EVALUATING CONTROLLER USE OF ADVANCED WEATHER PRODUCTS BY EVALUATING USER INTERACTION PATTERNS

Ferne Friedman-Berg, Ph.D.
The Titan Corporation, Mays Landing, NJ, USA

Ulf Ahlstrom, Ph.D.
Federal Aviation Administration, Atlantic City International Airport, NJ, USA

We evaluated how Terminal Radar Approach Control (TRACON) controllers interacted with advanced weather products when displayed on primary tactical radar displays or on auxiliary flat-panel displays. When we provide the controllers with access to appropriate weather tools, we potentially decrease workload and improve performance by increasing weather situational awareness (WSA). We assessed the benefit of providing weather information using Gantt charts to graphically depict user interaction patterns and by conducting an analysis of the frequency and duration of interactions with the weather tools. From these analyses, we identified two patterns of interacting with tools corresponding to heavy and light users. We correlated tool use with our measure of controller efficiency and found an increase in traffic throughput with increasing tool use. We also determined that the controllers were more likely to interact tactically with tools when placed on the Terminal Controller Workstation (TCW), turning them on when needed and turning them off when not needed. This method provides us with insight into how the controllers use the information on their displays to help provide them with tools to increase safety and efficiency.

INTRODUCTION

Unfavorable weather conditions can have a severe impact on both aircraft safety and system throughput. In the current National Airspace System (NAS), pilots have the primary responsibility for separating themselves from weather (Ahlstrom & Della Rocco, 2003). Most TRACON controllers receive weather information from supervisor briefings and by displaying up to six levels of precipitation on the Terminal Controller Workstation (TCW). However, this does not provide them with the detailed and timely information they need regarding weather conditions that have the greatest impact on operations, including thunderstorms, microbursts, gust fronts, and wind shifts (Ahlstrom, 2004). By providing controllers with direct access to the appropriate weather tools, we should be able to decrease controller workload and improve performance by increasing their weather situational awareness (WSA).

When we add weather tools to the controllers' displays, however, it is imperative that we also evaluate whether these tools take the controllers away from their primary duty of separating aircraft. If the controllers spend too much time interacting with a new tool, this could have a negative impact on performance. To truly evaluate these tools, it is not enough for us to simply evaluate the frequency and duration of user interactions. We must also look at the patterns of these interactions. In this study, our aim is to evaluate these patterns.

We evaluated how TRACON controllers interacted with advanced weather products when displayed either on their primary tactical radar display or on an auxiliary flat-panel display. We used interaction patterns to understand how controllers might incorporate weather information into a working mental model when they have direct control over when and how they access this information.

METHOD

Eleven non-supervisory, full performance level controllers participated in this study. Most of these controllers did not have access to advanced weather products at their home facilities.

Ahlstrom and Friedman-Berg (in press) conducted this simulation at the William J. Hughes Technical Center's Research Development and Human Factors Laboratory (RDHFL). The simulation configuration is designed to emulate Standard Terminal Automation Replacement System (STARS) functions. During each simulation run, two controllers managed traffic, with one responsible for East sector traffic and one responsible for West sector traffic. The controllers communicated with simulation pilots using appropriate air traffic control phraseology. The simulation pilots controlled aircraft from a remote room and maneuvered the aircraft targets displayed on the radar via a simulation configuration consisting of the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) and the Target Generator Facility (TGF). The TGF generates radar targets based on simulation pilot input which DESIREE then displays on the TCW.

There were three different simulation conditions. In the control condition, the controllers only had access to the basic precipitation levels normally found on the TCW. In the two experimental conditions, the controllers had access to advanced weather products. The primary difference between the experimental conditions was the placement of the advanced weather products. In the first experimental condition, we presented weather information on the TCW superimposed with traffic information. In the second experimental condition, we presented the weather information on an auxiliary Weather Information Display System (WIDS) above the TCW and presented traffic information on the TCW (see Ahlstrom, Keen, & Mieskolainen, 2004, for a detailed discussion of this display). To capture the complexity of real weather phenomena, we used two actual prerecorded Integrated Terminal Weather System (ITWS) samples to create two weather scenarios, Weather Scenario 1 (WS 1) and Weather Scenario 2 (WS 2). To ensure that the results were not due to the characteristics of a particular weather scenario, each controller ran traffic in all three tool conditions under both weather scenarios for a total of six simulation runs for each participant. We then analyzed results separately for WS 1 and WS 2. There were some differences between the results from WS 1 and WS 2 that are not relevant to the present paper, which are discussed elsewhere (Ahlstrom & Friedman-Berg, in press).

