual route diversion information pages helpful.	❍	❍	❍	❍	❍	❍	❍	❍
The "best" route I chose was acceptable.	❍	❍	❍	❍	❍	❍	❍	❍
The "best" route I chose is a route I would have picked myself if I did not have the ELP.	❍	❍	❍	❍	❍	❍	❍	❍
I had enough information to evaluate the quality of the routes.	❍	❍	❍	❍	❍	❍	❍	❍
It was easy for me to pick the best path.	❍	❍	❍	❍	❍	❍	❍	❍
It was easy for me to pick the worst path.	❍	❍	❍	❍	❍	❍	❍	❍
I understood the reasoning of why these diversions were given to me.	❍	❍	❍	❍	❍	❍	❍	❍
I found the risk statement helpful (If applicable)	❍	❍	❍	❍	❍	❍	❍	❍
I found the reasoning statements and pop-outs helpful (If applicable)	❍	❍	❍	❍	❍	❍	❍	❍

4) How important were the following factors in evaluating the quality of the given diversions? If there was a factor that is not listed that you felt was important in evaluating the diversion, please write it in the text space near "other" and rate it appropriately.

	Very important	Slightly important	Neutral	Slightly unimportant	Very unimportant	N/A
Flight path	❍	❍	❍	❍	❍	❍
Distance	❍	❍	❍	❍	❍	❍
The approach plates	❍	❍	❍	❍	❍	❍
Approach ceiling and visibility	❍	❍	❍	❍	❍	❍
Enroute weather	❍	❍	❍	❍	❍	❍
Approach weather	❍	❍	❍	❍	❍	❍
Airport facilities	❍	❍	❍	❍	❍	❍
Runway characteristics (length)	❍	❍	❍	❍	❍	❍
Weather at the runway	❍	❍	❍	❍	❍	❍
Difficulty of route and landing	❍	❍	❍	❍	❍	❍
The risk statement (if applicable)	❍	❍	❍	❍	❍	❍
The enroute/approach/runway ratings (ex. Enroute: Good) (if applicable)	❍	❍	❍	❍	❍	❍
The explanation of the enroute/approach/runway ratings (if applicable)	❍	❍	❍	❍	❍	❍
Other	❍	❍	❍	❍	❍	❍
Other	❍	❍	❍	❍	❍	❍

5) Any additional comments?

Appendix C: Post Condition Survey

1) Using your experiences with the condition you just tested, please rate the following statements from "strongly disagree" to "strongly agree."

	Strongly Agree	Agree	Slightly Agree	Neutral	Slightly Disagree	Disagree	Strongly Disagree
If the ELP gave me a top recommendation, I would rely to the top recommendation of the ELP without hesitation.	❍	❍	❍	❍	❍	❍	❍
I think using the ELP will lead to positive outcomes.	❍	❍	❍	❍	❍	❍	❍
I would feel comfortable relying on the recommendations of the ELP in the future.	❍	❍	❍	❍	❍	❍	❍
When the task was hard, I felt like I could depend on the ELP.	❍	❍	❍	❍	❍	❍	❍
If I were facing a very hard task in the future, I would want to have the ELP with me.	❍	❍	❍	❍	❍	❍	❍
I would be comfortable allowing this system to make a diversion decision for me.	❍	❍	❍	❍	❍	❍	❍
If I had my way, I would NOT let the system have any influence over issues that are important to the task.	❍	❍	❍	❍	❍	❍	❍
The ELP was transparent.	❍	❍	❍	❍	❍	❍	❍

2) Please move the slider to the number which best corresponds to how you rate your overall workload in this condition (with the given display and provided information) with 0 being "very low" and 10 being "very high. "As an example: "2" would be enroute, cruise phase of flight in level flight (only monitoring aircraft state and occasionally talking with ATC)"8" would be descent and approach phase with poor weather at destination (ATC issues a major reroute then issues holding instructions requiring numerous CDU/FMS entries, then configuring the aircraft for holding, briefing for next steps, and assessing fuel load to determine how long to hold)

Workload	0	1	2	3	4	5	6	7	8	9	10

3) Any comments about this condition?

Appendix D: Debriefing Survey

1) Which of the following display/information configurations were most useful to you?

□ Base information

□ Base information + Risk

□ Base information + Risk + Reason

2) Why was the above selected display configuration the most useful?

3) How helpful was each of the following information elements?

	Extremely helpful	Very helpful	A bit helpful	Not very helpful	Not helpful at all	Distraction
The main list of diversions	❍	❍	❍	❍	❍	❍
Individual diversion/runway information page	❍	❍	❍	❍	❍	❍
ATIS information	❍	❍	❍	❍	❍	❍
The approach plate information	❍	❍	❍	❍	❍	❍
Airport Facility Information	❍	❍	❍	❍	❍	❍
Computed ELP diversion path	❍	❍	❍	❍	❍	❍
The risk statement	❍	❍	❍	❍	❍	❍
The enroute/approach/runway ratings (ex. Enroute: Good)	❍	❍	❍	❍	❍	❍
The explanation of the enroute/approach/runway ratings	❍	❍	❍	❍	❍	❍

4) What aspects of the ELP were clear/easily understandable to you?

5) What aspects of the ELP were NOT clear/easily understandable to you? What can be done to make this clearer?

6) Was there any additional information you needed or was any information not clear to you in the risk condition?

7) For the risk statement, are there changes (different information, different presentation, etc.) that would have helped?

8) Was there any additional information you needed or was any information not clear to you in the risk + reasoning condition?

9) For the reasoning statements, are there changes (different information, different presentation, etc.) that would have helped?

10) What, if any, factors in your decision-making process of choosing a diversion should have been weighted more heavily by the ELP?

11) Were there any additional factors that should have been weighted by the ELP?

12) Please define transparency for automated systems, in your own words.

13) Think about your overall decision making process when selecting a diversion route and runway (without the ELP). Please rank the order of importance of the factors listed below when making a diversion decision. You can drag the choices to rank order from most important (what you think about first), to least important (what you think about last). If there are other key factors that you would like to include in this list, please fill out the blank text spot labelled "other" and rank it appropriately.

______ Approach Ceiling and Visibility (and minimums)
______ Distance to runway (pending fuel)
______ Enroute Weather
______ Approach Weather
______ Facilities available at the airport (refueling, hotels & rebooking for passengers, etc.)
______ Runway Characteristics (length, width, etc.)
______ Weather at the runway
______ Other Approach Plate information
______ Compatibility with Airline Company
______ Familiarity with that airport/runway
______ Difficulty of landing
______ Other
______ Other

14) Any other comments?

Appendix E: Demographic Survey

1) Name

2) Please report the airline in which you fly (or last flew if retired).

3) What position do you currently hold (or last held) there?

- □ Captain
- □ First Officer

4) Are you a retired pilot?

