DOT/FAA/TC-19/37

Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

Air Traffic Control Decision Support Tool Design and Implementation Handbook

Carolina M. Zingale, Ph.D., FAA Human Factors Branch, ANG-E25
Brion Woroch, Ph.D., SAIC.

September 2019

Technical Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-19/37		

4. Title and Subtitle
Air Traffic Control Decision Support Tool Design and Implementation Handbook

5. Report Date
September 2019

6. Performing Organization Code
ANG-E25

7. Author(s)

Carolina M. Zingale, Ph.D., FAA Human Factors Branch, ANG-E25
Brion Woroch, Ph,D., SAIC

8. Performing Organization Report No.
DOT/FAA/TC-19/37

9. Performing Organization Name and Address
Federal Aviation Administration
Human Factors Branch
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

Sent Bill an email 12. Sponsoring Agency Name and Address
Federal Aviation Administration
Human Factors Research and Engineering Group
800 Independence Avenue, S.W.
Washington, DC 20591

13. Type of Report and Period Covered
Technical Note/Technical Report

14. Sponsoring Agency Code
ANG-C1

15. Supplementary Notes

16. Abstract

Objective: The purpose of the Handbook is to provide guidelines for the development of air traffic Decision Support Tools (DSTs) planned for use in the National Airspace System (NAS) and how to best train users to work with those tools. **Background**: DSTs are typically not 100% accurate or reliable. Nevertheless, they can provide valuable assistance to users by helping them evaluate, select, and implement effective solutions. To do so, DSTs must be designed appropriately so that the tools themselves do not become distracting or add to workload. **Method**: The guidelines provided in this document were derived from several sources including, 1) literature on automation support in air traffic control and other complex domains, 2) findings from previous studies conducted to investigate DST use by novices, and 3) information obtained from six air traffic controllers who participated in familiarization workshops on several DSTs. **Results**: The Handbook consists of three sections. The first section provides guidelines for DST user interface design. The second section provides guidelines for training users to work with DSTs effectively. The third section provides an overview of research on human-automation teamwork. Human-automation teamwork will become increasingly important as automated systems continue to advance and artificial intelligence capabilities become more sophisticated. **Conclusion:** The guidelines in the Handbook will help system developers design DSTs that enable users to build appropriate trust in the tools and allow users to intervene effectively if system failures occur. **Applications**: The Handbook will be useful to human factors practitioners and systems engineers in FAA acquisition, including requirements developers, training developers, and others developing and testing air traffic control (ATC) systems. The results will also be useful to individuals and agencies who look to FAA standards for human factors guidance for DST use and integration.

17. Key Words
Decision Support Tools
Air Traffic Control
Human Factors
User Interface Design
Training

18. Distribution Statement
This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	35	

Form DOT F 1700.7 (8-72) **Reproduction of completed page authorized**

THIS PAGE IS BLANK INTENTIONALLY.

Table of Contents

THIS PAGE IS BLANK INTENTIONALLY.

Acknowledgments

We thank the air traffic control subject matter experts who provided comments and feedback on their use of Decision Support Tools and the training methods associated with them that contributed to the guidelines summarized in this document. We thank Tanya Yuditsky, ANG-E25, for providing information on several of the Decision Support Tools discussed in this document and for identifying potential sources of air traffic subject matter experts.

Contact Information

Comments, suggestions, or questions on this document should be addressed to:

Human Factors Branch
ANG-E25
Building 28 Human Factors Building
WJHTC
Atlantic City Airport, NJ 08405

Human Factors Standards Program

The FAA's Human Factors Division (ANG-C1) under the Assistant Administrator for NextGen, Portfolio Management and Technology Development Office, has program management responsibility and coordinates the development and maintenance of FAA Human Factors standards that are levied in acquisition programs. ANG-C1 point of contact: Dan Herschler, 202-267-9853

THIS PAGE IS BLANK INTENTIONALLY.

Executive Summary

The purpose of this document is to provide guidelines for the development of air traffic Decision Support Tools (DSTs) planned for use in the National Airspace System (NAS) and how to best train users who will work with those tools. DSTs provide system users (air traffic controllers, traffic flow managers, technical operations personnel) with recommended solutions or methods to evaluate potential solutions before they are implemented. As more data, systems, and procedures are introduced into the NAS, system users will need to assimilate increasing amounts of information, potentially leading to higher levels of user workload and degraded decision-making. DSTs can provide valuable assistance by helping users evaluate, select, and implement effective solutions. However, to achieve the anticipated benefits, the DSTs must be designed appropriately and effective user training must be provided so that the tools themselves do not become distracting and do not add to user workload.

The guidelines in this Handbook come from several sources, including, 1) literature on automation support in air traffic and other complex domains, 2) findings from previous studies conducted to investigate DST use, and 3) information obtained from several air traffic controllers who participated in workshops presented by the FAA Air Traffic Organization (ATO) Program Management Organization (PMO) that provided familiarization for several emerging DSTs such as Pre-Departure Rerouting (PDRR) at air traffic facilities throughout the US during 2018.

The amount of research on DST design and training and their use in aviation and air traffic control is not extensive. However, the guidelines provided in this Handbook are based on the available information and focus on the system design and user training principles that will enable users to build trust in the systems and intervene effectively when system failures occur.

The Handbook consists of three sections. The first section provides guidelines for DST user interface (UI) design and human-computer interaction (HCI). The second section provides guidelines for training users to work effectively with DSTs. The third section provides a descriptive overview of research on human-automation teamwork. Human-automation teamwork will become increasingly important as automated systems continue to advance and as artificial intelligence capabilities become more sophisticated. System designers will need to consider the evolving roles of humans and machines in order to develop DSTs that function optimally.

THIS PAGE IS BLANK INTENTIONALLY.

1. INTRODUCTION

This Handbook provides guidelines for the development of air traffic Decision Support Tools (DSTs) planned for use in the National Airspace System (NAS) and how to best train users to work with those tools. DSTs provide system users (air traffic controllers, traffic flow managers, technical operations personnel) with recommended solutions or methods to evaluate potential solutions before they are implemented. As more data, systems, and procedures are introduced into the NAS, system users will need to assimilate increasing amounts of information, potentially leading to higher levels of user workload and degraded decision-making. DSTs can provide valuable assistance by helping users evaluate, select, and implement effective solutions. However, to achieve the anticipated benefits, DSTs must be designed so that the tools themselves do not become distracting and do not add to workload.

DSTs must be developed appropriately, and effective user training must be provided, to optimize DST effectiveness. The guidelines provided in this Handbook are designed to assist DST system developers in achieving those objectives. The guidelines are based on human factors design principles, on available research pertaining to DSTs in aviation and other complex domains, and on feedback obtained from several traffic flow management specialists who participated in the "2018 Traffic Flow Management System (TFMS) Workshops." The Workshops provided familiarization on several DSTs including Pre-departure Rerouting (PDRR) and Airborne Rerouting (ABRR) to more than 400 Traffic Management Coordinators (TMCs), Supervisory Traffic Management Coordinators (STMCs), and other operational personnel and staff at air traffic facilities throughout the US during 2018. We interviewed six volunteers from the Workshops who provided their comments on the usefulness and usability of several DSTs including, but not limited to, those covered in the Workshops, how easy or difficult DSTs are to work with, the effectiveness of training, and suggestions for improvements for DST design and implementation.

This handbook provides guidelines pertinent to DSTs within the context of general user interface design recommendations that promote usability and effectiveness. Usability guidelines have been covered extensively elsewhere, both for general applications (e.g., Nielsen, 2013; Norman, 2013; Shneiderman et al., 2018), as well as for FAA systems (Ahlstrom, 2016). DST design should begin by utilizing these "best practice" UI design principles. We provide a summary of these core principles at the beginning of Section 2.1 Design Guidelines for Decision Support Tool User Interface/Human-Computer Interaction (HCI).

DSTs differ from other software tools in that they do not simply provide information or data to the user. DSTs go further by providing solutions or recommendations to problems or by providing methods for evaluating different potential solutions. However, DST solutions are typically not 100% accurate or reliable. They utilize data that may be uncertain because the data are noisy or vary over time (e.g., weather predictions). Therefore, DSTs may not provide optimal solutions in all conditions, and that may cause the user to distrust the tool. Once users distrust a system, they are less likely to use it.

