2007

Effectiveness of Weather Coding Schemes on Air Traffic Displays

Chen Ling

Yifei Dong

Desmond Harvey

Paul Brantmeier

Fabio Rojas

EFFECTIVENESS OF WEATHER CODING SCHEMES ON AIR TRAFFIC DISPLAYS

Chen Ling, Yifei Dong, Desmond Harvey, Paul Brantmeier, Fabio Rojas
University of Oklahoma
Norman, OK

Air traffic controllers need to reliably and quickly get weather information for aircraft safety. They typically use visual displays to acquire weather information. Current air traffic control displays use different schemes to encode weather information. The objective of this study was to explore the effectiveness of several coding schemes for precipitation levels. We studied three coding schemes that were adapted from the weather displays employed in the Automated Radar Terminal System (ARTS) Color Display (ACD), the Standard Terminal Automation Replacement System (STARS), and the Integrated Terminal Weather System (ITWS). We developed a computer simulation interface that displayed weather maps overlapping with aircraft datablocks. The subjects need to report the highest weather level presented on an 8x6 weather map. They reported the level for all weather blocks in the first type of tasks, and for the owned and point-out aircrafts for the second and third types of tasks. The dependent variables were the reaction time and correctness in identifying weather levels displayed. The results indicate that coding schemes, task type and their interactions all significantly contribute to the error rate of task performance. The coding scheme adapted from the STARS system resulted in lowest error rate and shortest response time.

Introduction

Adverse weather conditions create safety hazards for pilots and constrain the usable airspace for air traffic controllers (Ahlstrom, 2005). Air traffic controllers' tasks include maintaining separation among aircrafts to prevent accidents and ensuring an organized flow of traffic. Air traffic controllers also provide navigation and weather information to pilots in case of weather adversities (Xing, 2004). In case of severe weather, the controllers need to alert the pilots and answer questions regarding to weather conditions, while attending to other dynamic information from multiple sources and making important decisions. Effective way to present weather information on ATC displays is critical for the air traffic controllers to grasp the weather condition fast and accurately, and convey the information correctly. Poorly designed weather displays can be detrimental to the operation by exhibiting redundant visual displays and increasing the display clutter in the controller room (Ahlstrom, 2005). But proper presentation of weather information has proved to improve controller's efficiency with 6% to 10% increase in sector throughput (Ahlstrom & Friedman-Berg, 2006). Currently, several different weather-coding schemes are used in various ATC automation systems to represent weather precipitation levels, employing different visual stimuli dimensions: chromatic, brightness, color, and texture. Because these dimensions are also used to represent other ATC data, such as aircraft data block, flight path, regions, etc., there is a potential competition for the controller's cognitive resources among these data sources. It is interesting to examine how operator performance is affected by different weather coding schemes.

Weather information is represented by coding schemes in ATC systems. This information is crucial because it suggests whether the airplane is safe to land into the airport. Currently three different coding schemes are employed in different systems- Automated Color Display (ACD), the Standard Terminal Automation Replacement System (STARS), and the Integrated Terminal Weather System (ITWS). According to Xing (2006), the ACD system is one of the primary displays in many terminal facilities; the STARS system is the newly developed primary display for terminal facilities, and is deployed in a number of facilities. The ITWS system is the major weather display in terminal facilities. In order to evaluate the comparative effectiveness of these coding schemes, empirical experiment with task performance of response time and error rate can provide the most valuable information.

The objective of this study is to empirically determine the best coding scheme to represent different weather conditions. The scheme that is sought after should allow for human operators to quickly determine weather conditions fast and accurately. Such a scheme could make it easier for an air traffic controller to monitor the weather while focusing on all other required tasks.

Literature Review

Colors are being widely used in air traffic control displays (Xing, 2006). There are advantages in using color in displays. But basic human factors and color principles have not been employed sufficiently thus far. There are no standard coding schemes that developers go by. Colors are used for three main purposes in ATC displays: attention, identification, and segmentation. The disadvantage of the color also might be due to the following reasons: distraction, coding uncertainty, loss of integration, multiple coding schemes, experience interference, text readability, color naming consistence, and view-angle intolerance (Xing, 2006). Ahlstrom and Arend (2005) also noted the usability issues that might be caused by color coding in ATC systems: legibility, salience manipulation (clutter avoidance), and color recognition. The first two issues are directly related to the luminance contrast between the foreground and background of the information. Krebs and Ahumada (2001) studied the masking effect on text block for ATC weather displays. A masking metric can be used to calculate the equivalent contrast between the background and foreground, and is found to be a good indicator of percent correctness and response latency for tasks. One important factor to consider when choosing weather coding schemes is the contrast between the weather background and the information shown on the foreground.