- □ No
- □ Yes (if yes, how many years and months?) _____________________

5) Years in current position (or last position if retired)

- □ 0 - 1 year
- □ 1 - 5 years
- □ 5 - 10 years
- □ Over 10 years
- □ N/A

6) Total hours flown as line pilot

- □ 1 - 1000
- □ 1001 - 3000
- □ 3001 - 5000
- □ 5001 - 10000
- □ Over 10000

7) Total hours flown in Boeing glass

- □ 1 - 1000
- □ 1001 - 3000
- □ 3001 - 5000
- □ 5001 - 10000
- □ Over 10000
- □ N/A

8) Total hours flown in Boeing glass in last 6 months/last 6 months flown

- □ 0 - 300
- □ 301 - 400
- □ 401 - 500
- □ 501 - 600
- □ Over 600
- □ N/A

9) Total hours flown in Airbus glass

- □ 1 - 1000
- □ 1001 - 3000
- □ 3001 - 5000
- □ 5001 - 10000
- □ Over 10000
- □ N/A

10) Please list the last three aircraft type ratings you hold (or last held if retired)

- □ 1
- □ 2
- □ 3

11) Please provide your current aircraft qualification type (or last held if retired)

12) Do you have any military flying experience?

- □ Yes
- □ No

13) Have you participated in past studies with this lab or the NASA Airspace Operations Lab?

- □ Yes
- □ No

14) If yes, have you participated any of the Single Pilot Operations (SPO) studies in the past year?

- □ Yes
- □ No

15) How familiar are you with flying in and around the Denver area?

- □ Very familiar
- □ Familiar
- □ Somewhat familiar
- □ A little familiar
- □ Not familiar at all

16) Have you ever had to make a diversion decision while flying? If yes, please explain the circumstances of your most recent diversion.

17) Please supply your email address

Appendix F: Training Document

Training for Trust and Transparency Study

Briefing

- Introductions: Consent Form and Demographics Survey
- Introduction of the Emergency Landing Planner
  - What is the Emergency Landing Planner (ELP)
  - The purpose of ELP
  - How the ELP works
- Introduction of Transparency Project
  - Project Objectives
  - Scenarios and Conditions
    - Scenarios around Denver area
    - The different conditions:
      - C1: Input (baseline condition)
      - C2: Input and risk (risk condition)
      - C3: Input, risk, and reasoning (risk + reasoning condition)

Training

- Scenarios
  - Train how to use ELP tools for each condition and review if necessary
  - Explain purpose of situation briefing of each scenario: Explain the tasks they are required to do: Complete task worksheet – Give rating on diversions, select the best and the worst runway
- Explain when they will be completing the surveys
  - Post-Trial survey after every scenario (6 total)
  - Post-Condition survey issued after every condition (3 total)
  - Debriefing survey and interview
- Instructions to Use Tools
  - Getting Started
  - Main Summary Page Navigation
  - Tools and information
    - Automatic Terminal Information Service (ATIS)
    - Runway information
    - Airport Facilities
  - Condition 1
  - Condition 2
  - Condition 3

Introduction of the Emergency Landing Planner

The Emergency Landing Planner (ELP) is an automated recommender system that is designed to generate and evaluate the best diversion for a damaged aircraft in an emergency situation. The ELP was designed to generate a list of diversion in a ranked order in an emergency situation, however, in this study, it is being used as a diversion planner for off-nominal situations.

For this study, it is very important to remember that the diversions being shown are in a randomized order instead of a ranked order. Normally, the ELP will order the safety of each diversion option. The ELP compiles information from many different sources to generate a list of diversions. The ELP uses ATIS information at the airport, current weather, flyability of the aircraft, GPS location and terrain, airport/runway characteristics (approach plates), and the population density of the airport into consideration when generating diversions. The main factors that the ELP takes into consideration are: the approach ceiling, enroute distance, enroute turns, enroute weather, approach weather, population density of the airport (i.e. in case of crash), runway length, runway width, landing speed (tail wind + required approach speed), runway surface, approach visibility, landing crosswind, and terrain. The terrain data, urban development and urban population density is gathered by the Navy Fleet Numerical Oceanography Center (FNOC).The weather used by the METAR is gathered by the National Weather Service (NWS). However, the ELP does not know everything. It does not know information about the resources at an airport, the state of passengers, the state of the pilot, or preferences of the pilot.

In this study, you will be provided different types and amounts of information generated by the ELP. This will be further explained in the conditions. The options can be updated by the pilot to account for the changing location, altitude and velocity of the aircraft. The ELP can factor in subsequent degradation or failures that change the predicted control envelope, and updated weather and airport information. The best runways and approach paths generated are presented to the pilots in ranked order; the pilot makes the final decision on where to land. For this study, the options will be displayed in a randomized order and you will not be able to update the diversions given.

It is also important to know that the ELP, in this study, will only be focusing on 6 airports in the Denver area: Denver, Cheyenne, Colorado Springs, Eagle, Pueblo, and Grand Junction (see figure 2). You also may only be shown diversions to two airports. In this study, you will be show the top 5 recommendations in a randomized order. You will have to accept that if you are not show a runway or an airport, that diversion has too poor of weather or is a less safe diversion than the ones you are show. You must focus on the diversions you are given.