User trust is a critical component of effective DST use (Masalonis and Parasuraman, 1999). There are several ways that DST developers can help foster trust in the tool. One way is to provide information about the tool's reliability and, if applicable, the conditions under which the recommendations are expected to be more and less reliable. This information allows users to set appropriate expectations about tool effectiveness and when to use it. Another way to help foster trust in the DST is to provide users access to information about how the tool operates and arrives at recommendations. This transparency allows users to understand the factors that the DST has considered in arriving at a solution. Users may have to consider other factors in some situations. System transparency helps users maintain situation awareness so they can assess whether or not the tool is operating properly or if human intervention is required. It is critical to provide information about the tool's operating status so that the user is continuously aware as to whether the tool is operating normally or experiencing a failure or anomaly that requires user intervention. This is especially important for complex systems that have multiple modes of operation. If users are unaware of the system's current operating mode, its capabilities, and its operating status they will not be able to work with it effectively.

DST use and effectiveness are influenced by other factors, including user workload, experience level, and skill. Under lower workload conditions, the user presumably has more time and cognitive resources to evaluate DST recommendations and to consider alternative options. However, under higher workload conditions, the user may implement the DST recommendation more readily because time and cognitive resources are more limited. Additionally, a DST user with more skill, knowledge, and experience is able to more critically evaluate DST recommendations and has more resources to bring to bear on evaluating alternatives before implementing a solution. The interactions of these factors creates a complex environment in which to develop and train DSTs.

The guidelines in this Handbook come from sources that include a review of previous research on air traffic DSTs, part-task studies with air traffic novices (Woroch, Zingale, & Masalonis, 2017), and a review of available air traffic DST training materials (Zingale, Masalonis, Woroch, & Yuditsky, 2017). The part-task studies illustrate the complex nature of DST use. Those studies implemented a simplified traffic management rerouting task in which novice participants (no air traffic experience) were taught to reroute traffic on a computer-based software system that had been developed specifically for the study and was based on the Integrated Departure Route Planning Tool (IDRP), a concept that utilizes information about weather and traffic demand to provide alternative routing options (DeLaura, Underhill, Hall, & Rodriquez, 2012; Davison Reynolds & DeLaura, 2011). Participant workload was manipulated by including additional monitoring and messaging tasks that simulated the types of tasks performed by traffic managers. The DST created for the study provided one or three recommendations about which alternative route to select. The DST route recommendations were based on weighted criteria, including the effect of weather on the route, impact of arrival time at the aircraft destination, and route congestion. The participants had to decide whether to accept one of the recommended solutions or to evaluate the available routes on their own and make their own choice about which route to select. Selecting more efficient routes resulted in higher performance scores than selecting less efficient routes. Half of the participants in the study received supplementary training that informed them about the conditions under which the DST would be more and less reliable. The other half of the participants did not receive this training.

The results of the part-task study showed that DST reliability, participant workload, and training on DST reliability affected performance. Overall, higher workload decreased performance and increased the participants' reliance on the DST. Low DST reliability had little impact on performance when workload was low. But, when workload was high and tool reliability was low, performance was lowest. When workload was high, the participants who had been trained to understand DST reliability performed better than those who had not been trained. Information about tool reliability appears to be especially useful under workload-intensive situations, just the type of situations that are common in the air traffic environment. The results of the part-task study also indicated that when the DST was reliable, providing a single recommendation resulted in higher performance than providing multiple recommendations. When the automation was not as reliable, participants performed better when the DST made multiple recommendations. Once again, participant workload is a factor and adds to complexity. If workload is high, the participant has less time to review multiple recommendations. In summary, the results of Woroch et al.'s (2017) part-task study underscore the complex nature of factors that influence DST usefulness and effectiveness that must be considered when developing these systems.

A review of available computer and web-based training materials for some Traffic Flow Management (TFM) tools also provides input to the Handbook training guidelines (Zingale, Masalonis, Woroch, & Yuditsky, 2017). That review was based on principles of effective learning, memory, and problem solving, and the comments and feedback of air traffic subject matter experts. We found many positive and useful aspects of these training materials. They were generally easy to access and use, included helpful narratives and written summaries of key training points, and presented material in a manageable timeframe of about an hour or less. They also provided assessments at the end of sections or the end of the course so that the users could evaluate their knowledge.

Some of the negative aspects of these training materials involved their navigation structure that was sometimes cumbersome and required the user to return to the beginning of a previous section rather than back to a specific portion of the content. We also noted instances in which information coding (e.g., color-coding) was used inconsistently within an application, making it unclear as to whether different meanings were intended by the difference in coding. In addition, we found that these materials often presented the users with a series of steps required to perform an action but did not provide a context or explanation for the task objective that is essential for users to be able to apply the tool to their needs. Finally, the training materials did not typically provide information about tool reliability. The traffic flow management personnel we interviewed reported that they found the computer and web-based training materials generally useful as a first exposure to a new software tool, but that these materials were less useful for helping them understand the functionality of the tool and how it would meet their needs in the operational environment. For detailed training and a comprehensive understanding of the tool, users require hands-on experience and the ability to test how the tool functions for their purposes.

We separate the information in the Handbook into three sections. The first section provides Design Guidelines for Decision Support Tool User Interface/Human-Computer Interaction (Section 2.1). The second section provides Training Guidelines for Decision Support Tools (Section 2.2). The third section provides a summary of relevant research on human-automation teamwork which is becoming a more common context for DST use as automation and artificial intelligence capabilities evolve (Section 2.3).

Although there is not an extensive amount of research on DST design and training and their use in air traffic control, we provide the guidelines in this document with confidence as the best available to date. We remind those making use of these guidelines that each DST is unique as each is designed for a specific purpose and will make use of different data and algorithms to provide solutions and recommendations. Therefore, the guidance provided in this document must be utilized with the understanding that specific issues will be pertinent to some tools and not to others, or that involve issues that are not yet addressed by these guidelines. Each tool will have specific requirements needed to optimize its design, use, and training, and we recommend that DST designers work with human factors engineers to build and test design prototypes during system development so that key issues can be identified, addressed, and optimized early through iterative design and testing.

As DSTs and other automation tools become increasingly common in air traffic, we expect research to continue to address new issues and concepts and that guidelines for DST and automation development and deployment will need to be updated regularly. Metrics will need to be established to determine whether the DSTs are serving their intended purpose effectively. In other domains, such as medicine, clinical decision support tools also integrate vast sources of information (e.g., patient data, disease information) to make decisions about potential patient diagnoses and treatments, but they have yet to achieve their full potential and their impact on improving clinical outcomes has been questioned (Castaneda, Nalley, Mannion, Bhattacharyya, Blake, Pecora, et al., 2015). Care must be taken to ensure that the DSTs being developed and deployed are achieving their intended goals.

DSTs will continue to advance and to integrate greater levels of sophistication and capability. Users will be working with these systems differently than they typically do today, interacting with them more as teammates rather than simply delegating tasks to them. As the National Research Council's report on decision making and human-machine interaction states "…this area of inquiry is still in its infancy relative to where multi-disciplinary research could take it over the next generation" (National Research Council, 2014, p.1). Much remains to be learned about how automation will best interact with and support human users in solving problems and making decisions. With increasingly advanced and complex air traffic procedures and increasing amounts of data required to enable those procedures, humans and automation will need to work together to arrive at effective solutions.

2. DECISION SUPPORT TOOL GUIDELINES

2.1 Design Guidelines for Decision Support Tool User Interface/Human-Computer Interaction

Effective DST development must start with the implementation of core UI principles when designing the interface. Since the mid 1990's, the UI design principles put forward by Neilson, Norman, and Shneiderman have remained the foundations of effective UI design. The key principles, as summarized by Nielsen (2013), are:

1. **Visibility of system status**
   The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world**
   The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom**
   Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards**
   Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.

5. **Error prevention**
   Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. **Recognition rather than recall**
   Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use**
   Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. **Aesthetic and minimalist design**
   Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. **Help users recognize, diagnose, and recover from errors**
   Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation**
    Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

Another key resource for air traffic DST designers and developers is the Human Factors Design Standard (Ahlstrom, 2016) that includes a comprehensive summary of guidelines for FAA software and hardware systems as well as references from other government organizations including the Department of Defense and National Aeronautics and Space Administration (NASA). The Human Factors Design Standard includes important principles in many relevant categories such as screen design, coding, and navigation that provide a strong foundation on which to build DSTs. It also includes a section on "Decision Aids" (see Appendix A) that provides recommendations about when decision aids/tools should and should not be used and stipulates that the user should ultimately determine when and how to use them. This concept aligns with FAA Order JO7110.65X (FAA, 2017), Section 2-1-3 Procedural Preference which directs air traffic personnel to "Use automation procedures in preference to non-automation procedures when workload, communications, and equipment capabilities permit." This approach provides the flexibility necessary to address the specific circumstances at hand so that the most operationally effective actions can be taken.