The characteristics of good coding scheme include: detectability, discriminability, meaningfulness, standardization, and use of multidimensional codes (Sanders & McCormick, 1993). According to Sanders and McCormick (1993), the stimuli used in the coding scheme should be detectable by the human sensory scheme (detectability). Each code symbol need to be discriminable from other code symbols (discriminability). If possible, codes should be meaningful to the user (meaningfulness). The same coding scheme should be kept the same across different usage situation (standardization). Also, the use of the multidimensional codes can increase the number and discriminability of coding stimuli (use of multidimensional code). The differences in the three coding systems indicate that the coding standardization is not well observed in ATC system. The three schemes: STARS, ACD, and ITWS systems, vary in many of the coding characteristics in their designs, and may result in different effectiveness.

Method

Participants

Thirty (30) participants were fortuitously sampled for the experiment. The participants were primarily students at the University of Oklahoma. The ages of participants are between 18-24 years. Because of the nature of the study, none of the participants have any kind of vision impairment. Some of the participants were eligible for course credit. To avoid learning effect across different schemes, each participant was randomly assigned to use only one of the three coding schemes. Therefore, ten participants were used for each coding schemes.

Experimental Design

Independent variables. The independent variables for this study are the three coding schemes adapted from ATC systems and the task types. The three coding schemes being tested include those adapted from the Automated Color Display (ACD), the Standard Terminal Automation Replacement System (STARS), and the Integrated Terminal Weather System (ITWS). This is a between-subject variable because each person only use one coding scheme. ACD, STARS, and ITWS systems are major displays currently used in ATC terminal facilities. The coding schemes in ACD, STARS, and ITWS, are all used to represent six weather levels, ranging from moderate to severe weather conditions.

Different ways of coding are used in these three systems to represent weather severity. In the ACD system (see Figure 1), gray color is used to represent moderate weather (levels 1 and 2), orange is used for heavy weather (levels 3 and 4), and orange red color is used for severe heavy weather (levels 5 and 6). To differentiate the two levels within the same type of weather, gray stipples are used for levels 1, 3, and 5, whereas uniformly filled colors are used to represent levels 2, 4, and 6. Orange and orange-red are used to draw controller's attention to severe weather condition. This is a multidimensional coding scheme with three colors and two textures to represent six weather levels.

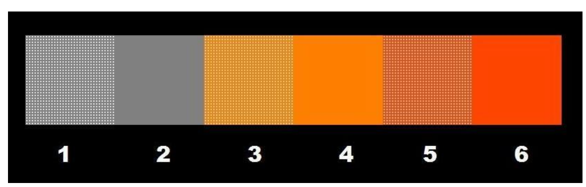


Figure 1. Coding scheme for Weather adapted from ACD System

STARS system is a primary display used in terminal facilities. In STARS system (see Figure 2), dark grey blue is used for levels 1, 2, and 3, and dark mustard is used for weather levels 4, 5, and 6. Low density bright gray stipples are overlaid on the color to represent levels 2 and 5, whereas high density bright gray stipples are overlaid on the color to represent levels 3 and 6. This is a multidimensional coding scheme with two colors and three textures to represent six weather levels.

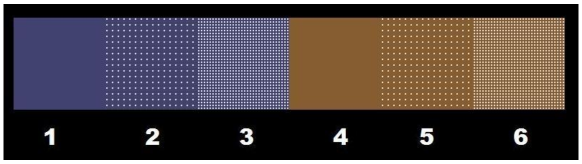


Figure 2. Coding scheme for Weather adapted from STARS System

In ITWS system (see Figure 3), precipitation is displayed using a six-level scale developed by the National Weather Service (NWS). The six weather precipitation levels are indicated with light green, green, yellow, light orange, orange, and red. This is a single dimension coding scheme with six colors representing six weather levels. The purpose of colors is mainly identification, but the red color used for the sixth level is also for drawing attention.

