2009

Cue-Based Training Effects on Visual Scanpaths During Weather-Related Decision Making

Michael W. Sawyer

Scott A. Shappell

CUE-BASED TRAINING EFFECTS ON VISUAL SCANPATHS DURING WEATHER-RELATED DECISION MAKING

Michael W. Sawyer
Clemson University
Clemson, South Carolina
Scott A. Shappell
Clemson University
Clemson, South Carolina

Flight into adverse weather remains a leading cause of fatal accidents in general aviation. The situation assessment hypothesis suggests that pilots continue into adverse weather because they fail to accurately recognize the weather conditions present. In this study 20 participants' eye movements were tracked as they viewed various weather scenes before and after training. The results showed that after training participants made decisions using fewer visual fixations and less total gaze time. Further, the average time until first fixation on critical weather features was decreased after training. Participants were effectively taught what weather features are important, thus allowing participants to make quicker, more efficient decisions. Eye tracking was found to offer objective evidence of cue-based training's ability to affect and improve the decision making process.

Adverse weather is consistently cited as a cause in general aviation (GA) accidents; especially in fatal accidents. The majority of these fatal weather accidents, up to 90%, occur when pilots who are certified to fly only according to visual flight rules (VFR) continue flight into instrument meteorological conditions (IMC) (Coyne, Baldwin, & Latorella, 2008). VFR allows pilots to fly only in conditions which allow the pilot to control the aircraft by visual reference to the environment outside the cockpit. Specific weather minimums have been established in terms of visibility and cloud ceiling to ensure that pilots can control their aircraft. Regardless of these regulations, VFR pilots continue to fly into IMC, often with tragic results. VFR into IMC accidents have been shown to be result in at least one fatality 75% of the time, compared to 18% for other types of GA accidents (Wiegmann & Goh, 2000). What's more, these accidents should technically never happen as they involve pilots flying in conditions they are not certified for.

The majority of GA accidents are a result of pilot performance, as opposed to mechanical or structural failures (O'Hare, Wiggins, Batt, & Morrison 1994). There is strong evidence suggesting that pilots continue into IMC because they fail to accurately recognize the severity of the conditions ahead of them (Goh & Wiegmann, 2001; Wiegmann, Goh, & O'Hare, 2002). Further, several studies have shown that given the same written information about a flight, pilots tend to make consistent, predictable decisions (Driskoll, 1998; Hunter, 2003). This suggests that pilots must be having trouble gathering and identifying weather conditions accurately.

Training

Research suggests that training is the best method for improving situational assessment (Gaba, 1995). Cue-based training in particular has established itself as an effective method for training decision-making in many industries, including: emergency response (Ash & Smallman, 2008, aviation (Wiggins & O'Hare, 2003), cognitive rehabilitation (Hampstead, Sathian, Moore, Nalisnick, & Stringer, 2008, medicine (Jenkins, Shields, Patterson, & Kee, 2007), and law enforcement (Santarcangelo, Cribbie, & Hubbard, 2004). Cue-based training identifies and teaches specific cues that signify a change in system state which require a specific response (Smith, Giffen, Rockwell, & Thomas, 1986). One such program, WeatherWise, has been developed specifically to teach weather-related decision making (Wiggins & O'Hare, 2003).

WeatherWise was developed through a series of research studies aimed at understanding how pilots make weather decisions. Interviews with expert pilots were used to identify the key features of weather available to pilots as they made weather related decision (Wiggins & O'Hare, 2003). An online survey was then created to allow pilots to rate the importance of each cue. From the results of the study it was identified that the presence of three or more cues indicated that pilots should divert from their flight path. These studies thus served as the basis for the creation of a training program. An initial test of the program was completed by a group of pilots in 2003. The study found the program increased the pilots' subjective importance ratings of all nine weather cues. Further, pilots who received the training condition outperformed the control group on a decision making task involving a simulated flight. While the program increased the pilots' subjective importance ratings for weather cues, a more thorough analysis of its affect on pilot behavior is needed.