The DST guidelines that follow expand on the basic UI principles. The guidelines include examples and references so that DST designers and developers will have additional information on which to base their decisions and so that they can identify sources of other relevant information as needed. The UI examples provided in this document are not intended to serve as recommendations for specific design elements (colors, graphical indicators, features, or formats) because those elements will depend on the specific tool needs and on the context in which the tool is used. The Human Factors Design Standard is a useful source for determining what to consider when initially considering design elements, including general guidelines on the use of color and color combinations to avoid.

2.1.1 Clearly designate recommended option(s) from non-recommended options.

Clearly designate the recommended option so that it stands out from non-recommended options and can be immediately identified. For example, when providing graphical reroute options, the recommended route should be presented in a different color, different brightness, or different line width than the non-recommended options.

If more than one option is recommended, clearly present the most to least recommended options in a hierarchical manner, such as in a list or by numerical designation. Clearly differentiate any recommended options from non-recommended options by separating them from one another such as in separate sections of a table or separate sections within a list. The objective is to make salient functions obvious (Ahlstrom, 2016; section 4.2) and to structure them logically (Ahlstrom, 2016; section 5.1.4.4).

2.1.2 Provide information about the reliability or certainty of recommendations to help build user trust.

Provide information about the reliability or certainty of the DST recommendations. Understanding when the tool is more and less reliable is necessary for users to develop appropriate levels of trust in the DST and enable them to use the tool effectively. Without reliability information, users are likely to distrust the tool if/when they encounter situations in which the recommendations do not appear useful. User trust in automation is difficult to build, but easy to break down, and once trust deteriorates users may abandon use of the tool altogether (Masalonis and Parasuraman, 1999). Information we obtained from air traffic subject matter experts who had participated in the TFMS Workshops also indicated that once a tool is found to be unreliable or faulty, trust and use are diminished.

Research has shown that people often trust automated systems at the start (Dzindolet, Peterson, Pomranky, Pierce, & Back, 2003), but that trust deteriorates quickly once the system exhibits errors. Dzindolet et al. found that user trust increased when the system provided explanations as to why an error occurred. Trust is complex, involving each individual's personal tendency to trust (dispositional trust) which is relatively stable over time, situation trust that involves the environment in which the tool is used (including the task, the system complexity, and user workload), and learned trust that develops over time (Hoff and Bahir, 2015). Although dispositional trust cannot be controlled, DST designers can help foster situation and learned trust by developing systems that behave consistently under specified conditions so that users can learn the circumstances under which they can expect the tool to provide more and less useful recommendations.

Presenting information about reliability, such as the likelihood of an event, is particularly beneficial for complex and workload intensive tasks (Sorkin, Kantowitz, and Kantowitz, 1988). For example, Thomas, Wickens, and Rantanen (2003) found that pilots using cockpit displays to evaluate potential future aircraft conflicts were more accepting of false conflict alerts when the alerts indicated severity based on time until the predicted conflict. The pilots accepted that the earlier an alert was presented, the greater the probability that it may prove to be false because the situation would be likely to change over time. Understanding that the conflict was less likely to occur enabled the users to accept those false alerts in the system more readily.

A tool that provides reliability information in the TFM environment is the Route Availability Planning Tool (RAPT). RAPT uses Corridor Integrated Weather System (CIWS) weather data to provide decision support for managing departures during convective weather. CIWS calculates (in real time), the accuracy of its past 30, 60, and 120 minute precipitation and echo tops forecasts. Some Traffic Managers have reported displaying forecast accuracy scores for the airport being managed to determine how much credence to give to the forecast. The accuracy scores can be depicted on the map or displayed in a table that can be moved independently to any position on the screen.

2.1.3 Provide reliability information to promote effective use of the tool.

Provide information about tool reliability to increase overall utility. Wiegmann (2002) demonstrated that users who do not have information about tool reliability may adopt ineffective strategies for use, thereby reducing the tool's overall utility. In Wiegmann's study, participants were unaware of the tool's actual reliability (accurate on 80% of trials). They could agree or disagree with the tool's recommendation on each trial. Some participants adopted a "maximization" strategy in which they accepted a high proportion of the tool recommendations (average of 95%) whereas other participants adopted a "probability-matching" strategy in which they accepted a lower proportion of tool recommendations (average of 82%) as they attempted to match the accuracy rate they encountered through use. The participants who utilized probability matching performed more poorly than those who utilized the maximization strategy. Poorer performance resulted because the participants who adopted the probability-matching strategy could only be correct 80% of the time if they agreed with the DST on the right trials. In order to be 80% correct overall, they would have to agree with the recommendation on every trial.

Reliability information can be depicted graphically in a number of ways, such as by using different colors to indicate a more reliable (green) or a less reliable (yellow) recommendation. Alternatively, bar or circle charts can depict different fill levels to indicate higher or lower reliability. If more than one option is provided, the options should be presented in a manner that clearly designates the highest from the lowest level of reliability so they can be readily differentiated from one another.We included several questionnaires during the experiment; after each reroute, after each scenario, and at the end of all scenarios. We provide screen shots of those surveys in Appendix D.

2.1.4 Provide a recommendation if it is at least 70% reliable.

DST recommendations are typically not 100% accurate or reliable (e.g., weather). However, research indicates that even less reliable automation may be better than no automation as long as a

70% threshold of reliability is achieved. Trapsilawati et al. (2015) found that 100% reliable automation for air traffic conflict resolution advisories resulted in better performance than less reliable automation. But, even the less reliable automation was successful at resolving conflicts and avoiding the creation of new conflicts 80% of the time. Both the reliable and the less reliable automation led to better performance than no automation. Research indicates that alerts that report the likelihood of an event with at least 70 – 75% accuracy are useful and provide benefit, particularly when workload levels and task demands are high (e.g., Dixon & Wickens, 2006).

2.1.5 Provide information about the conditions under which recommendations will be more and less reliable.

Reliability information should include information about specific conditions and circumstances under which the tool is expected to be more and less reliable. Woroch, Zingale, and Masalonis (2017) found that novice participants in a simulated traffic management study performed better on a rerouting task when they were provided with information as to when the DST recommendation was more likely to be correct. Woroch et al. provided text at the top of the display to indicate which algorithm was in use. The text was meaningful to participants who had been trained to understand the differences between the algorithms, but it was not meaningful or useful to participants who had not been trained to understand the distinction. The participants who were trained to understand that different rerouting algorithms performed better for aircraft traveling in one direction than the other (e.g., Algorithm X for westbound traffic; Algorithm Y for eastbound traffic) were better able to make efficient use of the simulated DST than participants who had not been trained.

Information about the conditions should be immediately available on the primary display, in a prominent location, so that it is highly salient. Users will have a direct understanding as to when the tool is more and less likely to provide a helpful recommendation and can use the tool more effectively.

2.1.6 Provide information about factors contributing to DST recommendations.

Provide the user access to information that indicates the factors involved in the DST recommendation (e.g., weather, sector congestion). This information helps users determine whether those factors are the most appropriate and how they are weighted in the decision. Understanding how the tool arrived at its recommendation will help users evaluate different options if the situation requires an alternative resolution. For example, for DST aircraft rerouting recommendations, the user should understand how the route recommendation was determined. The explanation may indicate that weather is the primary factor and that airport arrival time is also considered. Having this information allows the user to make an alternative decision if necessary (e.g., notification of airport arrival runway closure).

The information should be presented at the user's request, such as via selection of a display item or a menu option, so that it is available on demand rather than continuously presented. This helps to minimize display clutter. Such system "transparency" helps foster trust in the tool (Dadashi, Stedmon, and Pridmore, 2012) and improves task performance (Bass, Baumgart, & Shepley, 2013).

2.1.7 Provide a single recommendation when DST reliability is high, multiple recommendations (if available) when DST reliability is low.

Woroch, Zingale, and Masalonis (2017) found that when DST reliability was high on an aircraft rerouting task, providing a single recommendation led to better performance than providing three recommendations. However, when DST reliability was low, providing three recommendations led to better performance on the rerouting task. When tool reliability is lower, users likely need to consider additional factors to identify the most appropriate solution for the situation. Providing multiple possible solutions for them to consider may be useful in these circumstances.

2.1.7.1 Provide fewer, more reliable recommendations under higher workload conditions.

Woroch, Zingale, and Masalonis (2017) found that higher workload can decrease performance and increase reliance on DST recommendations on an aircraft rerouting task. Therefore, fewer and more reliable recommendations should be provided when workload is high. Of course, user workload is subjective and influenced by a number of factors such as the task itself and its complexity, other simultaneous tasks or subtasks the user is engaged in, and the complexity of the DST. DST designers can best help keep user workload manageable by ensuring that the design adheres to good general UI principles so that the DST interface is as easy to use as possible and information is presented in an optimized manner – key information is highlighted and easy to find, error messages are obvious, meaningful, and presented in a timely manner, and the state of the system is continuously conveyed (see also section 2.1.6). Additionally, DST designers should consider the context in which the DST is used so that they can gauge what user workload level is likely to be and provide recommendations accordingly. Alternatively, DST designers can provide more advanced implementations by providing a means by which the user can select a mode of operation based on his perceived workload level so that the number of recommendations provided can be tailored to the current needs.