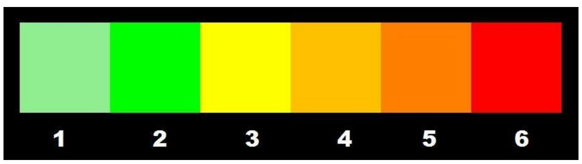


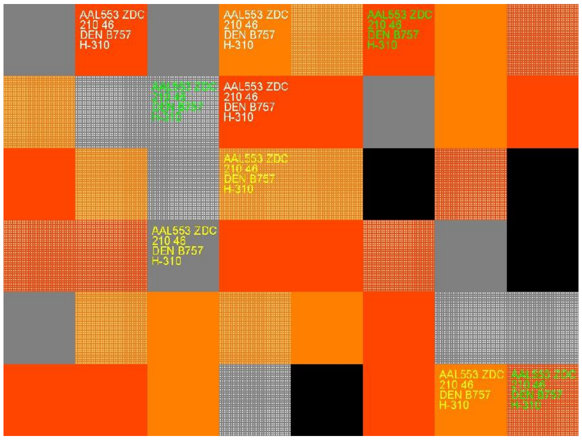
Figure 3. Coding scheme for Weather adapted from ITWS System

The second independent variable in this study is the task type. The task types are manipulated by the nature of questions posed to the participants about the weather levels. In the first type of task (task level 1), the focus is just to study the differences between weather coding schemes. Participants are asked to identify the highest weather level presented on the weather map. In order to answer the question, the participants need to first look at the display with 8x6 blocks and then determine out of all weather levels presented, what the highest level is. This is similar to what the controllers face in the real-world scenario, where they need to identify the highest weather level existing with the system, and know which level it is. For the second type of tasks, on the other hand, the purpose is to test not only the effectiveness of parsig weather information but also the interaction between the weather coding and the information overlaid on the weather background. This is even more representative for tasks with the ATC displays where information is presented on the weather background. The illumination contrast between the background and information might affect how fast and accurately the information is viewed. In this type of task, participants are asked to identify the highest weather level of the owned (task level 2) or point-out aircrafts (task level 3). The owned category refers to the aircrafts that the controller is responsible for on the screen, whereas the point-out category refers to the aircraft that will soon be the responsibility of the controller. Another color coding system is used to represent the ownership of the aircrafts. The data block color for owned aircrafts is yellow, and for point-out aircrafts is white. Another possible color of data blocks is green, which represents the aircrafts owned by another controller. When participants need to pay attention to both the aircraft data block color and the background color denoting the weather level, the task becomes more challenging because two color coding systems are used simultaneously. If the weather coding interfere with the data block color, it would be even harder for the participants to carry out the tasks. Because each participant carry out both types of tasks, this variable is a within-subject variable.

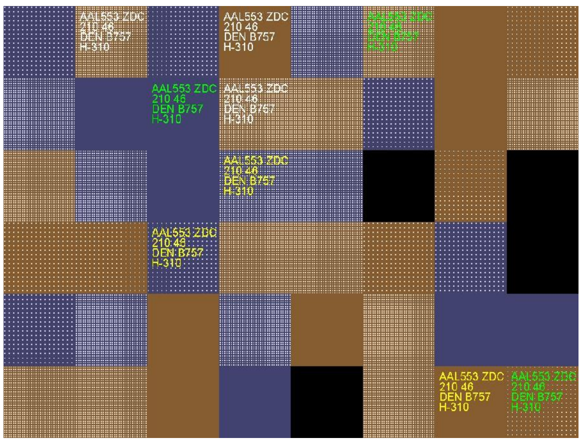
Dependent variables. The dependent variables measured are the response time of the participants to answer each question and the error rate of the answers by the participants.

Equipment

The real experimental sessions follow after the participants have completed the training. Pre-generated weather maps rendered in the three coding schemes adapted from three ATC systems are displayed by DMDX, an accurate timing display program developed at the University of Arizona by K.I. Forster and J.C. Forster. Each display consists of a grid of 8x6 cells with randomly assigned colors in the scheme. The cell can also be black, which denotes no weather is present in that cell (level 0). 9 aircraft data blocks are overlaid on the weather map background, and they belong to one of the three categories: owned, un-owned and point-out, with 3 cells in each category. The participants go through three trials with different task types (task levels 1-3). Within each trial, 40 weather maps are shown to the participant. The highest weather level presented on these maps is randomly distributed among the six possible levels. Data from 120 runs are collected with each participant. The participants are asked to answer the highest weather level that is presented in the map by pressing the corresponding number as soon as possible. The correctness of the answer and the response time are derived from the raw data recorded by the DMDX program. Figure 4 shows the example screens coded with adapted ACD, STARS, and ITWS coding schemes for the experiment trial with aircraft datablocks. All the schemes in the figure use the same weather data.