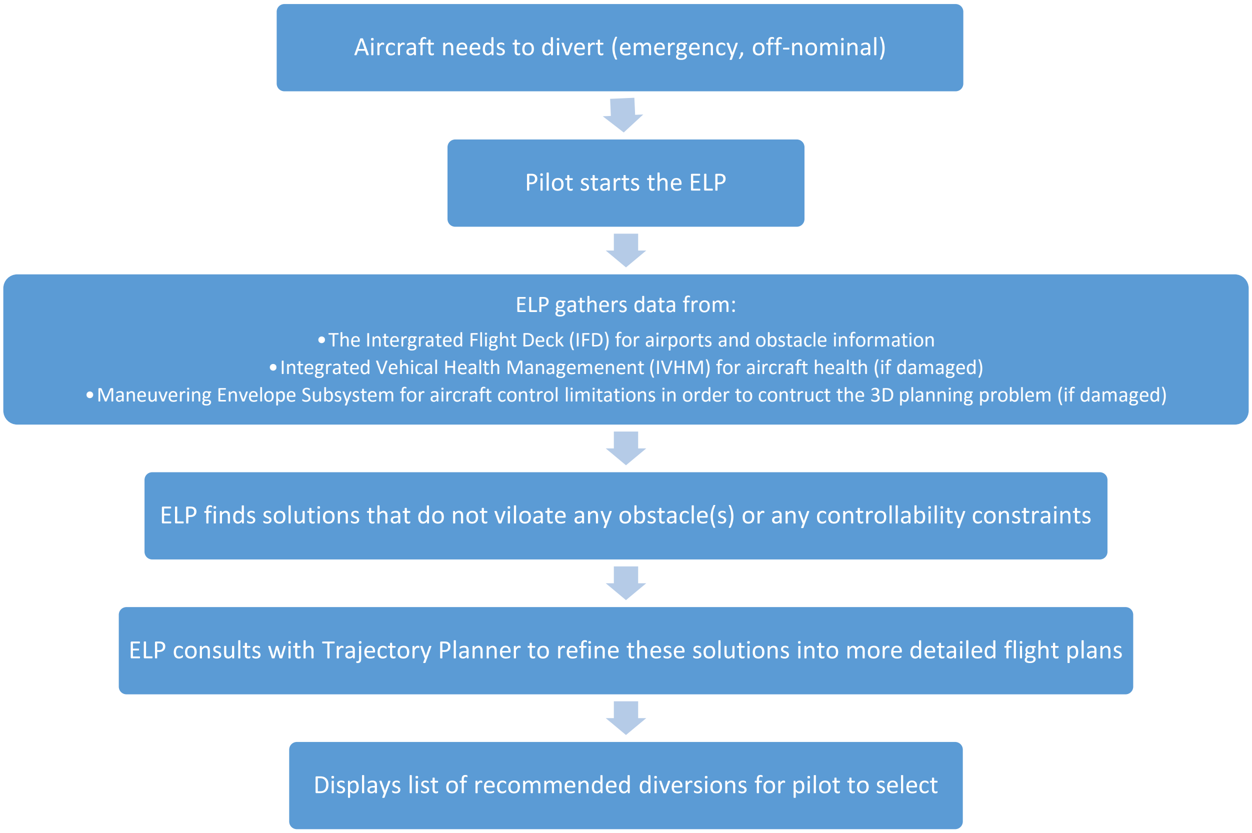


Figure 1. ELP’s process

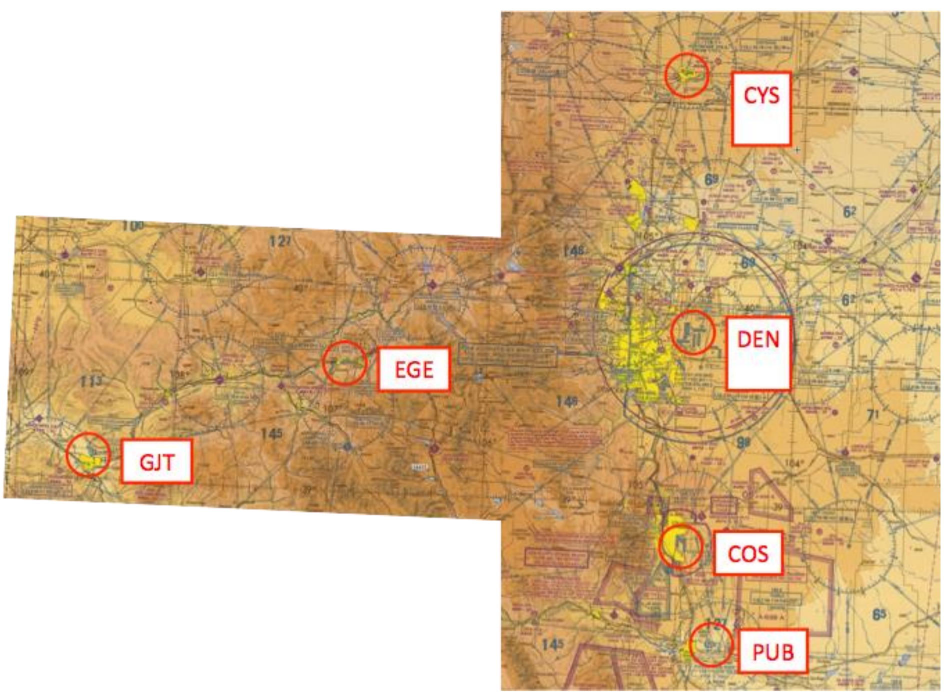


Figure 2. Map of Denver area and airports involved in this study.

Introduction of Transparency Project

Project Objectives

1) How does the type of explanation (transparency) effect pilot performance and trust?

Scenarios and Conditions

There will be a total of six different scenarios all in the Denver area. Flights will be either arriving or departing from the Denver area. In each scenario, there will be an event which requires the pilot to divert to another airport.

For each condition, the Emergency Landing Planner will provide a different level of explanation.

Conditions

C1	Input (baseline) information
C2	Input and risk
C3	Input, risk, and reasoning statement

Training

Scenarios

There will be a brief training period at the beginning of each condition to ensure you understand how to properly use the tools provided. At the beginning of each scenario, you will receive a briefing of your flight. There will be two scenarios for each condition. The order of the conditions and scenarios will be randomized. Each scenario requires you to rate the diversions based on safety as well as indicating the 1st (best), 2nd, 3rd, 4th and 5th (worst) diversion from a list of randomized recommendations. In addition, we would like to record your reasoning and decision process. We want to understand why and how you are choosing a diversion as the best diversion or as an unacceptable diversion path. While going over each diversion option, please discuss your decision making process **aloud**. For example, if you are trying to find the best diversion, we would like to hear you reason out loud while you are going through each option such as, “I like this diversion path because it has the longest runway” or “I don’t like this diversion path because it doesn’t have excellent airport facilities.”

During each scenario, you will be given a task worksheet to complete. You will be asked to give a ranking on how safe you think each diversion is as well as choosing which diversion you think is the best to the worst. You will also be asked to give explanations to your answers. Once you complete the task worksheet, you have finished the scenario.

Surveys

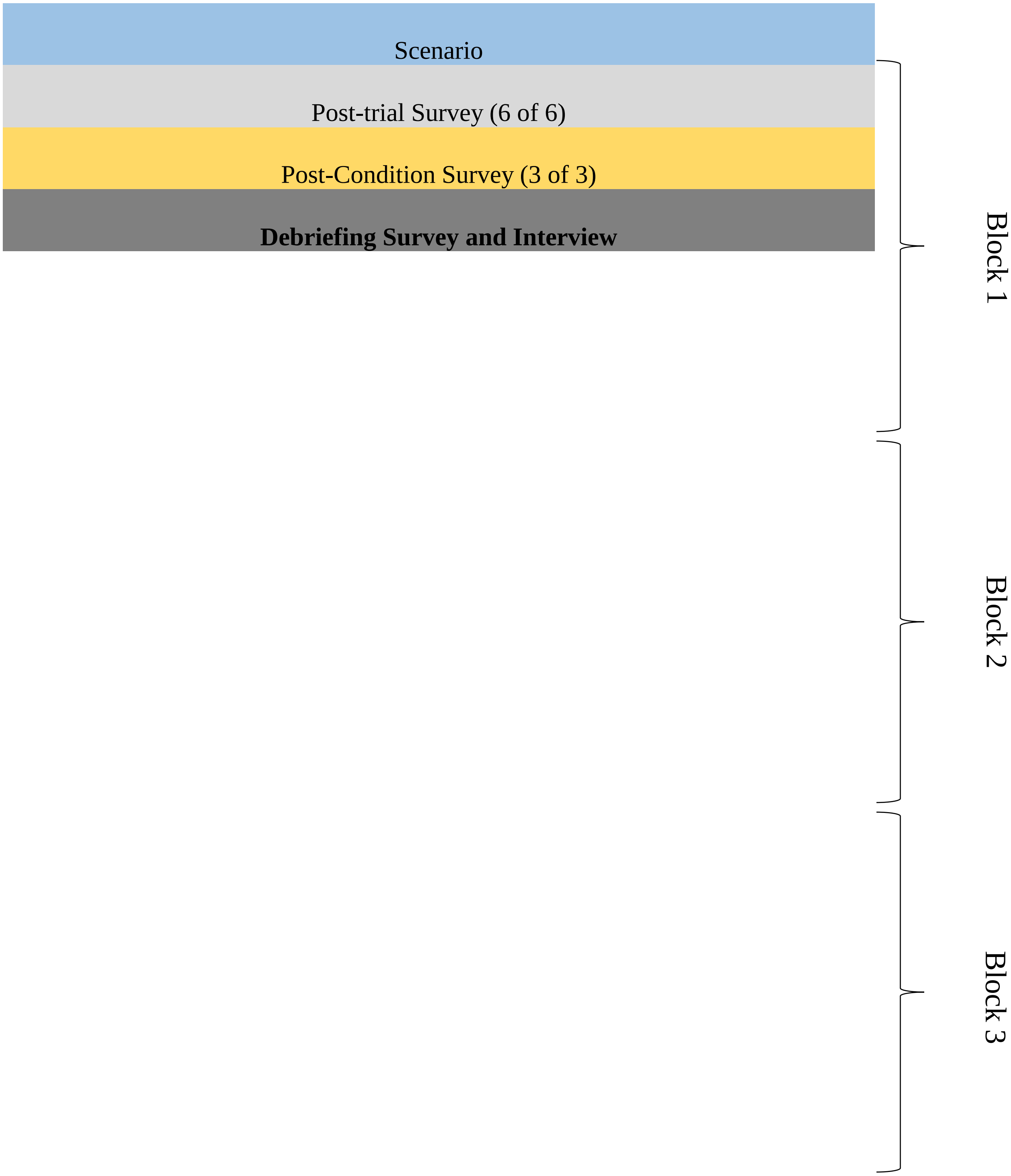
After each scenario, you will complete a brief post-trial survey. There will be a total of six (6). These are short questions based on your experiences completing the scenario.