This study, therefore aims to analyze the effects of cue-based training on weather-related decision making. Eye movements will be utilized as the primary measure to study the effects of training. Eye movements are strongly correlated to interest (Starker, 1990), and have been shown to provide insights into a person's decision making process (Land, 2007). Eye movements, therefore should provide a novel method of assessing the training program's effectiveness.

Methodology

Participants

This study involved 20 participants recruited from Clemson University and surrounding areas. There were nine males and eleven females participants. The participants were an average age of 25.5 years old with a standard deviation of 4.8 years (min = 21, max = 42). None of the participants had accumulated any flight hours or certifications. All subjects in the study reported having normal or corrected to normal vision.

Apparatus

A Tobii ET-1750 eye tracking monitor was used to collect all eye movement data. The ET-1750 has a 17" monitor and samples at a rate of 50 hz with 0.5º accuracy. The resolution of the monitor was set at 1280 x 1024 pixels. Eye movement data was collected using the software program ClearView 2.7.0 developed by Tobii Technology.

Training Program

The training program WeatherWise was used in this experiment to teach weather-related decision making. WeatherWise is a cue-based training program developed by Dr. David Hunter, Dr. Mark Wiggins, and Dr. David O'Hare. The program was produced by the Federal Aviation Administration, Office of Aerospace Medicine for the Aviation Safety Program of the Flight Standards Service, with the assistance of The Ohio State University, The University of Western Sydney, The University of Otago, and King Schools. The program is available free in the public domain. WeatherWise and its development are well documented (Wiggins & O'Hare, 2003).

Weather Images

In preparation for this study a group of 120 royalty-free or creative commons weather images were gathered from various photo websites. The pictures collected showed a variety of weather scenes a pilot might encounter, ranging from relatively clear skies to severe thunderstorms. From the original 120 images, the 36 best images were chosen. These images were then sent out for external validation by a panel of five certified flight instructors with over 5,000 flight hours. The panel rated the conditions of each image as either VFR or IFR conditions. From their ratings, the 10 images which best represented each set of weather conditions were chosen. These 20 images were then randomly divided into two even groups.

Procedures

Participants were first given an initial briefing about the nature and goals of this study. They then read and signed an informed consent form. Each participant then completed a brief demographic questionnaire.

Participants were then given an introduction to GA and weather decision making. Participants were given a description of VFR including the specific requirements for daytime flight in class G airspace. Participants were told to assume they were on a cross-country VFR daytime flight in class G airspace. Participants were then asked if they had any questions or needed any clarification about VFR. The introduction was read from a script to ensure consistency.

The first task consisted of showing participants a series of 10 randomly ordered weather images on the eye tracker. For each image, participants verbally responded to the question "could you continue along your current flight path while staying above VFR minimums?" The image was displayed until a response was stated. This process was continued for each of the 10 images.

Participants then completed the training program WeatherWise. Participants were allowed as much time as needed to fully complete the training program.

Participants then returned to the eye tracker to view another series of weather images. Participants first viewed a randomly ordered set of new weather images, again responding whether or not they could continue on their flight path while staying above VFR minimums. Participants then viewed and responded to the original 10 images from the first part of the study, again in a random order.

Subjects were then thanked for their time, compensated and dismissed.

Data

The experiment was setup to capture differences resulting from training. The two types of data collected in this experiment consisted of the verbal responses to each weather scene and the eye movement data collected during each trial. The accuracy of each decision, both overall and in terms of signal detection, was calculated by comparing participants' responses to the responses identified by the panel of experts of each image.

Eye tracking data was continuously collected during each of the 1,800 weather decisions made in the study. Areas of interest (AOIs) were defined in each image prior to testing. In each image the following AOI features were identified: terrain, clear sky, clouds, horizon, cloud base, and cloud darkness. These features were chosen to be consistent with previous research findings (Wiggins, 1999). The eye tracking data was analyzed in terms of fixations, dwell time, and time until first fixation.