The advanced weather tools available to the controllers in the experimental conditions included an Echo Top prototype, a Gust Front prototype, a Precipitation Forecast prototype, a Storm Motion prototype, a Weather Loop prototype, and a Wind Shear prototype. The controllers could access all of these tools throughout the scenario. In addition, the controllers in all conditions still had access to the six precipitation levels normally found on the TCW, with Level 1 being the least severe and Level 6 being the most severe. We expected that the controllers would access the most useful tools more frequently or for longer durations than less useful tools. We were also interested in seeing if certain tools were used in combination with other tools, if tool usage varied based on tool location, and if tool usage varied based on the underlying weather conditions. For these evaluations, we recorded every controller interaction with the weather tools during the simulation. These interactions were all time stamped so they could be correlated with the weather scenario.

RESULTS

Gantt Charts

We created Gantt charts to show tool usage over time. Gantt charts were initially developed to schedule and manage projects (see http://www.ganttchart.com for more background on Gantt charts), but they are well suited to creating a useful graphical representation of how users interact with tools over time.

Our Gantt charts showed some systematic variations in how users interacted with both the advanced weather products and with the six precipitation levels. In our evaluation of precipitation level use, we found that the controllers in the experimental conditions typically turned on some precipitation levels (usually Levels 4, 5, and 6) and left them on, or turned some off and left them off (usually Levels 1, 2, and 3 - see Figure 1). In the control condition, users appeared to interact more with the six precipitation levels, turning them on and off more frequently (see Figure 2). From this, we conclude that when the controllers could access more advanced weather information to enhance their WSA, they had less need to access basic precipitation levels.

We also used the Gantt charts to evaluate how the controllers used the advanced weather products. We found that in the experimental conditions, the controllers primarily used the tools in one of two ways. Some controllers made very little use of the tools (see Figure 3) or turned on one of the advanced weather products and left it on (Storm Motion and Gust Front prototypes could be turned on and left on). Other controllers made extensive use of the weather tools, accessing the Gust Front, Weather Loop, Precipitation Forecast, and Storm Motion prototypes frequently throughout the 50-minute scenarios (see Figure 4). We hypothesize these patterns may correspond to two types of user populations, those who are early technology adopters and those who are slower to adopt new technology. It is important that we recognize if there are two distinct subpopulations of users because we can then tailor training to target both groups.

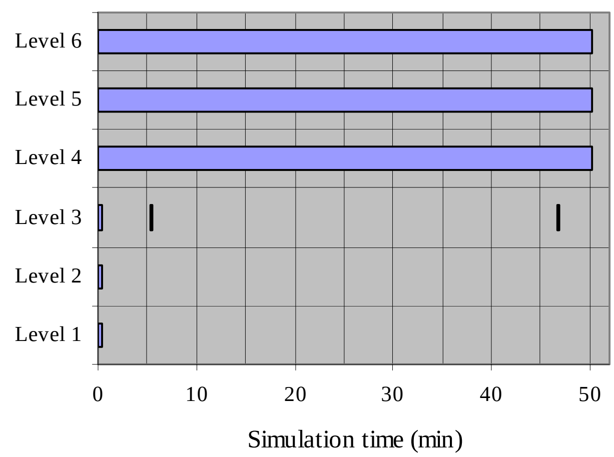

Figure 1. A typical example of one controller’s light use of precipitation levels in the experimental conditions.