(a) with coding scheme adapted from ACD System



(b) with coding scheme adapted from STARS System

(c) with coding scheme adapted from ITWS System

Figure 4. Display of the same weather and ATC scenario using three coding schemes

Procedure

Before the experiment, the experimenter makes sure that the participant has no vision impairment. The experimenter then explained the purpose of the experiment and how the coding schemes are used in the experiment. The participants are randomly assigned to learn one of the three coding schemes. After understanding the scheme, participants go through a computerized training session. In the training, the participants are first presented with one block with weather one weather level represented and asked which level is represented. Feedback of the correct answer is given if any mistakes are made. Then the participants are presented with three blocks with aircraft data block, and are asked to report the weather level of a particular category of aircrafts (owned, point-out). After the training session, the participants proceed to complete tasks with the computer program. Answers to the questions and the response time are recorded by the computer program. After the experiment sessions, the participants are thanked for their time.

Results and Analysis

Based on the data collected from the 30 participants, separate analyses were completed with time and amount of errors as dependent variables respectively.

Error rate as dependent variable. In terms of error rate of the tasks, the coding schemes were found to be significantly different from each other (p=0.043). Post-Hoc Analysis shows that STARS system (M=5.17%, SD =0.221) has significantly lower error rate than the ITWS coding schemes (M=22.33%, SD=0.417). There was no significant difference between ACD (M=18.50%, SD=0.389) and ITWS coding schemes. And there were also no significant differences between the ACD and STARS coding schemes.

The task type factor was also significant (p=0.0003). The task level 1 (M=12.67%, SD =0.333) resulted in significantly less error rate than the other two levels of tasks (M=15.17%, SD=0.359 for level 2 and 18.17%, SD=0.3857 for level 3).

The scheme*complexity interaction was also significant (p=0.001). From Figure 5, it can be seen that the effects of task types on error rate are different for the different coding schemes. For the STARS system, there is not much effect, but for the ITWS and ACD system, the effect is more obvious. The reason might that the contrast between the data block colors (yellow and white) and the weather levels color is not as much in the ITWS and ACD systems as in STARS system, making them more error prone. The level 2 of ACD has the highest error rate among three task levels for ACD. Because the level 2 task require controllers to identify highest weather level for owned (yellow) data block. The higher levels of

ACD coding scheme use orange and orange red, which are similar in chromaticity with yellow. This might be the reason for the error rate. Similarly, the level 3 task has a very high error rate in the ITWS system. This task requires participant to identify the highest weather level with the point-out (white) aircrafts data blocks. The low contrast between the white text and background and the large number of colors (6) used in the system might both contributed to the high error rate.

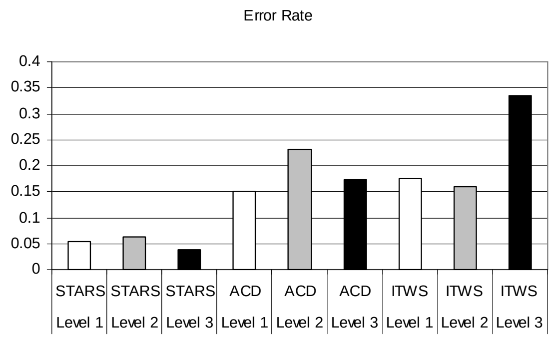


Figure 5. The error rate of different adapted coding schemes and task levels

Response time as dependent variable. In terms of response time it took to complete the tasks, the coding schemes were found to be significantly different from each other (p=0.012). Again, the color coding used in STARS system (M=3.098s, SD =1.898s) resulted in significantly shorter response time than the ITWS (M=4.800s, SD=3.158s) and ACD (M=3.558s, SD =2.988s) systems.

The task level (p=0.0445) were also significant. The task level 1 (M=3.548s, SD=2.912s) takes shorter time to respond than the 2 other levels of tasks, including answering the highest weather level for point-out aircrafts (M=3.848s, SD=2.495s) and owned aircrafts (M=4.060s, SD =3.037s). There is no significant difference between the level 2 and level 3 tasks in terms of response time.

The scheme*complexity interaction was also significant (p=0.004). From Figure 6, again, the effects of task type on response time are different for the different coding schemes. The task type did not make much difference for the STARS system, but the effect is more obvious for the ITWS and ACD system. The reason can be attributed to the difference in contrast between the data block color and the weather levels color as well. The ITWS system takes longer response time for level 2 and level 3. It implies bigger interaction between the data block text color and weather level background. Because of the differences in luminance contrast, the same data block color (e.g. yellow or white) may look different on different background. Therefore, it takes longer to identify the 3 data blocks of particular category before determining the highest weather level among the three.