2.1.8 Provide information about DST operational status to support situation awareness.

The DST should clearly indicate its operational status, current mode of operation (if applicable), and provide notifications of system failures or errors. To use DSTs effectively, users must maintain situation awareness. Situation awareness is a complex concept that involves comprehension of the current state of the system, the task, and the environment as well as the ability to project near-term future states (Endsley, 1995). Situation awareness has been shown to degrade with increasing levels of system automation (Sebok & Wickens, 2017). As automation increases, the user is often less directly engaged in the task, resulting in reduced situation awareness. Situation awareness can be supported by good UI design that ensures that required information is accessible to the user when needed and salient system status notifications or error messages and conditions are prominently displayed (Endsley & Jones, 2004). This requires that the DST provide dynamic updates of information so that the user is continuously aware of system status.

2.1.8.1 Provide notifications regarding system failures or errors.

The DST should provide clear indications of system failures or outages. Highly salient notifications should be used to capture the user's attention. To effectively intervene during a system failure, the user must have adequate situation awareness, and this can be challenging for users working with highly automated systems. As Sebok & Wickens (2017) reported, "In routine operations, an increased degree of automation supports performance, but in failure conditions, increased automation results in more significantly impaired performance"(p. 189). Sebok and Wickens noted that automation failures are often unexpected and can go unnoticed. If failure events are rare and unexpected, users may become complacent and may continue to rely on the automation without noticing that the failure has occurred. Therefore, it is critical that the system present highly salient cues (visual, auditory) to indicate that the system has identified a failure and that the human user must step in to take action. Higher workload will further degrade the user's ability to detect unexpected events as cognitive resources may be allocated elsewhere. Therefore, notifications such as a flashing red "No Data" message is an example of one way to help the user quickly redirect attention.

2.1.8.2 Provide a clear indicatation of system mode.

To help the user maintain situation awareness, DST systems that operate in multiple modes must provide a clear indication of the current mode of operation and notification of a mode change. If the system mode automatically changes, the user may not notice the change. This may result in the user expecting to have access to data, information, or processes that are now unavailable. If DST mode changes go unnoticed, the user may expect to receive a recommendation that the system is not currently able to deliver.

Numerous aviation accidents have been attributed to a lack of pilot awareness of aircraft mode in the automated cockpit (e.g., German and Rhodes, 2016). These mode changes often go unnoticed. Resnick (2012) found that if a change on a visual display goes unnoticed for more than 10s after it has occurred, it is unlikely to be detected at all ("change blindness"). Pilots have long been noted to experience difficulties identifying mode changes on the flight deck which has led to pilot confusion, unexpected system behaviors, and inappropriate system interaction (Lyall & Funk, 1998). Any automated systems that feature multiple operating modes can be susceptible to these problems.

Sebok, Wickens, Sarter, Quesada, Socash, & Anthony (2012) developed the "Automation Design Advisor Tool" (ADAT) to examine human performance models associated with various Fight Management Systems (FMS) to guide FMS design and help avoid user interface problems and difficulties with mode changes. Many of the ADAT guidelines center on fundamental UI guidelines such as presenting related information in close proximity on the display to minimize search time, placing highly salient or frequently used information in the user's direct line of sight, and clearly designating a change in mode by providing a different background or boarder around the information so that it appears noticeably different from other operational modes.

In the TFM environment, the Reroute Impact Assessment (RRIA) and the Collaborative Trajectory Options Program (CTOP) tools both have a "Model Mode" which allows Traffic Managers to view models of what would occur if a reroute or CTOP initiative was implemented. When in this mode, both tools provide a distinct appearance to the displays to raise users' awareness that what they are viewing is not the current, actual situation. Both use a yellow background color in textual displays and add a blue border along with a "Model Mode" label with a time stamp (indicating the time of the model) to the map.

It is important to note that higher workload levels also affect whether and how quickly users detect mode changes. In high workload situations, users may be focused on completing a specific task (e.g., "tunnel vision") and have less attention to devote to more subtle changes to the interface. Lack of knowledge of system mode can obviously have dire consequences if the user is expecting the system to perform a function that it is not currently in a state to perform.

2.1.9 Use terminology and symbology that is meaningful to the user and use the terms and symbols consistently.

A central tenet of effective UI design is to use standard terminology and symbology as they are used in the field and to use them consistently throughout the DST. Misuse or inconsistent use of terminology is confusing to system users and can cause them to respond erroneously to system output. Sebok et al. (2012) found the confusability of terms and symbols to be a key source of error for pilots in the automated cockpit. Unfamiliar or nonstandard symbols and abbreviations can cause confusion, uncertainty, and delay user actions. Air traffic subject matter experts we interviewed who had participated in the TFMS Workshops noted occasions in which they encountered this problem with some of the DSTs they have used. They noted that those situations are disruptive.

2.2 Training Guidelines for Decision Support Tools

Because training will typically be the users' first encounter with the DST, it is an opportunity to help them establish an understanding about the tool's purpose, objectives, the factors considered in the tool's decisions, and its reliability. Training is the optimal time to help users begin to develop their understanding of these critical features and to establish initial trust. "When operators lack knowledge about the purpose of a system or how it functions, they will likely have a more difficult time accurately aligning their trust to a system's real-time reliability." (Hoff and Bashir, 2015, pp.421-422).

2.2.1 Provide information about DST purpose.

Training materials should clearly state the purpose of the tool at the outset of training – what it is designed to accomplish and the problem it is trying to solve. Users need a clear understanding of what the DST is designed to do in order to help them develop appropriate expectations about its output.

2.2.2 Train users on DST logic and functionality.

Training materials should provide information about how the system operates and arrives at proposed recommendations. System transparency will help users maintain situation awareness and enable them to intervene more effectively when the DST experiences a failure. It will also help users understand why a DST solution may not be optimal and that another decision is needed. Information about the factors that are considered in the DST recommendation and the relative importance or weighting of the factors should be provided. Understanding how the system operates better enables users to "partner" with automation because they will better understand the system's priorities and how it arrived at its recommendation. For example, if weather is the top factor in a recommended traffic reroute, the system users may need to reevaluate whether to accept the recommendation if other circumstances are relevant or other priorities emerge (e.g., an airport closure). Such off-nominal circumstances would require a different approach and potentially make it inappropriate to select the solution provided by the DST. Training users to understand the system's functionality promotes an understanding about the type of response to expect from the system and when the response may not be useful.

2.2.3 Train users on DST reliability.

Training should provide information about DST reliability. This information should be explicitly addressed during training so that users have a solid foundation on which to build understanding and trust. Training users to understand tool reliability, including the conditions under which the tool is expected to be more and less reliable will help them establish trust in the DST (e.g., Dadashi, Stedmon, & Pridmore, 2012; Koustanaï, Cavallo, Delhomme and Mas, 2012; Masalonis, 2003). Training should include explanations about the conditions under which the tool is expected to be more and less reliable, and it should also include examples and practice problems that the users can work through to develop an understanding of the type of output provided and conditions in which they may need to consider alternative options. Such experience and practice will better prepare users for what to expect and what options to consider when working with the tool in the field..

2.2.4 Provide opportunities for hand-on learning.

Hands-on learning has long been regarded as a highly effective method of acquiring new information (e.g., Dewey, 1963). Hands-on learning affords users the opportunity to evaluate whether they have mastered the concepts and procedures presented during training as they work on simulated tasks. This active learning process (i.e., engagement) helps users achieve a deeper level of understanding, allowing them to make more effective and efficient use of the tool in the field. Hands-on learning is especially important for DST training because it allows users to gain a better understanding of tool reliability before using it in the field. Comments from air traffic subject matter experts we interviewed who had participated in the TFMS Workshops indicated that the opportunity for hands-on learning provided was highly beneficial in helping them develop a good understanding of the tools.

2.2.4.1 Focus on helping users understand the reasons for human-system interactions, rather than simply presenting how to execute the actions.

Web-based courses and computer-based instruction often illustrate the sequence of steps required to complete a task but do not explain why those steps need to be taken. Without providing the rationale for the actions, users will not acquire a deeper level of understanding and will not retain the information as effectively. Presenting information as to why the steps are executed helps users better understand the system's operations and enables them to better maintain situation awareness. A review of several air traffic computer-based and web-based courses found that there was often an emphasis on training a sequence of "button presses" rather than on describing why the options need to be selected (Zingale et al., 2017). Air traffic subject matter experts we spoke with who participated in the TFMS Workshops reiterated this concern and felt that some of the training packages they had worked with were limited in value because the instruction focused on how to execute steps ("buttonology") rather than on helping them understand why they needed to take the steps required.