After every condition, you will complete a post-condition survey. There will be a total of three (3). These are short questions based on your experiences on both scenarios in the condition.

Lastly, after all of the scenarios have been competed, you will complete a debriefing survey and conduct a short interview to discuss your responses and answer any questions you may have.

Orientation Training
ELP Training
Scenario
Post-trial survey (1 of 6)
Scenario
Post-trial Survey (2 of 6)
Post-Condition Survey (1 of 3)
ELP Training
Scenario
Post-trial survey (3 of 6)
Scenario
Post-trial Survey (4 of 6)
Post-Condition Survey (2 of 3)
ELP Training
Scenario
Post-trial survey (5 of 6)

Rests and bathroom breaks are allowed. We ask that you please do so in between conditions.


Scenario
Post-trial Survey (6 of 6)
Post-Condition Survey (3 of 3)
Debriefing Survey and Interview
Block 1
Block 2
Block 3

Getting Started

Please read the short briefing updating you on your current scenario. The yellow cog/circles on the map represent all the airports in the area. The red airplane represents your current position and the direction of the aircraft indicates the heading. The purple on the map represents the radar observed weather for the area. After you have finished reading the briefing, click the space bar to continue.

Please read the alert and then select “Start ELP” to continue.

Main Summary Page Navigation

On the left hand side of the screen, there are a list of randomized diversion options. The Emergency Landing Planner will provide the top five recommendations. The diversions will only be to the 6 airports in the Denver area (normally it includes all airports, but for this study, we are only including 6). ELP is generating diversion routes that are under 250 miles away from their current location. To view more information about each option, click on the option you would like to view. There are also no limits to how many times you can view each option. To return to the main screen, click the blue arrow located underneath the runway information.

On the right hand side of the main summary page, there will be a map with the weather and diversion paths. You will not be able to read the ATIS report for other airports that are not recommended by the ELP. If they are not recommended they are either too far, or not as safe as the other recommendations. The blue lines on the map represents the alternative flight paths that are on the left-hand side of the page. The yellow circles represent the airports that are included in the diversion paths. The red circle indicates the closure of airport that you either just departed from or were arriving too.

The weather can be removed from the map by clicking the blue button that says “remove weather”. To display the weather again, select “show weather”. There is no limit for how many times you can remove and show the weather. The map for each diversion option will have the weather, the alternative flight paths, and the yellow circle to indicate where the airport is located.

Tools and information

There will be tools that will be consistent in each condition. Each scenario will have the automatic terminal information service (ATIS) report, runway information, information about the airport facilities, and the approach plates for each runway. This information will always be on the individual diversion information pages.

ATIS Report

To display and remove the ATIS information, click on the ATIS button located at the top left of the screen when you are on the individual diversion information page. The ATIS is comprised of decoded METAR information about the airport. You still have the option to remove or show the weather on the map as well. You can click the ATIS button again to close the window.

Runway Information

The runway information is located on the left hand side of the screen directly under the ATIS button. The runway information consists of the runway approach, runway number, the runway length, distance, and the bearing.

Airport Facilities

The airport facilities for each diversion option are listed at the bottom left of the page. This information does NOT contribute to the calculations of the ELP. They have been color coded for ease of visibility. The different colors are used to represent excellent, good, fair, poor and unavailable airport facilities. The facility ratings will not be consistent for every scenario and have been manipulated for this study. Please look at the airport facility ratings for each scenario. If you are familiar to with the airports in the scenarios, please rely on the information provided and not your current knowledge. The ratings are as followed:

Excellent = Blue

Good = Green

Fair = Yellow

Poor = Orange

Unavailable = Red

Medical**:** Based on the medical facilities in the surrounding area and at the airport.

> An airport's medical is rated excellent if there is 100% likelihood of being able to satisfy medical needs. The airport has medical facilities and medical staff on site and hospitals in the surrounding area that are open 24/7. They are able to care for both major and minor injuries and illness. The facilities are large enough to accommodate for a large population.

> An airport's medical is rated good if there is 80% likelihood of being able to satisfy medical needs. The airport has medical facilities and medical staff on site and hospitals in the surrounding area, but are only open for business hours. They are able to care for major and minor injuries and illness, but there is a limited number of staff.

> An airport's medical is rated fair if there is 50% likelihood of being able to satisfy medical needs. The airport has medical facilities and medical staff on site to treat minor injuries and illnesses and are open for business hours only. The nearest hospital is an half an hour away from the airport, but they can't accommodate for a large population.

> An airport's medical is rated as poor if there is 25% likelihood of being able to satisfy medical needs. The airport has limited medical facilities and medical staff on site to treat some minor injuries and illnesses. They are only open when they have arriving and

departing flights. The nearest hospital is about an hour away, but can't accommodate for large populations.

An airport's medical is rated as unavailable if there is 0% likelihood of being able to satisfy medical needs. An airport would not have any medical facilities or medical staff available on site. The nearest hospital is over an hour away and only has the resources to treat minor illnesses and injuries.

***Maintenance*:** Determined by the aircraft maintenance facilities available at airport.

Maintenance would be rated as excellent if there is 100% likelihood that the airport could satisfy any maintenance needs. The airport has facilities that offer ground support for major and minor repairs. The facilities are open at all hours.

Maintenance would be rated as good if there is 80% likelihood that the airport could satisfy any maintenance needs. The airport has crews that have most parts in stock (won't need to fly over) and are capable of fixing major and minor repairs to the aircraft.

Maintenance would be rated as fair if there is 50% likelihood that the airport could satisfy any maintenance needs. The airport will most likely will not have the part you need in stock for repair. They are capable of doing minor repairs and some major repairs.

Maintenance would be rated as poor if there is 25% likelihood that the airport could satisfy any maintenance needs. Most likely, any part your plane would need would not be at the airport and they are only able to do minor repairs.

Maintenance would be rated as unavailable if there is 0% likelihood that the airport could satisfy any maintenance needs. The plane will not be able to be serviced at all at this airport. The plane will either have to be moved or crew, tools, and parts need to be flown over to the aircraft.