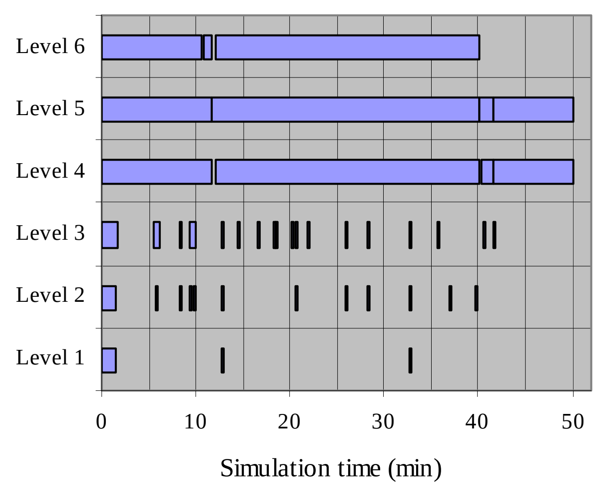

Figure 2. A typical example of one controller’s heavy and frequent use of precipitation levels in the control condition.

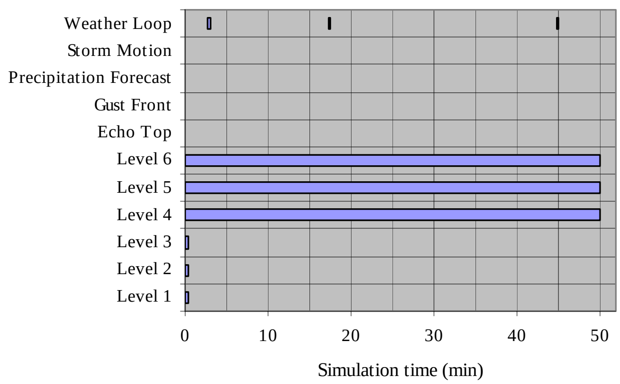

Figure 3. A typical example of one controller's light use of advanced weather tools in the experimental conditions.

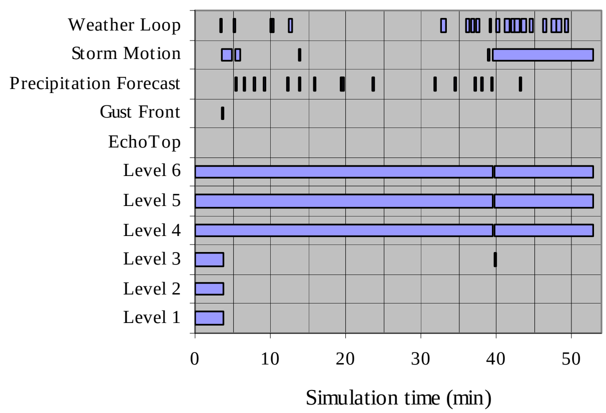

Figure 4. A typical example of one controller's heavy and frequent use of advanced weather tools in the experimental conditions.

We correlated the frequency of tool interactions in the TCW and WIDS conditions for both scenarios with our measure of controller efficiency, the number of completed flights (i.e., the number of instrument operations), to evaluate whether efficiency increased or decreased with increasing weather tool interactions. Although there might be some concern that frequent use might negatively impact performance, we did not find this to be the case.

As can be seen in Figure 5, most controllers interacted with weather tools between 19 and 83 times, but three controllers interacted with weather tools between 116 and 187 times. We assumed these outliers reflected tool experimentation by the controllers rather than operational use to facilitate control operations. Therefore, we performed the correlation twice, once including the outliers and once excluding them. In Figure 5, we plotted the best fitting regression line both with (solid) and without (dotted) the outliers. When including the outliers, we found there was a nonsignificant correlation between the number of weather tool interactions and the number of completed flights. However, when we excluded the outliers we found a significant correlation between the number of tool interactions and the number of completed flights, $F(1, 19) = 5.60$, $p = .03$, $r = .48$. As the number of weather tool interactions increased, there was a corresponding increase in the number of completed flights, indicating that controller efficiency increased with increasing tool use. Given that neither slope was negative, we found no evidence of a decrement in performance related to increased weather tool usage.