2.2.4.2 Provide opportunities for users to test their knowledge.

During training, users should be provided the opportunity to test their knowledge to evaluate whether they are executing the correct actions and obtaining the expected output from the DST. Training should provide sample problems and feedback to help them identify any errors made as well as provide explanations as to why the selections were incorrect. Ideally, problems should closely reflect the types of situations that the user would encounter in the field, and should reflect a range of difficulty levels. Practice and feedback will help shape users' use of the tool and prepare them for solving problems using the tool in the field.

2.2.5 Train different user groups on DST capabilities that will be most useful in their specific work environment.

DSTs are often very complex and offer more capabilities than most users are likely to need. It is critical that training focus heavily on those capabilities that will most benefit the user's specific work environment. Air traffic controllers in different domains and at different facilities will have very different needs for the tool, so advanced, specialized training should be provided to help them become proficient with the most important system functions for their location and their operational needs. The DST users will be able to make better use of the tool in the field if they are informed and trained specifically on how it will best be used in their domain (e.g., en route or terminal) and at their facility.

In the TFM environment, the Air Traffic Control System Command Center (ATCSCC) uses the CTOP tool to implement CTOP initiatives while the ARTCCs use the tool to model and monitor initiatives. For this reason, there were two separate training modules for CTOP – one that focused on the ATCSCC users and another that focused on the ARTCC users. Thus, each user group was provided with the targeted information that would be most applicable to their needs.

2.2.5.1 Provide a forum for users from different facilities to share how they use the tool.

Training should provide a forum for users to share with others how they use different DST capabilities. These "lessons learned" can provide a helpful means to extend the utility of the tool beyond current use at a facility. Air traffic subject matter experts that we interviewed who had participated in the TFMS Workshops provided this suggestion because they want to understand how other facilities are using the tool, to discuss pitfalls they have encountered and potential solutions, and to determine what other DST functions may be useful to their operations. Because the DSTs are typically complex and have numerous capabilities, many of the functions may not be used unless the users learn what value those functions may have and in what contexts they are used.

2.2.6 Train users to manage system failures.

Provide users with training scenarios that allow them to manage a DST system or subsystem failure. Providing this type of training will allow users to more effectively intervene when a system failure or anomaly occurs and they must solve a problem without the DST (Sebok and Wickens, 2017). It is important to provide training that allows users to practice the skills necessary to recover from a system failure so that they can keep these skills current and arrive at an effective solution. Practice on these procedures will enable users to more readily step in and take effective action when the system encounters problems.

2.3 Consider Decision Support Tools to be "Team Players."

Sections 2.1 and 2.2 offered concrete advice on the design of DST interfaces and training procedures for those who will use the systems. This section presents a more theoretical discussion of a topic that we believe is informative to DST development: human-automation teamwork. Although there are fewer points of specific advice to provide, understanding and considering these concepts during development of DST design and training will enhance DST use and effectiveness. DSTs should be considered part of a problem-solving team that includes both the human and the automation. Cuevas (2007) defined human-automation teams as "the dynamic, inter-dependent coupling between one or more human operators and one or more automated systems requiring collaboration and coordination to achieve successful task completion"(p. B64). As automated systems continue to develop and include more advanced capabilities, this coupling will become increasingly important. Automated systems will be "complementing and augmenting human capabilities" (Cuevas, 2007, p. 1) as they evolve. Developing automation to partner with humans in solving problems leads to more effective tool development (Wilson & Dougherty, 2018). The evolution of such partnering is most apparent in fields like manufacturing, in which robots are already partnering with humans in the assembly process to manage jobs that require strength or repetition while the human handles tasks requiring dexterity and judgment. In the future, we may anticipate that intelligent systems will step in and take over for the human when the human encounters difficulty, much as the human takes over for the automation when the automation experiences a problem or failure today or as humans take over for one another when needed.

Automation often changes the way that tasks are accomplished. In air traffic, Corver and Aneziris (2015) found that flight-strip automation drastically decreases both the number of possible errors and the realistically probable errors when compared to conventional strip operations, affecting perception, judgment and decision-making, memory tasks, and action execution tasks. The addition of automation to a system alters the role of the human user. The move from paper flight strips to electronic strips provides one example of the way in which new automation altered the way controllers perform their job.

"For example, controlling without flight progress strips (relying more on the indications presented on the radar screen) asks controllers to develop and refine new ways of managing airspace complexity and dynamics. In other words, it is not the technology that gets transformed and the people who adapt. Rather, people's practice gets transformed and they in turn adapt the technology to fit their local demands and constraints." (Dekker & Woods, 2002, p. 6)

The resulting new "team" of the human controller and automated flight strips offered not just a performance improvement, but introduced new cognitive challenges and altered the overall ATC processes. The workload analysis by Corver and Anaziris (2015) found that the controller's mental workload was reduced in automated strip environments for tasks such as monitoring and tracking. However the reduction in workload in these areas allowed the user to spend more time detecting potential conflicts and optimizing traffic, resulting in the overall workload remaining about the same. The introduction of the automation improved performance, but also altered the types of tasks being performed. DST development efforts should consider not just how the automation can improve current ATC procedures but how their use may introduce new task demands. The role of the DST in these emergent task demands should be considered during tool development and training,

A traditional view of automation is that it should replace a human when the ability of the automation is greater than that of the human operator or to assign the tasks to whichever processor, human or automation, does the task better. Dekker & Woods (2002) suggest that there is a need to move beyond this "MABA" (man-is-better-at/machine-is-better-at) way of thinking. The idea that new technology, automation, and DSTs can simply substitute for a human operator in a one-to-one way oversimplifies the complexity of the tasks and decision-making processes required. Instead of a division of labor and developing new tools to replace tasks done by humans, the focus should be on developing tools that act as team players.

New DSTs will not simply improve existing abilities and processes, they will enable new abilities and processes to emerge. It is important that new automated tools are not simply considered incremental improvements to the current system or replacements for a single function, but as capabilities that may change the team dynamic, for better or worse. Langan-Fox et al. (2009) investigated the potential role of automation in air traffic management (ATM) by putting together a theoretical framework for human-automation teamwork that focused on the interplay between human and automation. Potential tasks such as targeting, monitoring, regulating, and tracking can be assigned to and performed by various team members in order to maintain overall situation awareness and form and execute plans. They emphasized that the team outputs or operating states that emerge from such teams would be greater than the simple sum of the individual player's contributions and are expected to enhance overall performance. These emergent states might result in a better team mental-model, enhanced team situation awareness, and increased team cohesion. Their conclusion was that teamwork coordination between humans and automation is critical in future ATM settings. However, this report provided only a conceptual framework which lacked specifics and concrete examples. Therefore, future empirical testing is required to obtain specific advice on DST design and training.

An example of the interplay between humans and automation is illustrated by the Reroute Impact Assessment (RRIA) tool. A traffic manager can use the tool to evaluate potential air traffic reroutes. The tool provides information about the reroute to the human user and informs the user's current awareness of the situation and updates the user's mental model. In turn, the updated mental model may spawn new ideas for routes that can be assessed or that may help the user decide on a course of action. The more a DST can be adaptable to the user's needs as situations unfold, the better the overall "team" performance will be.

The efficacy and utility of human-automation teamwork design is mostly theoretical since larger-scale empirical testing has not been performed. However, we feel that the theoretical underpinnings are strong and it is worth considering how DSTs can be conceptualized and designed as team players during design and training. The next section contains advice on how to design automation that focuses on the teamwork perspective.

2.3.1 Design automation to be good team players.

There are several factors that can enable automated DSTs to function as good teammates with human users. Three useful concepts to keep in mind during development of new DSTs are that good teammates – whether automated or human – provide mutual predictability, common ground, and directability (Klein et al., 2004).

Mutual predictability means that a team member can expect the type of information other team members will provide or the actions they will take.

"For example, in older, hardwired control centers, individual controllers can often infer what other controllers are working on just by observing which displays or control panels they are attending to." (Christoffersen & Woods, 2002, p. 5)

For a DST team member, mutual predictability means that the human user can expect the type of information or actions a DST will provide, just as a software system can "expect" the type of information a human user will provide. More complex tools can adapt their performance based upon the current situation. In providing DST user interface design guidance, we suggest that the ideal number of alternate recommendations provided by a DST depends upon its reliability and accuracy as well as the user's current workload (see section 2.1.4). The human user should be able to predict how the automation is likely to perform in certain situations. Over time, by being predictable, the user's trust in the automation will increase. Similarly, the automation needs to receive consistent input from its users in return. DSTs should either have the flexibility and adaptability to accept multiple types of inputs or different methods of use. If not, clear and explicit input rules and procedures should be established. The DST needs to "know" what type of information it will receive. The human user needs to be able to predict the type of supporting information a DST will provide.