***Airline*:** Depends on the amount of support of your airline company at the airport.

Airline would be rated as excellent if there was a 100% likelihood the airline company services at the airport will satisfy the needs of the passengers. The airport has arriving and departing flights all throughout the day and night. Flights could easily be re-booked for passengers.

Airline would be rated as good if there was an 80% likelihood the airline company services at the airport will satisfy the needs of the passengers. The airport has arriving and departing flights only during business hours. Flights can be re-booked for passengers.

Airline would be rated as fair if there was a 50% likelihood the airline company services at the airport will satisfy the needs of the passengers. The airport has some flights arriving and departing throughout the day, but maybe not to passenger's original destination. Re-booking services would be limited and passengers would have long layovers.

Airline would be rated as poor if there was 25% likelihood the airline company services at the airport will satisfy the needs of the passengers. The arriving and departing flights at the airport are limited and it will be difficult for passengers to get flights towards their original destination.

Airline would be rated unavailable if there is 0% likelihood the airline company provides services at that airport. The company does not have any arriving or departing flights.

***Passenger*:** Based on convenience for passengers. Examples of these would be passenger's access to transportation, lodging, and services at the airport and in the area.

An airport would be rated at excellent if there was 100% likelihood that passenger's needs would be satisfied. Passengers have access to all transportation options (shuttles, rental cars, etc.) at all hours. They also have access to hotels near the airport. The airport has restaurants, fast-food, convenience stores, and bars located throughout.

An airport would be rated at good if there was 80% likelihood that passenger's needs would be satisfied. Passengers would have access to all transportation options (shuttles, rental cars, etc.) at all hours. They also have access to hotels near the airport. The airport restaurants, fast-foods, convenience stores, and bars are only open during business hours.

An airport would be rated at fair if there was 50% likelihood that passenger's needs would be satisfied. Passengers have some access to transportation options (shuttles, rental cars, etc.) but all are not all hours and may not be able to accommodate all passengers. Hotels and lodging options are not conveniently located to passengers. The airport may have limited restaurants, fast-foods, convenience stores, and bars are only open during business hours.

An airport would be rated at poor if there was 25% likelihood that passenger's needs would be satisfied. Passenger's access to ground transportation is limited. There is limited rental car service and shuttles are not always available. Hotels and lodging options are not conveniently located to passengers. The airport most likely does not have restaurants and stores available to the passengers.

An airport would be unavailable if there is 0% likelihood that passenger's needs would be satisfied. Passengers would not have any access to transportation or lodging and the airport would not have any services at the airport available to passengers.

***Fuel*:** Whether the airport has the required fuel for the aircraft.

Fuel would be rated as excellent if there is 100% likelihood of satisfying fueling needs. The airport would have the required jet fuel and fueling staff available at all hours.

Fuel would be rated as good if there is 80% likelihood of being able to satisfy fueling needs. The airport has the required jet fuel, but fueling services are only available during business hours.

Fuel would be rated as fair if there is 50% likelihood of being able to satisfy fueling needs. The airport may not have the required jet fuel and is only available during business hours.

Fuel would be rated as poor if there is 25% likelihood of being able to satisfy fueling needs. The airport may not have the required jet fuel and fueling arrangements need to be in advance for fueling service.

Fuel would be rated as unavailable if there was 0% likelihood of satisfying fueling needs. The airport would not have the required fuel for the aircraft and there would not be any

Condition 1

Condition 1 has limited features. For condition 1, you will be given the ATIS report, runway information, and airport facility ratings. For more information regarding the tools provided, see the section titled, *Tools and information*.

Completing the task

Please go through each option and look at all of the information provided even if you are already familiar with the airport. Airport facilities may be different from what you are used to and differ between scenarios. Also, please remember to discuss your reasoning out loud.

Complete your task worksheet. Give a rating of the safety of each diversion given. Also, please decide which runway you think is the 1st (best), 2nd, 3rd, 4th, and 5th (worst) option. Please complete the task worksheet before selecting “end scenario”.

Click “end scenario” located on the main summary page of runway options to complete the scenario.

Please notify me when you have reached the following screen and click the link to continue filling out the survey. While completing the survey, please continue to discuss aloud your decisions.

Prior to starting scenario, participant should be able to:

- ☐ Start ELP
- ☐ Show and remove weather
- ☐ Know how to maneuver between diversion options
- ☐ Show and remove ATIS report
- ☐ Know where airport information is located
- ☐ Know where Airport facilities are located on the screen and what they mean
- ☐ Know how to properly end the scenario

Condition 2

For condition 2, you will be given the ATIS report, runway information, airport facility ratings, and the risk statement. The risk statement is generated by the ELP. The ELP calculates this risk percentage by examining many factors (these are discussed in the introduction of the ELP). This statement represents the likelihood that you will successfully complete the approach and landing under current conditions on your first attempt. In the study's scenarios, you are the pilot of a flyable/healthy aircraft. Realistically, if you did not make the first attempt, you could attempt to land again or you could alter the diversion path. The risk statement does not represent the likelihood of you crashing.

The risk statement has been color coded by the percentage of success. This means if a diversion had an 87%, the ELP predicts that you will be able to successfully land the diversion 87% of the time. Again, it is important to remember that this represents the first attempt at a diversion and you would, in reality, have the choice to attempt the landing again or choose a different diversion. The percentage corresponds with a ranking (excellent, good, fair, poor, or unacceptable), which corresponds with a color. For example, if you were given a risk statement with 30%, it would be considered unacceptable and be color coded red.

100% - 90% = Excellent = Blue

89% - 76% = Good = Green

75% – 56% = Fair = Yellow

55% - 41% = Poor = Orange

40% and below = Unacceptable = Red

Completing the task

Please go through each option and look at all of the information provided even if you are already familiar with the airport. Airport facilities may be different from what you are used to and differ between scenarios. Also, please remember to discuss your reasoning out loud.

Complete your task worksheet. Give a rating of the safety of each diversion given. Also, please decide which runway you think is the 1st (best), 2nd, 3rd, 4th, and 5th (worst) option. Please complete the task worksheet before selecting "end scenario".

Click "end scenario" located on the main summary page of runway options to complete the scenario.

Please notify me when you have reached the following screen and click the link to continue filling out the survey. While completing the survey, please continue to discuss aloud your decisions.

Prior to starting scenario, participant should be able to:

- ☐ Start ELP
- ☐ Show and remove weather
- ☐ Know how to maneuver between diversion options

- Show and remove ATIS report
- Know where airport information is located
- Know where Airport facilities are located on the screen and what they mean
- Know where the risk statement is located and understand what it implies
- Know how to properly end the scenario

Condition 3

For condition 3, you will be given the ATIS report, runway information, airport facility ratings, the risk statement, and the reasoning statement for the risk statement. The risk statement is generated by the ELP. The ELP calculates this risk percentage by examining many factors (these are discussed in the introduction of the ELP). This statement represents the likelihood that you will successfully complete the approach and landing under current conditions on your first attempt. In the scenarios you are the pilot of a flyable/healthy aircraft. Realistically, if you did not make the first attempt, you could attempt to land again or you could alter the diversion path. The risk statement does not represent the likelihood of you crashing.