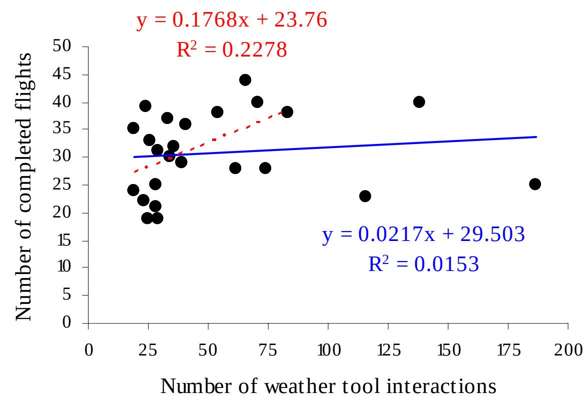

Figure 5. The correlation between the number of weather tool interactions and the number of completed flights.

Frequency and Duration Analyses

We conducted an analysis of frequency of use and duration of use for the weather tool interaction data. We performed planned comparisons of the frequency and duration of use for each advanced weather product, to see if usage differed when the products appeared on the WIDS compared to the TCW.

To avoid the clutter that could result from displaying weather information and traffic data on the TCW, the controllers could have adopted one (or more) of 3 strategies: 1) turning weather products on for shorter durations, 2) turning weather products on and off with greater frequency, or 3) turning weather products on infrequently. If the controllers in the TCW condition were able to learn how to use the tools in the most tactical way, we should find evidence of a decrease in the mean activation duration for each instance of a tool's use. Given that controllers might leave on weather information when it is spatially separate from traffic data, we might expect to find that the sum of activation durations is longer in the WIDS condition compared to the TCW condition. Because we expect these longer total activation durations in the WIDS condition, we would also expect to find a corresponding increase in the frequency of tool interactions in the TCW condition compared to the WIDS condition. However, if the controllers in the TCW condition were unable to learn how to use the tools tactically we might find less frequent interactions, resulting in the frequency in the TCW condition being lower than or the same as in the WIDS condition.

The results clearly demonstrated that in the WIDS condition, the controllers were more likely to turn on weather products and leave them on. In the TCW condition, they were more likely to use the tools tactically, turning them on for short durations only when needed. In the first analysis, we looked at whether the controllers were more likely to leave a weather product on for the duration of the scenario on the WIDS compared to the TCW. There were two weather products that could potentially be turned on and left on: Gust Front and Storm Motion. If a controller left either product on for more than half the scenario, they were coded as leaving a product on. As predicted by our model of tactical use, the controllers in both WS 1 and WS 2 were more likely to leave an advanced weather product on in the WIDS condition than in the TCW condition (WS 1: $t(10) = 2.39$, $p = .037$ and WS 2: $t(10) = 2.39$, $p = .037$). For WS 1, 54.5% of users left a tool on in the WIDS condition, compared to only 18.2% in the TCW condition. For WS 2, 63.6% of users in the WIDS condition left a tool on, compared to 27.3% in the TCW condition.

We also analyzed whether the mean number of interactions or the duration of the average interaction for each weather product differed for the WIDS and TCW conditions, using the planned comparison of (WIDS – TCW). For the Weather Loop, Precipitation Forecast, and Echo Top, there were no differences in the mean interaction frequencies or mean duration of interactions. For Gust Front, we did find a significant contrast for WS 2, $t(10) = -2.67$, $p = .024$, $\lambda = -1.18$, 95% *CI* [-2.17 to -.19], indicating there were more interactions with the Gust Front tool in the TCW condition, as predicted. Although there was some variability, we also found a significant difference in the total duration of interactions for WS 2 for Storm Motion, $t(10) = 2.73$, $p = .02$, $\lambda = 17.57$, 95% *CI* [3.21 to 31.92], indicating that when the controllers turned on Storm Motion, it was on for an average of 17 minutes longer in the WIDS condition compared to the TCW condition. This supports the previous finding that the controllers were more likely to leave weather products on for the duration of the scenario in the WIDS condition as compared to the TCW condition.