One way to establish mutual predictability is for the automation to be observable. That is, the human user should be able to "look behind the curtain" to understand what it is that the automation is doing. The types of information that would be helpful in supporting system "transparency" are described throughout section 2.1. The following information helps users understand system operations and supports human-automation teamwork and situation awareness.

- The difficulty encountered in arriving at a solution
- The alternative solutions that were considered
- How close the system was to failing

- How optimal the solution is
- The reliability of the solution

Recommendations vary about how much information to provide the user. Simple output by the automation gives the human little information about how it is performing. However, the cognitive workload needed for a human to extract useful information from complex output may outweigh the benefits. Data availability does not equal informativeness; providing more information can do more harm than good if the information cannot be processed quickly and efficiently. The optimal method for facilitating effective and efficient feedback is not yet fully understood. However, there is evidence that providing information about the current reliability of the automation is useful (see sections 2.1.2 & 2.1.4). In addition, Christoffersen & Woods (2002) offer the following guidelines, stating that feedback from automation should be:

- Event-based — highlights when the system changes states and focuses on intermediate steps
- Future-oriented — communicates what steps the system will take next
- Pattern-based — displays DST's output quickly so that the user can pick up on abnormalities or unexpected conditions rather than have to mentally integrate individual pieces of information

Common Ground between human and automation refers to a set of instructions and procedures that include shared knowledge, beliefs, assumptions, and goals. In order to generate and maintain common ground, the human user needs to understand what the automation is trying to accomplish and how it is proceeding to the goal; essentially, to understand the situation from the automation's point-of-view. The goal is for the human and automation to have the same understanding of the problem space; what the problem and the solutions will look like. Common ground can be established by ensuring that the automation is observable and by providing feedback on operational status. Common ground can be established over time as the human learns more about how the automation operates. It can also be established by specifically addressing it when training new DSTs. The recommendations in the training section of this document can enhance both the mutual predictability and common ground between human users and automated DSTs (see section 2.2).

Directability refers to the human user's ability to override the automation as well as influence its operations. If humans are ultimately responsible for the outcome of the human-automation team, they will also need ultimate control over how problems are solved. Traditionally, if a DST is not providing useful information, the human user would switch the automation off and go into "manual" mode.

"For example, pilots in highly automated commercial aircraft have been known to simply switch off some automated systems in critical situations because they have either lost track of what the automation is doing, or cannot reconcile the automation's activities with their own perception of the problem situation. Rather than trying to sort out the state of the automation, they revert to manual or direct control as a way to reclaim understanding of and control over the situation. The uncooperative nature of the automated systems forces the pilots to buy this awareness and control at the price of abandoning the potentially useful functions that the automation performs, thus leaving them to face the situation unaided." (Christoffersen & Woods, 2002, p. 9)

Human users should be able to "substantially influence" automation's activity; perhaps by switching the automation's modes or changing the nature of its output (Christoffersen & Woods, 2002). We should move away from the idea of all-or-none automation (manual or automatic) and focus on objectives that rely on human-automation collaboration. It is not effective for the automation to simply be a passive tool for the human user, and it is not time-effective nor workload efficient for the human to take the role of micro-manager, controlling the work of several automated DSTs.

"What is required are intermediate, cooperative modes of interaction which allow human operators to focus the power of the automation on particular sub-problems, or to specify solution methods that account for unique aspects of the situation which the automated agent may be unaware of. In simple terms, automated agents need to be flexible and they need to be good at taking direction." (Christoffersen & Woods, 2002, p. 7)

The specifics on how to best accomplish good directability are not well understood. However, one framework that researchers Miller and Parasuraman (2007) found useful was a playbook analogy allowing human users to rapidly and effectively delegate tasks to automation. A sports playbook is a set of routines and actions to be undertaken by various team members at different times. Rather than delegate a task to each automated team member separately, one action or "play" can be initiated to assign roles and tasks to many DSTs. Following the advice from section 2.1.4, under high workload conditions, the human can call for a low-reliability DST to provide multiple options and for a high-reliability DST to provide one option. Under a different set of circumstances, the human could call a different "play" that causes the DST team-members to perform or behave differently. The human user gets the enhanced performance of flexible DSTs without the workload cost of micro-managing the behavior of each one individually.

3. SUMMARY

This Handbook provides guidelines for the design of air traffic DSTs in the NAS and how to best train users who will work with those tools.. The guidelines come from sources including: literature on automation support in air traffic and other complex domains, previous research on DST use and implementation, and on the comments and feedback about DSTs obtained from several air traffic controllers who participated in ATO PMO workshops that provided familiarization on several DSTs such as PDRR and ABRR throughout the US in 2018.

The first two sections of this Handbook focus on DST UI design and training guidelines that promote effective use of these tools. DSTs that follow these guidelines will help users develop an understanding of DST purpose and objectives, how DSTs arrive at recommendations, and help users establish trust in the tool. The third section of the Handbook provides an overview of existing research on human-automation teamwork. Human-automation teamwork will become increasingly important as DSTs continue to advance in complexity, particularly as increasing levels of automation and artificial intelligence are introduced. Overall, the future of DSTs will require that a philosophy of human-automation teamwork be adopted in order for successful systems to be developed. As automation capabilities continue to advance, subsequent research will enable increasingly specific guidance on DST development and training.

References

Ahlstrom, V. (2016). *Human Factors Design Standard* (DOT/FAA/HF-STD-001B). Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center. https://hf.tc.faa.gov/publications/2016-12-human-factors-design-standard/full_text.pdf

Bass, E. J., Baumgart, L.A., & Shepley, K.K. (2013). The effect of information analysis automation display content on human judgement performance in noisy environments. *Journal of Cognitive Engineering and Decision Making, 7*, 49-65. doi: 10.1177/1555343412453461

Castaneda, C., Nalley, K., Mannion, C., Bhattacharyya, P., Blake, P., Pecora, A., Goy, A., & Suh K.S. (2015). Clinical decision support systems for improving diagnostic accuracy and achieving precision medicine. *Journal of Clinical Bioinformatics, 5(4),* 1-16. http://dx.doi.org/10.1186/s13336-015-0019-3

Christoffersen, K., & Woods, D. D. (2002). How to make automated systems team players. In *Advances in human performance and cognitive engineering research*. United Kingdom: Emerald Group Publishing, Limited. http://dx.doi.org/10.1016/S1479-3601(02)02003-9

Corver, S. C., & Aneziris, O. N. (2015). The impact of controller support tools in enroute air traffic control on cognitive error modes: A comparative analysis in two operational environments. *Safety Science*, *71*, 2-15. http://dx.doi.org/10.1016/j.ssci.2014.07.018

Cuevas, H. M., Fiore, S. M., Caldwell, B. S., & Strater, L. (2007). Augmenting team cognition in human-automation teams performing in complex operational environments. *Aviation, Space, and Environmental Medicine, 78(5)*, B63-B70.

Dadashi, N., Stedmon, A. W., & Pridmore, T. P. (2012). Semi-automated CCTV surveillance: The effects of system confidence, system accuracy and task complexity on operator vigilance, reliance, and workload. *Applied Ergonomics, 44*, 730–738. http://dx.doi.org/10.1016/j.apergo.2012.04.012

Davison Reynolds, H. J., & DeLaura, R. A. (2011). *Concept of operations for the Integrated Departure Route Planning (IDRP) tool* (Project Report No. ATC-379). Lexington, MA: Lincoln Laboratory, Massachusetts Institute of Technology. https://www.ll.mit.edu/sites/default/files/publication/doc/2018-12/Reynolds_2011_ATC-379_WW-22138.pdf

DeLaura, R. A., Underhill, N. K., Hall, L. M., & Rodriguez, Y. G. (2012). *Evaluation of the Integrated Departure Route Planning (IDRP) tool 2011 prototype* (Project Report No. ATC-388). Lexington, MA: Lincoln Laboratory, Massachusetts Institute of Technology. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.295.527&rep=rep1&type=pdf

Dekker, S. W., & Woods, D. D. (2002). MABA-MABA or abracadabra? Progress on human–automation co-ordination. *Cognition, Technology & Work, 4(4)*, 240-244. doi:10.1007/s101110200022