The risk statement has been color coded by the percentage of success. This means if a diversion had an 87%, the ELP predicts that you will be able to successfully land the diversion 87% of the time. Again, it is important to remember that this represents the first attempt at a diversion and you would, in reality, have the choice to attempt the landing again or choose a different diversion. The percentage corresponds with a ranking (excellent, good, fair, poor, or unacceptable), which corresponds with a color. For example, if you were given a risk statement with 30%, it would be considered unacceptable and be color coded red.

100% - 90% = Excellent = Blue

89% - 76% = Good = Green

75% – 56% = Fair = Yellow

55% - 41% = Poor = Orange

40% and below = Unacceptable = Red

In condition 3, you will also be given the reasoning statement for the risk statement. To view the reasoning statements, click the white arrow underneath the risk statement. On the next screen Enroute, Approach, and Runway will be visible. Next to each feature, there will be the ranking (Excellent, Good, fair, poor, or unacceptable) for each and the color that corresponds with that ranking.

Example:

ENROUTE: Good

APPROACH: Excellent

RUNWAY: Unacceptable

The reasoning statement explains how the ELP calculated that risk number. The ELP generates a numerical percentage (similar to the risk statement) for each factor listed earlier, which we then give a value (ex. 95% is excellent). The ELP calculates the interaction between all the factors taken into consideration. Each issue is not detrimental on its own, but the interaction with another issue can make that stage in the flight path unacceptable. For example, bad approach weather alone will not make the approach unacceptable, but bad weather and poor visibility will. If you scroll over Enroute, Approach, and Runway, you will be given the reasoning statement for why it received that ranking. For the enroute path, the ELP take the distance, turns, enroute weather and terrain into consideration. For the approach, it takes the approach ceiling, visibility, and weather. Finally, for the runway, it takes the runway length and width, landing speed, runway surface, and landing crosswind into its calculations.

Enroute	Approach	Runway
Enroute Distance	Approach ceiling	Runway Length
Enroute Turns	Approach Weather	Runway Width
Enroute Weather	Approach Visibility	Landing Speed
Terrain		Runway Surface
		Landing Crosswind

100% - 90% = Excellent = Blue

89% - 76% = Good = Green

75% – 56% = Fair = Yellow

55% - 41% = Poor = Orange

40% and below = Unacceptable = Red

If you would like to view the reasoning statement and the ATIS report at the same time, move the cursor over the risk factor you would like to see. Once the reasoning appears, move the cursor to the ATIS button without scrolling over any other reasoning factor.

To return to the risk statement, select the white arrow in the box.

Completing the task

Please go through each option and look at all of the information provided even if you are already familiar with the airport. Airport facilities may be different from what you are used to and differ between scenarios. Also, please remember to discuss your reasoning out loud.

Complete your task worksheet. Give a rating of the safety of each diversion given. Also, please decide which runway you think is the 1st (best), 2nd, 3rd, 4th, and 5th (worst) option. Please complete the task worksheet before selecting “end scenario”.

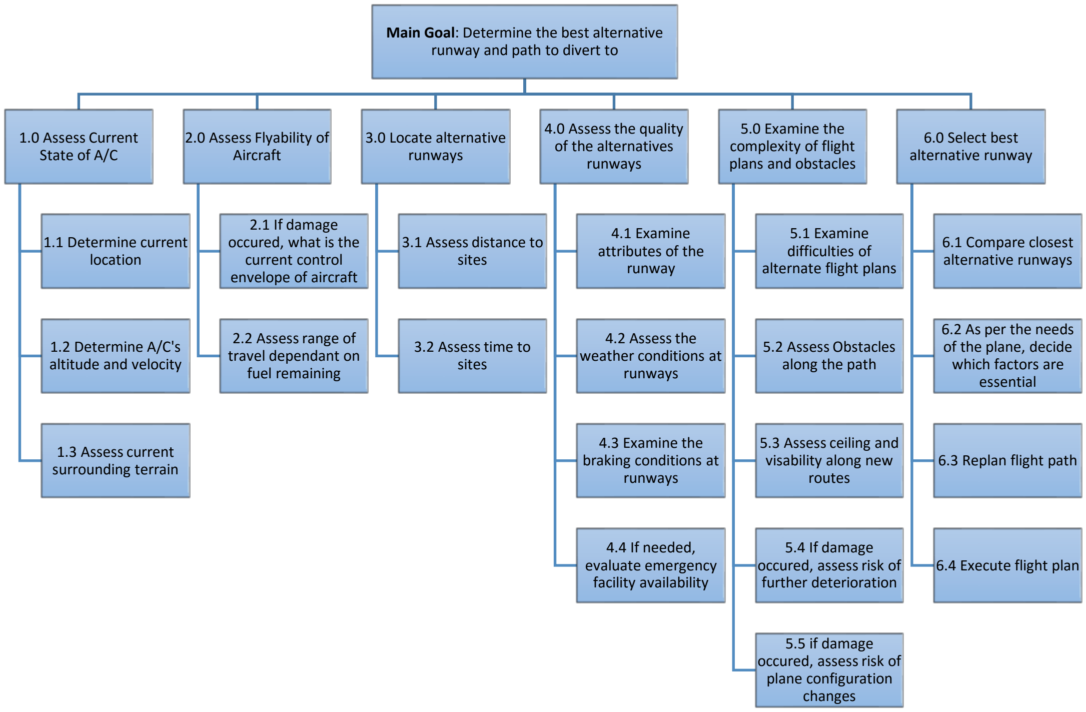
Click “end scenario” located on the main summary page of runway options to complete the scenario.

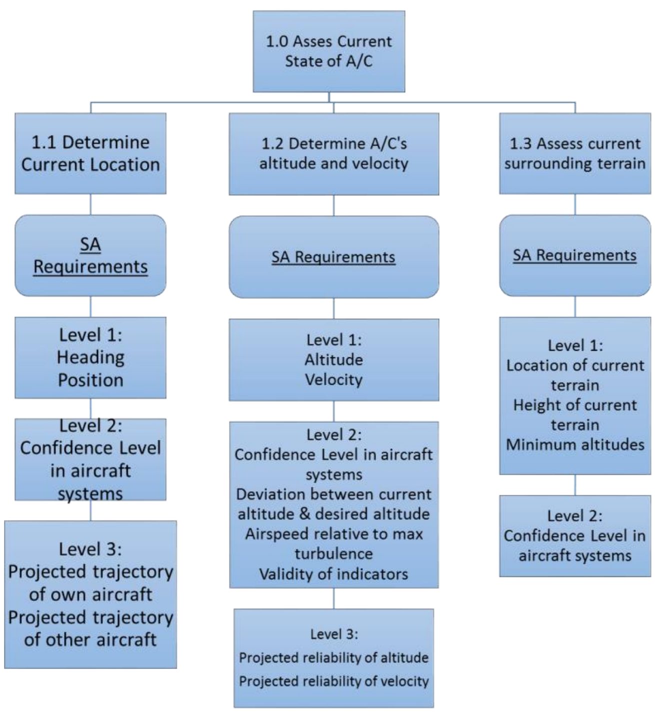
Please notify me when you have reached the following screen and click the link to continue filling out the survey. While completing the survey, please continue to discuss aloud your decisions.