Results

Response Accuracy

Correct trials occurred when participants' responses to a weather scene matched the responses of the expert pilot panel. When the scene conditions were below VFR minimums (i.e. adverse weather was present) correct responses were coded as hits, while incorrect responses were misses. When conditions were above VFR minimums (i.e. adverse weather not present), correct responses were correct rejections (CRs), while incorrect responses were false alarms (FAs). An analysis of the responses to each weather scene is shown below in Table 1.

Table 1. *Response accuracy for all trials*

	n	Correct	Incorrect	Hit	Miss	CR	FA	d'	c
Pre-Training	200	143 (71.5%)	57 (28.5%)	66	34	77	23	1.151	0.163
Post-Training New	200	149 (74.5%)	51 (25.5%)	82	18	67	33	1.355	-0.238
Post Training Repeat	200	147 (73.5%)	53 (26.5%)	81	19	66	34	1.290	-0.233
All Trials	600	439 (73.2%)	161 (26.8%)	229	71	210	90	1.241	-0.096

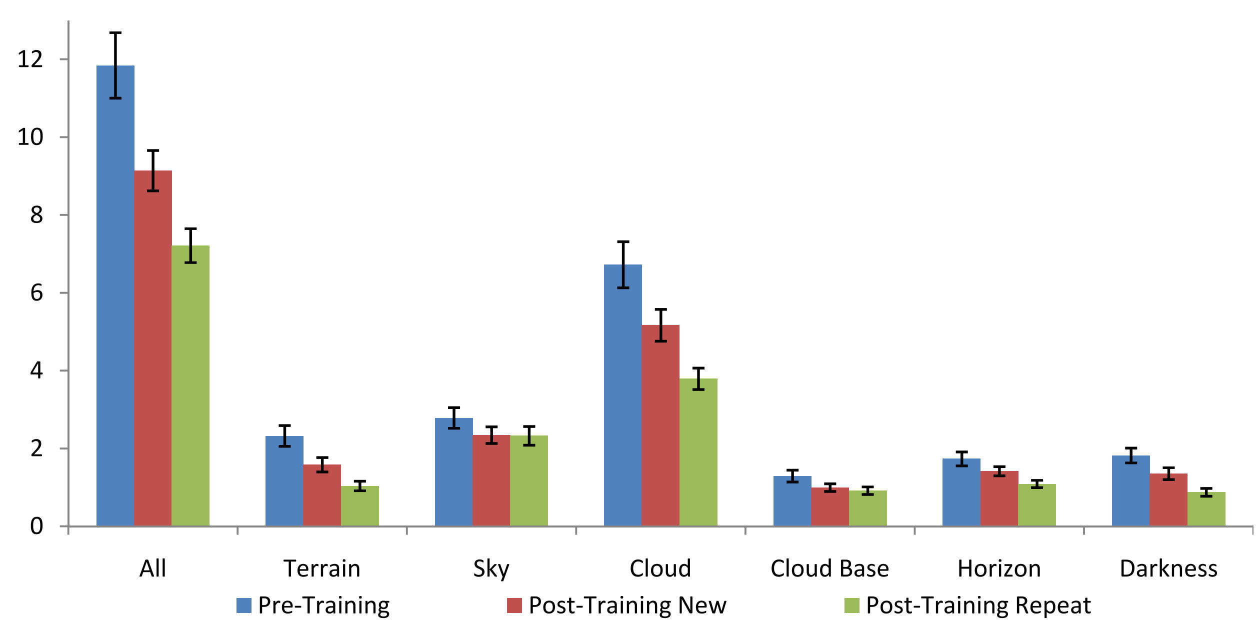
While an increase in response accuracy was observed after training, a 2-proportion z-test found the change was not significant between the pre-training condition and either the new image condition ($z = 0.68$, ns) or the repeated image conditions ($z = 0.45$, ns). Therefore, based on these data, it cannot be said that training improved decision accuracy.