Dewey, J. (1963). *Democracy and Education*. New York: MacMillan. http://dx.doi.org/10.1086/453309

Dixon, S.R. & Wickens, C.D. (2006). Automation reliability in unmanned aerial vehicle flight control: A reliance-compliance model of automation dependence in high workload. *Human Factors, 48*, 474-486. doi: 10.1518/001872006778606822

Dzindolet, M.T., Peterson, S.A., Pomranky, R.A., Pierce, L.G., & Beck, H.P. (2003). The role of trust in automation reliance. *International Journal of Human-Computer Studies, 58*, 697-718. http://dx.doi.org/10.1016/S1071-5819(03)00038-7

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 1(37),* 32-64. http://dx.doi.org/10.1518/001872095779049543

Endsley, M. & Jones, D.G. (2004). *Designing for Situation Awareness: An Approach to User Centered Design, 2nd ed.* Taylor & Francis: Boca Raton, FL. http://dx.doi.org/10.1201/9780203485088

Federal Aviation Administration. (2017). Air Traffic Control (JO 7110.65X). Retrieved from http://www.faa.gov/regulations_policies/orders_notices/index.cfm/go/document.informatio n /documentID/1031850

German, E,S. & Rhodes, D.H. (2016). Human-model interactivity: what can be learned from the experience of pilots with the glass cockpit? *2016 Conference on Systems Engineering Research,* Massachusetts Institute of Technology (MIT). From: http://seari.mit.edu/documents/preprints/GERMAN_CSER16.pdf

Hoff, K. & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors*, *57,* 407-434. http://dx.doi.org/10.1177/0018720814547570

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffman, R. R., & Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *IEEE Intelligent Systems, 19(6)*, 91-95. http://dx.doi.org/10.1109/MIS.2004.74

Koustanaï, A., Cavallo, V., Delhomme, P., & Mas, A. (2012). Simulator training with a forward collision warning system effects on driver-system interactions and driver trust. *Human Factors*, *54*, 709–721. https://doi.org/10.1177%2F0018720812441796

Langan-Fox, J., Canty, J. M., & Sankey, M. J. (2009). Human–automation teams and adaptable control for future air traffic management. *International Journal of Industrial Ergonomics*, *39(5),* 894-903. http://dx.doi.org/10.1016/j.ergon.2009.04.002

Lyall, B. & Funk, K. (1998). Flight Deck Automation Issues. In M.W. Scerbo & M. Moulousa (eds.) *Proceedings of the Third Conference on Automation Technology and Human Performance.* Norfolk, VA. https://doi.org/10.1207/s15327108ijap0902_2

Masalonis, A.J. (2003). Effects of training operators on situation specific automation reliability. *IEEE International Conference on Systems, Man and Cybernetics*, *2,* 1595–1599. http://dx.doi.org/10.1109/ICSMC.2003.1244640

Masalonis, A. J., & Parasuraman, R. (1999). Trust as a construct for evaluation of automated aids: Past and future theory and research. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting,* 184–188. http://dx.doi.org/10.1177/154193129904300312

Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors*, *49(1),* 57-75. http://dx.doi.org/10.1518/001872007779598037

National Research Council (2014). *Complex Operational Decision Making in Networked Systems of Humans and Machines: A Multidisciplinary Approach.* The National Academies Press. Washington, D.C. doi: 10.17226/18844

Nielsen, J. (2013). Ten Usability Heuristics for User Interface Design. From: https://www.designprinciplesftw.com/collections/10-usability-heuristics-for-user-interface-design

Norman, D. (2013). *The Design of Everyday Things*: Revised and Expanded Edition. Basic Books. New York, New York.

Resnick, R.A. (2002). Change detection. *Annual Review of Psychology*, *53*, 245-277. https://doi.org/10.1146/annurev.psych.53.100901.135125

Shneiderman, B., Plaisant, C., Cohen, M. Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2018). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (6th Edition). Pearson Education Limited: Essex, England. http://dx.doi.org/10.1145/25065.950626

Sebok, A. and Wickens, C. (2017). Implementing Lumberjacks and Black Swans Into Model-Based Tools to Support Human–Automation Interaction. *Human Factors, 59 (2),* 189–203. http://dx.doi.org/10.1177/0018720816665201

Sebok. A., Wickens, C., Sarter, N., Quesada, S., Socash, C., & Anthony, B. (2012).The Automation Design Advisor Tool (ADAT): Development and Validation of a Model-Based Tool to Support Flight Deck automation Design for NextGen Operations. *Human Factors and Ergonomics in Manufacturing & Service Industries. 22 (5)*, 378-394. http://dx.doi.org/10.1002/hfm.20389

Sorkin, R.D., Kantowitz, B.H., and Kantowitz, S.C. (1988). Likelihood Alarm Displays. *Human Factors, 30 (4),* 445-459. http://dx.doi.org/10.1177/001872088803000406

Thomas, L., Wickens, C., and Rantanen, E. (2003). Imperfect Automation in Aviation Traffic Alerts: A Review of Conflict Detection Algorithms and Their Implications for Human Factors Research. *Proceedings of the Human Factors and Ergonomics Society, 47th Annual Meeting.* http://dx.doi.org/10.1177/154193120304700320

Wiegmann, D.A., (2002). Agreeing with Automated Diagnostic Aids: A study of users' concurrence strategies. *Human Factors and Ergonomics Society, 44 (1),* 44-50. http://dx.doi.org/10.1518/0018720024494847

Wilson, H.J. & Dougherty, P.R. (2018). *Collaborative Intelligence: Humans and AI are Joining Forces.* Harvard Business Review, July-August. https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces

Woroch, B., Zingale C.M., & Masalonis, A. J. (2017). *Traffic Flow Management Tools: Guidance for Use, Integration, and Training: Part Task Experiment 1* (DOT/FAA/TC-17/37). Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center. https://hf.tc.faa.gov/publications/2017-09-01-traffic-flow-management-tools-phase2/TC-17-36.pdf

Zingale C.M., Masalonis, A. J., Woroch, B., & Yuditsky, T. (2017). *Traffic Flow Management Tools: Guidance for Use, Integration, and Training: Phase II – Human Factors Assessment of Computer and Web-based Training for Decision-Support Tools* (DOT/FAA/TC-17/36). Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center. https://hf.tc.faa.gov/publications/2017-09-01-traffic-flow-management-tools-phase2/TC-17-36.pdf

Acronyms

ABRR	Airborne Rerouting
ADAT	Automation Design Advisor Tool
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATM	Air Traffic Management
ATO	Air Traffic Organization
CTOP	Collaborative Trajectory Options Program
CWIS	Corridor Integrated Weather System
DST	Decision Support Tool
FAA	Federal Aviation Administration
FMS	Flight Management System
HCI	Human-Computer Interaction
IDRP	Integrated Departure Route Planning Tool
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
PDRR	Pre-departure Rerouting
PMO	Program Management Organization
RAPT	Route Availability Planning Tool
RRIA	Reroute Impact Assessment
SA	Situation Awareness
STMC	Supervisory Traffic Management Coordinator
TFM	Traffic Flow Management
TFMS	Traffic Flow Management System
TMC	Traffic Management Coordinator
UI	User Interface

Appendix A

Section 5.1.14 Decision Aids, from the Human Factors Design Standard (Ahlstrom, 2016). (References for works cited are provided in the HFDS)

Definition. Decision aids (sometimes referred to as decision support systems) are automated systems that provide support to human decision-making processes either unsolicited or by user request. Decision aids can narrow the decision alternatives to a few or suggest a preferred decision based on available data. [Source: Wiener, 1988]

- **5.1.14.1 When to use.** Decision aids should be used
  a. for managing system complexity;
  b. for assisting users in coping with information overload;
  c. for focusing the user's attention;
  d. for assisting the user in accomplishing time-consuming activities more quickly;
  e. when limited data results in uncertainty;
  f. for overcoming human limitations that are associated with uncertainty, the emotional components of decision-making, finite-memory capacity, and systematic and cognitive biases; and for assisting the user in retrieving, retaining, representing or manipulating large amounts of information, combining multiple cues or criteria, allocating resources, managing detained information, performing computations, and selecting and deciding among alternatives. [Source: AHCI, 1998; DISA, 1996]
- **5.1.14.2 When to avoid.** Decision aids should not be used
  g. when solutions are obvious;
  h. when one alternative clearly dominates all other options;
  i. when there is insufficient time to act upon a decision;
  j. when the user is not authorized to make decisions; or
  k. for cognitive tasks in which humans excel, including generalization and adapting to novel situations. [Source: AHCI, 1998]
- **5.1.14.3 Let users determine decision aid use.** Users should be able to determine when and how the decision aid should be used. [Source: Parasuraman & Riley, 1997]
- **5.1.14.4 Use terms and criteria appropriate to users.** Decision aids should use terminology and criteria appropriate to the target user group. [Source: DISA, 1996]
- **5.1.14.5 Reduce number of response options.** Decision aids should reduce the number of response options. [Source: Barnes, 1985]

  Discussion. The number of options that the user must consider is expected to decrease when a decision aid is used. Reducing the response options focuses the user's attention onto the most viable options.