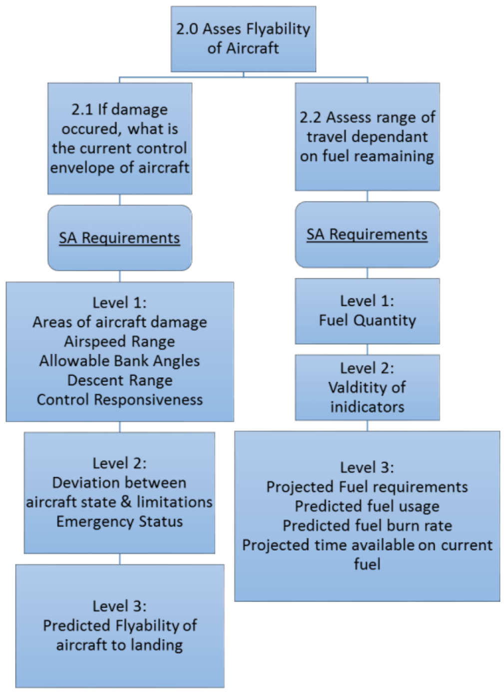
Prior to starting scenario, participant should be able to:

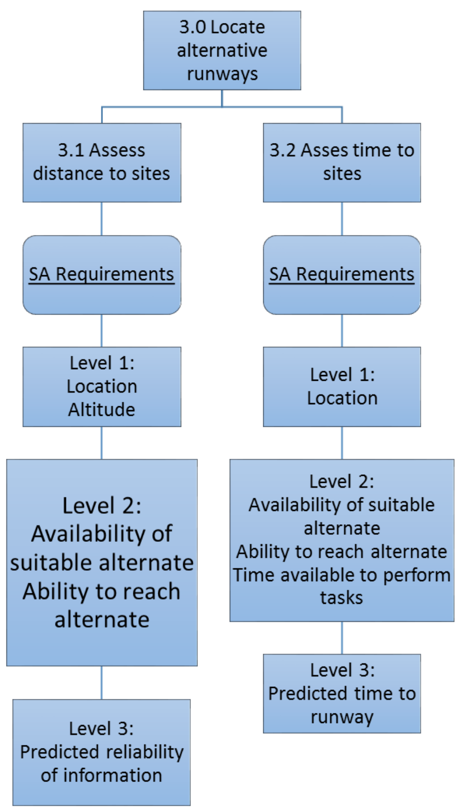
- ☐ Start ELP
- ☐ Show and remove weather
- ☐ Know how to maneuver between diversion options
- ☐ Show and remove ATIS report
- ☐ Know where airport information is located
- ☐ Know where Airport facilities are located on the screen and what they mean
- ☐ Know where the risk statement is located and understand what it implies
- ☐ Know where the reasoning statement is located and be able to understand how the Enroute, Approach and Runway
- ☐ Know how to properly end the scenario

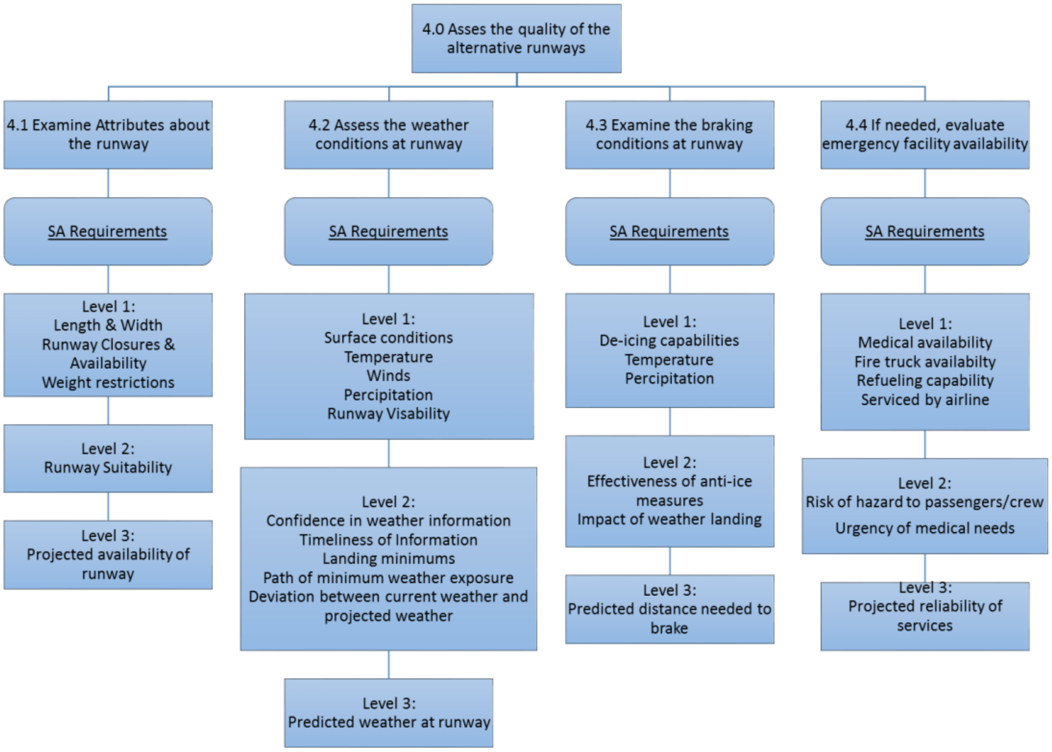
Appendix G: Task Analysis of Emergency Landing Procedure