Behavioral Performance Data

An analysis of the behavioral performance data indicated that the controllers performed more efficiently with the weather tools than they did in the control condition (see Figure 6). They completed significantly more flights in WS 2, $t(5) = 2.99$, $p = .015$, $\lambda = 2.75$, 95% *CI* [.90 to 4.61], for the WIDS and TCW conditions compared to the control condition. This corresponded to an increase in throughput of 10% in the WIDS condition and 6% in the TCW condition.

Because we found an increase in efficiency, we recommend providing controllers with access to advanced weather tools. However, the efficiency data did not clearly favor either the WIDS display or the TCW display (see also Ahlstrom & Friedman-Berg, in press). We also found no clear user preference for either display location in the subjective data. The interaction data indicated that the controllers in the TCW condition did learn how to tactically use information overlaid on their radar display. Using point-of-gaze oculometer data, we found that the percentage of time controllers spent looking at the WIDS display during a 50 minute scenario averaged only 3.2% and 9.0% for WS 1 and WS 2 respectively with corresponding average fixation durations of 576 msec (*SE*=102) and 558 msec (*SE*=58). Based on these current findings, we do not have clear evidence favoring either display format.

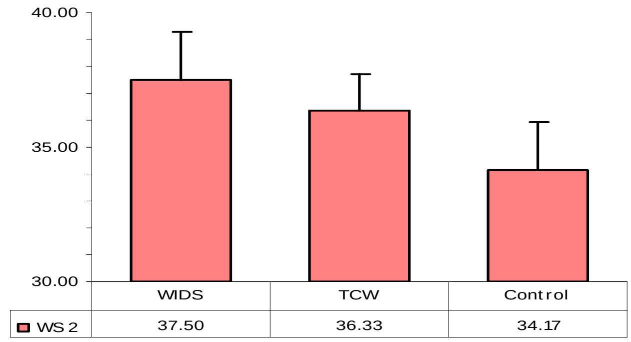

Figure 6. Mean number of completed flights by weather scenario and simulation condition. The error bars are *SE*.

Given that different TRACON facilities may have different needs due to different workstation configurations and space limitations on the controller workstation, we can recommend the placement of advanced weather products on either the primary radar display or on a WIDS display. However, due to issues related to mentally integrating radar targets on the TCW with weather data on the WIDS display, we also recommend adding radar targets to the WIDS display (see Sauer et al., 2002, for related findings on problems integrating spatially separate information).

Transitional Probability Diagrams

We also created transitional probability diagrams to examine the typical sequence of button activations for weather levels. In Figure 7, we have represented the transitional probabilities (for the WIDS condition) that a controller will activate a precipitation level (e.g., Level 2), given they have just activated another precipitation level (e.g., Level 1). This diagram looks very different from the transitional probability diagram for the control condition (see Figure 8). It looks like the controllers typically activated precipitation levels in sequence in the WIDS condition, but in the control condition had a greater number of nonsequential transitions. These diagrams show clearly different patterns for how controllers activate tools. It is important to understand typical sequences of tool use for the different conditions so that we can determine optimal button placement for tools given a specific operational setting.

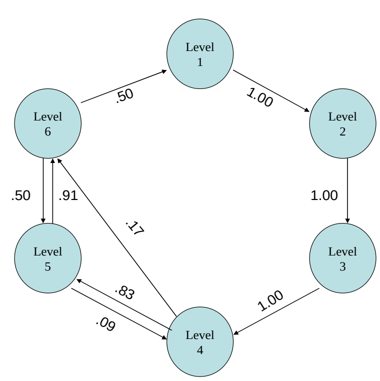

Figure 7. Transition probability diagram for precipitation levels on the WIDS display.

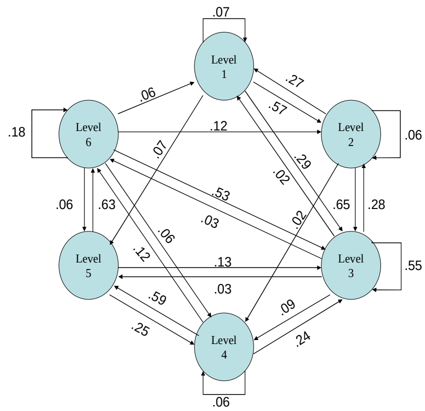

Figure 8. Transition probability diagram for precipitation levels in the control condition.