  5.1.14.6 Assist user decisions. Decision aids should assist, rather than replace, human decision makers by providing data for making judgments rather than commands that the user must execute. [Source: AHCI, 1998; DISA, 1996; Parasuraman & Riley, 1997]
- **5.1.14.7 Make support consistent with mental models.** The support provided by decision aids should be consistent with user cognitive strategies and expectations (mental models). [Source: NUREG 0700, 2002]

  Definition. A **mental model** is an individual's understanding of the processes underlying system operation. [Source: NRC, 1998; Parasuraman & Mouloua, 1996]
- **5.1.14.8 Do not cancel ongoing user tasks.** Use of decision aids should not require ongoing user tasks to be cancelled. [Source: NUREG 0700, 2002]
- **5.1.14.9 Minimize query of user.** Decision aids should minimize query of the users for information. [Source: NUREG 0700, 2002]
- **5.1.14.10 Minimize data entry.** Decision aids should minimize user data entry requirements. [Source: DISA, 1996**]**

- **5.1.14.11 Provide ability for planning strategy or guiding process.** Decision aids should be capable of planning a strategy to address a problem or guide a complex process. [Source: NUREG 0700, 2002]
- **5.1.14.12 Accept user direction.** Decision aids should accept direction from the users on which problem solving strategy to employ when alternative strategies are available. [Source: NUREG 0700, 2002]
- **5.1.14.13 Prioritize alternatives.** When more than one alternative is available, the decision aid should provide the alternatives in a recommended prioritization scheme based on mission and task analysis. [Source: AHCI, 1998]
- **5.1.14.14 Alert user when unable to process.** Decision aids should alert the user when a problem or situation is beyond its capability. [Source: NUREG 0700, 2002]
- **5.1.14.15 Be flexible in type and sequence of input accepted.** Decision aids should be flexible in the types and sequencing of user inputs accepted. [Source: NUREG 0700, 2002]
- **5.1.14.16 Estimate uncertainty and rationale.** Decision aids should estimate and indicate the certainty of analysis and provide the rationale for the estimate. [Source: NUREG 0700, 2002]
- **5.1.14.17 Make derived or processed data accessible.** When information used by a decision aid is derived or processed, the data from which it is derived should be either visible or accessible for verification. [Source: Billings, 1996]

  Discussion. Data that are not critical for operation can be made available only upon request.
- **5.1.14.18 Provide hard copy of decision aid use.** The user should be able to obtain hard copy print outs of data including screen displays, rules and facts, data employed, hypotheses tested, and summary information. [Source: NUREG 0700, 2002]
- **5.1.14.19 Allow access to procedural information.** Decision aids should give the user access to procedural information used by the aid. [Source: Morris, Rouse & Ward, 1985; NUREG 0700, 2002]
  - **Discussion. Procedural information** is information about the rules or algorithms used by the decision aid. Knowledge of procedural information fosters user acceptance of the aid because the user is able to understand how the aid functions. As the user becomes more familiar with a given situation, he or she requires less procedural information. [Source: Morris, Rouse & Ward, 1985]
- **5.1.14.20 Provide user controlled level of explanation detail.** When the system provides explanations to the user, it should supply a short explanation initially, with the ability to make available more detail at the user's request, including access to process information or an explanation for the rules, knowledge-basis, and solutions used by the decision aid. [Source: DISA, 1996; NUREG 0700, 2002]

  Discussion. Process information is the information about how the aid accomplishes a task. This information is required by users to decide whether to use the aid in unfamiliar situations and for identifying the nature and extent of malfunctions. [Source: Morris et al., 1985]
- **5.1.14.21 Provide clear explanations to user.** When the system provides explanations to the user, the explanation should use terms familiar to the user and maintain consistency with the immediate task. [Source: DISA, 1996]
- **5.1.14.22 Present information with appropriate detail.** Decision aids should present information at the level of detail that is appropriate to the immediate task, with no more information than is essential. [Source: AHCI, 1998]
- **5.1.14.23 Avoid repeated information.** Decision aids should avoid repeating information that is already available. [Source: AHCI, 1998]
- **5.1.14.24 Integrate decision aids.** Decision aids should be fully integrated and consistent with the rest of the computer-human interface. [Source: NUREG 0700, 2002]
- **5.1.14.25 Alert to newly available information.** Decision aids should alert the user to changes in the status of important system information such as when critical information becomes available during decision aid utilization. [Source: NUREG 0700, 2002]

  Discussion. Critical information in this standard refers to information that may have a significant impact on task completion.
- **5.1.14.26 Alert to meaningful events or patterns.** Decision aids should automatically notify the user of meaningful patterns or events such as when it predicts a future problem. [Source: AHCI, 1998]
- **5.1.14.27 Predict based on historical data.** Decision aids should be able to predict future data based on historical data and current conditions. [Source: AHCI, 1998]

- ▫ **5.1.14.28 Provide ability to represent relationships graphically.** Decision aids should be able to graphically represent system relationships, its rules network, and reasoning process. [Source: NUREG 0700, 2002]
- ▫ **5.1.14.29 Identify simulation mode.** When decision aids have a simulation mode, entering the simulation mode should require an explicit command and result in a distinguishable change in output. [Source: NUREG 0700, 2002]
- ▪ **5.1.14.30 Provide knowledge of intent.** Each element in an intelligent human-machine system shall have knowledge of the intent of the other elements. [Source: Billings, 1996; NRC, 1998; Parasuraman et al., 2000]

  Discussion. Monitoring of the system by the user and the user by the system can only be effective if each knows what the other one is trying to accomplish. [Source: Billings, 1996]
- ▫ **5.1.14.31 Adapt with situational demands.** When adaptive decision aiding is used, the level of decision aiding should change with the situational demands in order to optimize performance. [Source: Rouse, 1988]

  Discussion. The criticality of a given task can change dramatically depending on the current situation.
  - ▫ **5.1.14.32 Adaptive decision aiding implementation.** Adaptive decision aiding should be applied when resource loading, performance, error frequency, and deviations from intent exceed threshold levels. [Source: Andes, 1987]
    - ▫ **Discussion.** Resource loading, performance, errors, and deviations from intent can be used as indicators to determine when the user might need the help of the automated decision aid. The threshold levels of these indicators, specifying the optimal time to implement decision aiding may need to be determined on a system-by-system basis, possibly through simulation.
- ▫ **5.1.14.33 Provide planning assistance.** Adaptive decision aiding interfaces should allow the user to receive direct assistance in planning how to carry out the intended task. [Source: Tyler & Treu, 1989
- ▫]**5.1.14.34 Allow user to initiate automation implementation.** The user should be able to initiate automated aids even if system-initiated automation is the norm. [Source: Billings, 1997]

  Discussion. User acceptance of automation centers on whether the user feels in ontrol of the system. [Source: Rouse, 1988]

Appendix B

<u>Design Guidelines for Decision Support Tool User Interface/Human-Computer Interaction</u>

- Clearly designate recommended option(s) from non-recommended options.
- Provide information about the reliability or certainty of recommendations to help build user trust.
- Provide reliability information to promote effective use of the tool.
- Provide a recommendation if it is at least 70% reliable.
- Provide information about the conditions under which recommendations will be more and less reliable.
- Provide information about factors contributing to DST recommendation(s).
- Provide a single recommendation when DST reliability is high, multiple recommendations (if available) when DST reliability is low.
    - Provide fewer, more reliable recommendations under higher workload conditions.
- Provide information about DST operational status to support situation awareness.
    - Provide notifications regarding system failures or errors.
    - Provide a clear indication of system "mode."
- Use terminology and symbology that is meaningful to the user and use the terms and symbols consistently.

<u>Training Guidelines for Decision Support Tools</u>

- Provide information about DST purpose.
- Train users on DST logic and functionality.
- Train users on DST reliability.
- Provide opportunities for hands-on learning.
    - Focus on helping users understand the reasons for human-system interactions, rather than simply presenting how to execute the actions.
    - Provide opportunities for users to test their knowledge.
- Train different user groups on DST capabilities that will be most useful in their specific work environment.
    - Provide a forum for users from different facilities to share how they use the tool.
- Train users to manage system failures.

<u>Consider Decision Support Tools to be "Team Players."</u>

- Design automation to be good team players.