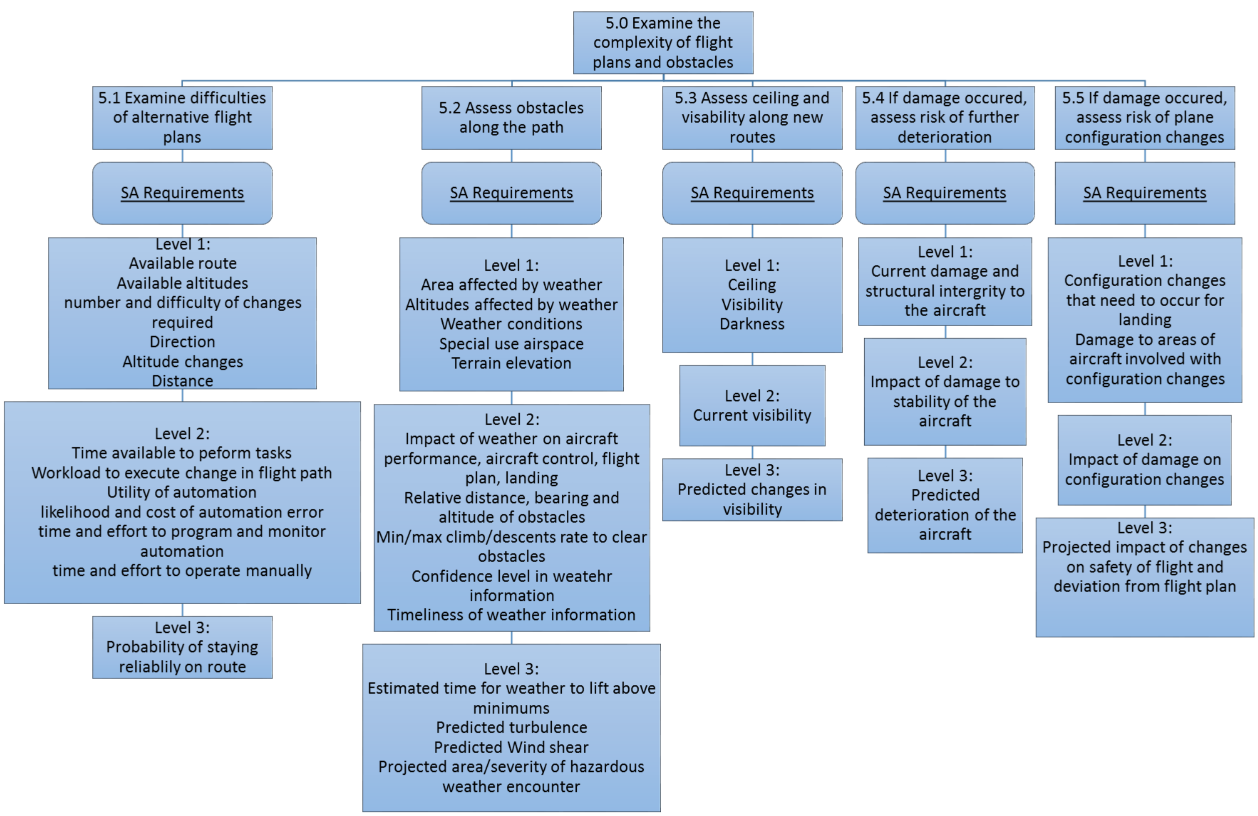


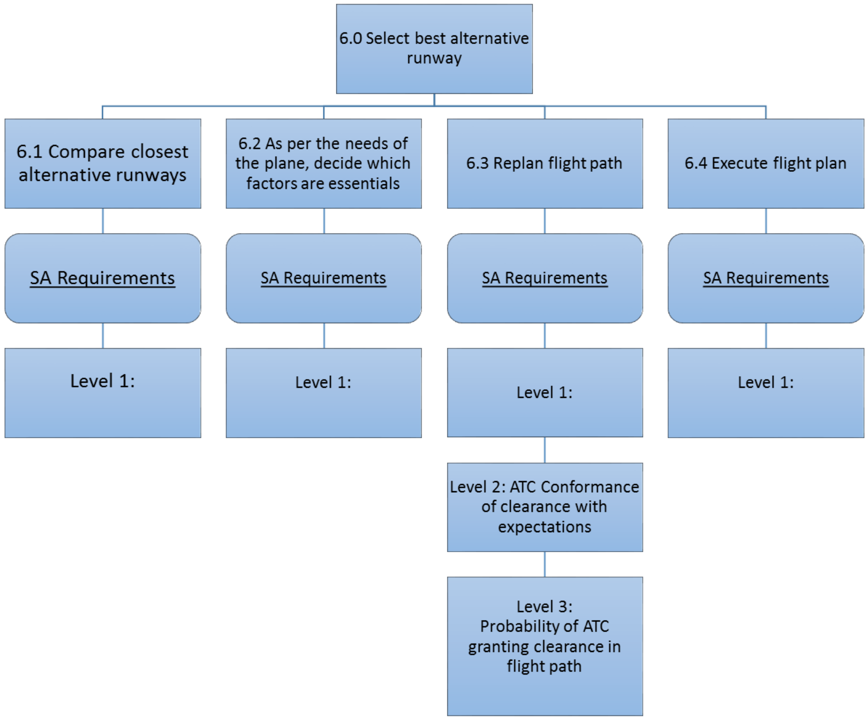



2.0 Asses Flyability of Aircraft
2.1 If damage occured, what is the current control envelope of aircraft
2.2 Assess range of travel dependant on fuel reamaining
SA Requirements
SA Requirements
Level 1: Areas of aircraft damage Airspeed Range Allowable Bank Angles Descent Range Control Responsiveness
Level 1: Fuel Quantity
Level 2: Valdity of inidicators
Level 2: Deviation between aircraft state & limitations Emergency Status
Level 3: Projected Fuel requirements Predicted fuel usage Predicted fuel burn rate Projected time available on current fuel
Level 3: Predicted Flyability of aircraft to landing


3.0 Locate alternative runways
3.1 Assess distance to sites
3.2 Asses time to sites
SA Requirements
SA Requirements
Level 1: Location Altitude
Level 1: Location
Level 2: Availability of suitable alternate Ability to reach alternate
Level 2: Availability of suitable alternate Ability to reach alternate Time available to perform tasks
Level 3: Predicted reliability of information
Level 3: Predicted time to runway


4.0 Asses the quality of the alternative runways
4.1 Examine Attributes about the runway
SA Requirements
Level 1: Length & Width Runway Closures & Availability Weight restrictions
Level 2: Runway Suitability
Level 3: Projected availability of runway
4.2 Assess the weather conditions at runway
SA Requirements
Level 1: Surface conditions Temperature Winds Percipitation Runway Visability
Level 2: Confidence in weather information Timeliness of Information Landing minimums Path of minimum weather exposure Deviation between current weather and projected weather
Level 3: Predicted weather at runway
4.3 Examine the braking conditions at runway
SA Requirements
Level 1: De-icing capabilities Temperature Percipitation
Level 2: Effectiveness of anti-ice measures Impact of weather landing
Level 3: Predicted distance needed to brake
4.4 If needed, evaluate emergency facility availability
SA Requirements
Level 1: Medical availability Fire truck availabilty Refueling capability Serviced by airline
Level 2: Risk of hazard to passengers/crew Urgency of medical needs
Level 3: Projected reliability of services


5.0 Examine the complexity of flight plans and obstacles
5.1 Examine difficulties of alternative flight plans
SA Requirements
Level 1: Available route Available altitudes number and difficulty of changes required Direction Altitude changes Distance
Level 2: Time available to peform tasks Workload to execute change in flight path Utility of automation likelihood and cost of automation error time and effort to program and monitor automation time and effort to operate manually
Level 3: Probability of staying reliablily on route
5.2 Assess obstacles along the path
SA Requirements
Level 1: Area affected by weather Altitudes affected by weather Weather conditions Special use airspace Terrain elevation
Level 2: Impact of weather on aircraft performance, aircraft control, flight plan, landing Relative distance, bearing and altitude of obstacles Min/max climb/descents rate to clear obstacles Confidence level in weatehr information Timeliness of weather information
Level 3: Estimated time for weather to lift above minimums Predicted turbulence Predicted Wind shear Projected area/severity of hazardous weather encounter
5.3 Assess ceiling and visability along new routes
SA Requirements
Level 1: Ceiling Visibility Darkness
Level 2: Current visibility
Level 3: Predicted changes in visibility
5.4 If damage occured, assess risk of further deterioration
SA Requirements
Level 1: Current damage and structural intergrity to the aircraft
Level 2: Impact of damage to stability of the aircraft
Level 3: Predicted deterioration of the aircraft
5.5 If damage occured, assess risk of plane configuration changes
SA Requirements
Level 1: Configuration changes that need to occur for landing Damage to areas of aircraft involved with configuration changes
Level 2: Impact of damage on configuration changes
Level 3: Projected impact of changes on safety of flight and deviation from flight plan


6.0 Select best alternative runway
6.1 Compare closest alternative runways
6.2 As per the needs of the plane, decide which factors are essentials
6.3 Replan flight path
6.4 Execute flight plan
SA Requirements
SA Requirements
SA Requirements
SA Requirements
Level 1:
Level 1:
Level 1:
Level 1:
Level 2: ATC Conformance of clearance with expectations
Level 3: Probability of ATC granting clearance in flight path