DISCUSSION

The typical metrics of human-in-the-loop simulations in air traffic control consist of dependent variables like distance flown, number and duration of holds, number of completed flights, and measures of workload. Although the value of these metrics cannot be understated, they do not provide researchers with a complete picture of how air traffic controllers use weather tools.

Researchers often examine the impact of different configurations of the same or multiple tools when added to controllers' displays (Sollenberger, Willems, Della Rocco, Koros, & Truitt, 2004), but they typically do not evaluate interaction patterns. We can only fully understand how tools impact performance when we systematically evaluate user interaction patterns, examine how patterns of use differ for different tool configurations, and investigate how these patterns of use relate to performance. Therefore, in this study we have chosen to look more closely at these interactions.

Using basic interaction data, we created a graphical representation of the data using Gantt charts and conducted an analysis of the frequency and duration of interactions. From these analyses, we determined that when using weather tools, the controllers employed two distinct patterns of interacting with the tools corresponding to heavy and light use. We looked at performance data and found an increase in traffic throughput with increasing tool use. We also found that the controllers were more likely to engage in tactical use of the tools when they appeared on the TCW compared to on the WIDS, turning them on when needed and turning them off when not needed. We hypothesize that because of potential problems with display clutter, the controllers in the TCW condition learned to tactically activate weather tools to support planning when merging aircraft into arrivals streams or selecting runways, which in turn resulted in more efficient use of the tools.

There are additional methods available to researchers to gain a better understanding of how controllers use new tools and how this use impacts performance. Some researchers use oculometers to study the kinds of information controllers attend to while controlling traffic (Willems, Allen, & Stein, 1999) or measure blink rate and pupil diameter to evaluate workload (Van Orden, Limbert, Makeig, & Jung, 2001). In a subsequent analysis, we will use oculometer data from this study to obtain a more objective measure of workload. We plan to examine whether workload varies with the frequency or duration of tool use. We also hope to use log linear modeling (Moertl et al., 2002) to statistically evaluate sequential dependencies in the tool interactions patters and to determine if there are reliably different patterns of use across conditions. These methods will aid research and help to provide controllers with new tools to increase safety and efficiency, while identifying tools that may have a negative impact on controller performance.

REFERENCES

Ahlstrom, U. (2004). *TRACON controller weather information needs: II. Cognitive work analysis* (DOT/FAA/CT-TN04/09). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Ahlstrom, U., & Della Rocco, P. (2003). *TRACON controller weather information needs: 1. Literature review* (DOT/FAA/CT-TN03/18). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Ahlstrom, U., & Friedman-Berg, F. (in press). *TRACON controller weather information needs: II1. Human-in-the-loop simulation* (DOT/FAA/CT-05/12). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Ahlstrom, U., Keen, J., & Mieskolainen, A. J. (2004). Weather information display system (WIDS). *Journal of Air Traffic Control, 46*(3), 7-14.

Moertl, P. M., Canning, J. M., Gronlund, S. D., Dougherty, M. R. P., Johansson, J., & Mills, S. H. (2002). *Aiding planning in air traffic control: An experimental investigation of the effects of perceptual information integration.* (DOT/FAA/AM-02/22). Washington, DC: Office of Aerospace Medicine.

Sauer, J., Wastell, D. G., Hockey, G. R. J., Crawshaw, C. M., Ishak, M., & Downing, J. C. (2002). Effects of display design on performance in a simulated ship navigation environment. *Ergonomics, 45*(5), 329-347.

Sollenberger, R. L., Willems, B., Della Rocco, P. S., Koros, A., & Truitt, T. (2004). *Human-in-the-loop simulation evaluating the collocation of the user request evaluation tool, traffic management advisor, and controller-pilot data link communications: Experiment I - Tool combinations* (DOT/FAA/CT-TN04/28). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors, 43*(1), 111-121.

Willems, B., Allen, R. C., & Stein, E. S. (1999). *Air traffic control specialist visual scanning II: Task load, visual noise, and intrusions into controlled airspace* (DOT/FAA/CT-TN99/23). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.