It was found, however, that training did significantly reduce the number of flights into adverse weather as seen in the improved hit rate in both post-training conditions (new: $z = 2.62$, s, repeated: $z = 2.44$, s). No significant change was seen in the false alarm rate in either conditions (new: $z = 1.58$, ns, repeated: $z = 1.74$, ns).

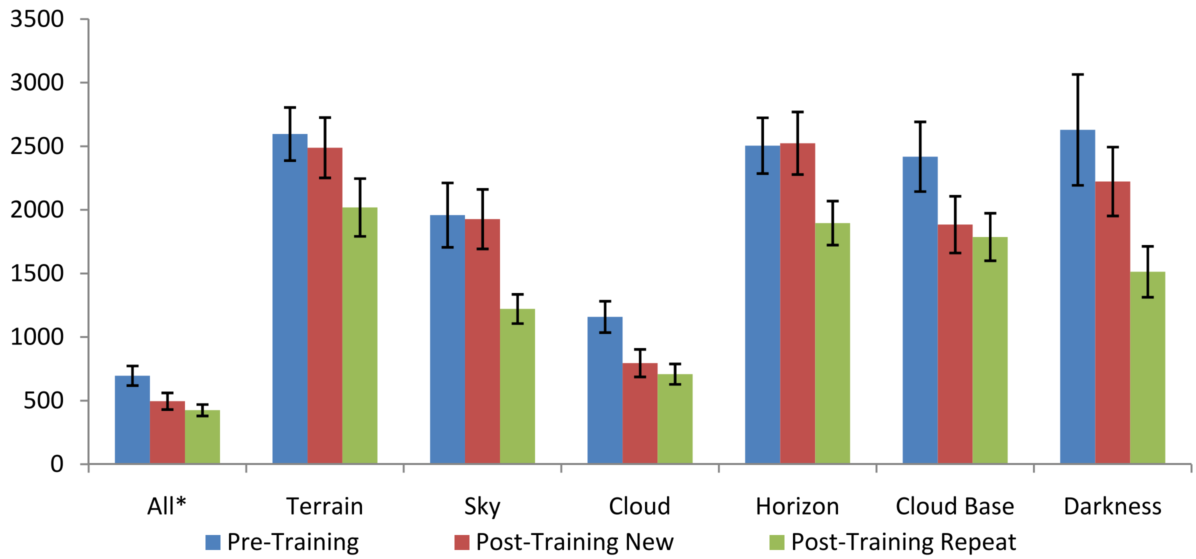
An overall shift in bias was observed after training, as seen by the change in c, from a positive value (a liberal bias) to a negative value (a conservative bias). Further, an increase in discriminability was also seen in the increased d' value in both post-training conditions.

Eye Tracking Data

Eye tracking data was continuously collected over all trials in this study. Of the 600 trials, there were 22 (3.6%) in which no fixations were recorded. Those trials were excluded from the analysis. The average numbers of fixations per trial are shown below in Figure 1. The *All* category represents the average number of total fixations per trial and is not simply a summation of other categories due to overlapping AOIs.



As can be seen, there was a large decrease in the number of fixations per trial in both post-training conditions. Differences between the pre-training and post-training new images best show the effects of training as the participants were not familiar the weather scenes. As participants had seen the images in the repeated condition before, a large decrease in the number of fixations would be expected regardless of training condition. There was a significant reduction in the number of visual fixations used to make a decision after training. The largest decrease in fixations was seen in the cloud group, followed by the terrain group.



*All represents the first fixation anywhere in an image

Figure 1 Average Time Until First Fixation

A decrease in the time until the participants' first fixation was seen after training. This signifies that after training participants scanned the image less before fixating on key features. This suggests that after training participants better knew what they were looking for in each image. Comparisons of the time until the first fixation within each category showed differences in the cloud and cloud base category.

Discussion

The training program WeatherWise was able to reduce the number of simulated flights into adverse weather, while also changing participants' visual scan behavior. While overall decision accuracy was not improved by the training program, the increased hit rate represents a successful decrease in inadvertent flights into adverse weather. In high-risk environments, this shift towards a conservative decision making bias should be considered a success.