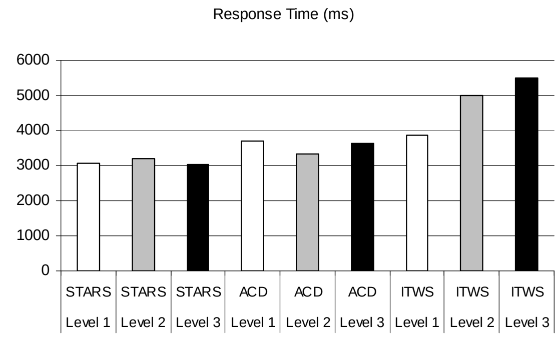


Figure 6. The response time (millisecond) of different adapted coding schemes and task levels

Conclusion and Recommendations

The adapted STARS coding scheme performed better than the other two coding schemes adapted from ITWS and ACD systems, with the lowest error rate in the shortest response time. The coding scheme of 2 color x 3 patterns proves to be easier to identify for the human operators.

Revisiting the characteristics of good coding scheme include: detectability of codes, discriminability of codes, meaningfulness of codes, standardization of the codes, and use of multidimensional codes (Sanders & McCormick, 1993), the coding scheme used in STARS is superior in many aspects compared to the other two schemes. The coding discriminability is better for STARS because the only two colors used, brown and blue, are different enough in chromaticity, and is easily discriminable. But in the adapted ACD system, the colors of orange and orange red are not sufficiently different, making it harder for users to differentiate between them. And the ITWS system used so many colors and the colors are not sufficiently different between light green and green, or light orange and orange for users to discriminate.

In regard to the meaningfulness of the code, the use of different densities of stipples in STARS to represent the higher levels of weather is intuitive, because user can imaging the stipples as precipitations, with higher density associated with more precipitation. The conceptual compatibility between the code and the referent is designed well, and might contribute to the effectiveness of the code. Both ACD and STARS systems used multiple dimensions in the scheme which derived better results than ITWS which used coding along single coding scheme. The use of multiple dimensions should be applied in the future system design.

This study not only helps to identify the “winning” coding scheme to represent weather precipitation conditions, but also studied the relationship between the data block color in the foreground and textured weather information in the background. When the luminance contrast is not sufficient, or when data block of the same color appear on different background, it is harder for human operators to identify the targets.

The limitation of the current study lies in that the interface used to test the coding schemes are not real ATC displays, and there is no continuation of the weather pattern across the adjacent blocks, which might be the case for the real weather conditions. Therefore, it is harder for the subjects to tell the weather condition with the current display than with the real situation.

Another limitation is that we only have access to informal, natural language description of the coding schemes. While we are confident with the experiment methodology, we need the precise definition of coding schemes, such as the color value (what exactly is “orange red”?) and stipples (patterns and density) to make formal recommendations. We also acknowledge that many other factors, such as controller’s familiarity with existing systems, should be considered in the selection of the best coding scheme.

Acknowledgements

This study was research collaboration with Federal Aviation Administration Civil Aerospace Medical Institute (CAMI), Oklahoma City. We express our sincere appreciation to Dr. Jing Xing for her invaluable insights and supports during the study.

References

Ahlstrom, U. (2005). Work domain analysis for air traffic controller weather displays. *Journal of Safety Research*, 36(2), 159-169.

Ahlstrom, U., & Arend, L. (2005). Color usability on air traffic control displays. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, 93-97

Ahlstrom, U., & Friedman-Berg, F. (2006). TRACON controller weather information needs. Washington, DC: Federal Aviation Administration, No: DOT/FAA/TC-06/10

Krebs, W. L., Ahumada, A. J. (2001) Air traffic control weather radar displays: Validation of a masking metric for prediction of text block identification. *Proceedings of Annual Meeting of the Human Factors and Ergonomics Society*, October 2001, Minneapolis MN.

Sanders, M. S., & McCormick, E. J. (1993). *Human Factors in Engineering and Design*, 7th Ed., McGraw-Hill Book Company, New York, N.Y.

Xing, J. (2006). Color and visual factors in ATC displays. Washington, DC: Federal Aviation Administration; No: DOT/FAA/AAM-06/15

Xing, J. (2006). Color analysis in air traffic control displays, Part I. Radar displays. Washington, DC: Federal Aviation Administration; No: DOT/FAA/AM-06/22