After training, participants required less visual information to make decisions with the same degree of accuracy. Participants' fixations tended to be spread more evenly among the features after training, reducing the number of redundant feature fixations. This indicates that after training participants learned how to interpret the significance of each weather feature more efficiently. Eye tracking was found to offer objective evidence of cue-based training's ability to affect and improve the decision making process. Eye tracking would be a suitable tool for assessing a training program's ability to affect behavior in many different environments. Further testing of the program's long-term effects on decision making bias and scanpaths are needed, however.

References

Ash, J. S., & Smallman, C. (2008). Rescue Missions and Risk Management: Highly Reliable or Over Committed? Journal of Contingencies & Crisis Management, 16(1), 37-52.

Driskill, W. E., Weissmuller, J. J., Quebe, J. C., & Hand, D. K. (1998). Evaluating the Decision-Making Skills of General Aviation Pilots. In O. o. A. Medicine (Ed.) (pp. 43): Federal Aviation Administration.

Gaba, D. M. H., Stephen K.; Small, Stephen D. (1995). Situation Awareness in Anesthesiology. Human Factors, 37(1), 12.

Goh, J., & Wiegmann, D. (2001). Visual Flight Rules Flight Into Instrument Meteorological Conditions: An Empirical Investigation of the Possible Causes. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, 11(4), 21.

Hampstead, B. M., Sathian, K., Moore, A. B., Nalisnick, C., & Stringer, A. Y. (2008). Explicit memory training leads to improved memory for face?name pairs in patients with mild cognitive impairment: Results of a pilot investigation. Journal of the International Neuropsychological Society, 14(5), 883-889.

Hunter, D. R., Martinussen, M., & Wiggins, M. (2003). Understanding How Pilots Make Weather-Related Decisions. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, 13(1), 15.

Jenkins, J., Shields, M., Patterson, C., & Kee, F. (2007). Decision making in asthma exaverbation: a clinical judgement analysis. Archives of Disease in Childhood, 92(8), 672-677.

Land, M. F., Van Gompel, R. P. G., Fischer, M. H., Murray, W. S., & Hill, R. L. (2007). Fixation strategies during active behaviour: A brief history. In Eye Movements (pp. 75-95). Oxford: Elsevier.

Santarcangelo, M., Cribbie, R. A., & Hubbard, A. S. E. (2004). IMPROVING ACCURACY OF VERACITY JUDGMENT THROUGH CUE TRAINING. Perceptual & Motor Skills, 98(3), 1039-1048.

Smith, P. J., Giffin, W. C., Rockwell, T. H., & Thomas, M. (1986). MODELING FAULT-DIAGNOSIS AS THE ACTIVATION AND USE OF A FRAME SYSTEM. Human Factors, 28(6), 703-716.

Starker, I. B., Richard A. (1990, April). A Gaze-Responsive Self-Disclosing Display. Paper presented at the CHI.

Wiegmann, D., Goh, J., & O'Hare, D. (2002). The Role of Situation Assessment and Flight Experience in PIlots' Decisions to Continue Visual Flight Rules Flight into Adverse Weather. Human Factors, 44(2), 9.

Wiegmann, D., & Shappell, S. (2000). Human Error Perspectives in Aviation. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, 11(4), 341-357.

Wiggins, M. (1999). The development of computer assisted learning systems for general aviation. In D. Ohare (Ed.), Human Performance in General Aviation (pp. 153-172). Aldershot, England: Ashgate.

Wiggins, M., & O'Hare, D. (2003). Expert and Novice Pilot Perceptions of Static In-Flight Images of Weather. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, 13(2), 15.

Wiggins, M. W., & Bollwerk, S. (2006). Heuristic-based information acquisition and decision making among pilots. Human Factors, 48(4), 734-746.

Wiggins, M. W., & O'Hare, D. (2003). Weatherwise: Evaluation of a Cue-Based Training Approach for the Recognition of Deteriorating Weather Conditions during Flight. Human Factors, 45(2), 9.