Graphical Weather Information System Evaluation: Usability, Perceived Utility, and Preferences from General Aviation Pilots

Kara A. Latorella
Crew/Vehicle Integration Branch, NASA Langley Research Center, Hampton, VA 23681

James P. Chamberlain
Crew Systems and Operations Branch, NASA Langley Research Center, Hampton, VA 23681

This Paper is a work of the U.S. Government and is not subject to copyright protection in the United States.

ABSTRACT

Weather is a significant factor in General Aviation (GA) accidents and fatality rates. Graphical Weather Information Systems (GWISs) for the flight deck are appropriate technologies for mitigating the difficulties GA pilots have with current aviation weather information sources. This paper describes usability evaluations of a prototype GWIS by 12 GA pilots after using the system in flights towards convective weather. We provide design guidance for GWISs and discuss further research required to support weather situation awareness and in-flight decision making for GA pilots.

INTRODUCTION

GENERAL AVIATION WEATHER ACCIDENTS - Eighty-five percent of the aviation accidents that occurred from 1990-1996, and nearly eighty-five percent of the accident fatalities, involved small GA airplanes. One major contributor to aviation accidents is hazardous weather. This equates to, on average, eleven weather-related GA accidents per week. Desktop simulation, and other laboratory experiments demonstrate pilot errors that corroborate the implication of this accident statistic [1,2,3,4,5,6]. Following the 1997 Gore commission on Aviation Safety, NASA initiated the Aviation Safety program (AvSP) with the goals of reducing the aircraft accident rate by a factor of 5 within 10 years and by a factor of 10 within 25 years. Within AvSP, the Aviation Weather Information (AWIN) program element aims to contribute to these goals by improving the weather information available to aviation users.

General aviation is particularly affected by convective weather. A survey of GA accidents from 1982 to 1993 [7] revealed that while only 3.5% of these accidents were directly attributed to thunderstorms, a large percentage of these accidents, 66%, resulted in fatalities. Convective weather is challenging because it can include severe/extreme turbulence, gusts, hail, icing, lightning, reduced ceiling and visibility, instrument meteorological conditions (IMC), and possibly severe downdrafts and microbursts. Such concomitant weather phenomena were analyzed separately in the AOPA accident analysis. Therefore the incidence of GA accidents attributed to convective activity, and the fatalities resulting from such weather systems is likely under-represented by the percentages cited for only thunderstorm effects.

GENERAL AVIATION WEATHER INFORMATION - Today's pilots of small GA aircraft principally rely on aural sources and external, or “out-the-window,” weather cues for weather information. Aural sources can include direct queries to Flight Service Station (FSS), En Route Flight Advisory Service (EFAS, or “Flight Watch”), and Air Traffic Control (ATC) personnel, as well as monitoring these radio frequencies to overhear other pilots' comments, queries, and the information supplied to them by ground-based professionals. Pilots can also tune in automated weather information services such as HIWAS, AWOS/ASOS, and ATIS to obtain a broadcast of conditions over a large area or at specific reporting stations. Unfortunately, the information available from broadcast aural sources is limited and, when weather becomes a problem, the frequencies used to obtain "live" aural information become saturated, making this information inaccessible at exactly the time it is most needed. When asking pilots to describe their current communication method for obtaining weather information during GA flights, 78.4% of the answers relied on radio communications [8]. Lack of clarity in (35%), and loss of (29%) these radio communications were most prominently mentioned as limitations on the effectiveness of this method for obtaining weather information [8].

Currently, pilots of small GA aircraft have limited in-flight information about convective weather activity, especially when compared to that available on larger
onboard sensors, data-link equipment, ground-based to determine what information is best acquired through display. Decision requirements therefore should be used in the information infrastructure that supports the GWIS. The design of a GWIS involves determining the requisite datalink capability, and onboard weather radar systems that are available for small GA aircraft are typically expensive, and limited in performance by size and power constraints. When available, these systems can provide improved weather awareness for severe weather hazards, but are limited in range and accuracy [9]. Onboard weather radar systems are workload-intensive to use accurately [10], are subject to attenuation, have a limited range, and provide information that is primarily forward of the aircraft and at the aircraft's altitude [11]. While these systems show severe local weather to avoid, they do not provide the more comprehensive weather picture required to fully support strategic planning or avoidance maneuvers. More accessible, complete, and usable weather information would benefit pilots’ situation awareness, decision-making, and safety. Graphical presentation is a more appropriate representation for this type of information [12], can more effectively be integrated with other flight information (e.g., terrain) and can be extended using symbols.

The FAA Flight Information Services Data Link (FISDL) program will soon make data-linked weather information systems widely available to GA pilots via commercial FISDL vendors. FISDL vendors provide, for no service charge, uplink of textual aviation weather products, including weather observations (METARS & SPECIs) and forecasts (TAFS) of terminal environments, as well as reports of severe weather conditions (SIGMETS, Convective SIGMETS, AIRMETS, and severe weather forecast alerts) and pilot reports (PIREPS). For a fee, GA pilots may augment this basic information. One of the first available graphical products is a national weather radar mosaic (NEXRAD). The FISDL textual and graphical weather information is broadcast by a network of VHF ground stations, and received and displayed by an onboard GWIS. NavRadio Corporation (now part of the Bendix-King Division of Honeywell International), in a cooperative agreement with NASA AWIN, developed the prototype GWIS used in this study and was subsequently selected as one of the two FISDL program participants. The other FISDL participant is ARNAV Systems, Inc.

GWIS ASSESSMENTS - The design of a GWIS involves many human factors considerations. The entire system must be designed such that information is available to the user that is reliable and temporally and spatially relevant to the decisions it supports. These considerations determine the requisite datalink capability, and onboard sensors and processing units, and constrain the design of the information infrastructure that supports the GWIS display. Decision requirements therefore should be used to determine what information is best acquired through onboard sensors, data-link equipment, ground-based sensors or observations; if uplinked, what the size of the packets are, the form of communication messaging (broadcast, request/reply), and the rate of uplinking new information. The content of the information available on a GWIS must also be selected to support flight decisions and support a pilot’s ability to assess the reliability of the information available. The manner in which this weather information is presented should facilitate interpretation of the location and intensity of weather phenomena, assist pilots in determining the relevance of weather to the mission, and assist pilots in projecting the location and intensity of weather phenomena over time. GWISs may also include aiding functionality to assist pilots in intelligently acquiring relevant weather information, determine when weather presents a hazard to the mission, and suggest or evaluate actions in response to hazards detected. Finally, the more mundane considerations of human-computer interaction must also be well-designed to result in a usable system. These considerations include selection of usable input devices, legible fonts, easily navigable menu structures, appropriate screen resolutions, timely system response to inputs, meaningful coding and symbols. Finally, the system must be robust to the ambient conditions existing in aviation, and therefore be usable in turbulent conditions, as well as in direct sunlight and at night. Considering the breadth of human factors concerns in designing a GWIS, there is a dearth of research that directly addresses these concerns.

Principally, prior research focuses on simply answering the question “Does graphical weather information improve pilot decision making?” As one might expect, based on theory, access to graphical weather information can assist pilots. Pilots using a prototype GWIS in static and dynamic desktop and flight simulation experiments were shown to be more likely to acquire trend data, have a more comprehensive awareness [4], make better go/ no-go decisions, rate hazard levels higher, have more confidence in weather-related decisions, make fewer calls to ground aviation weather personnel [13, 14], make more correct decisions with graphical weather information than with either verbal or text alerts [15, 16], and use 5% less fuel and clear thunderstorms by 3 times the margin of pilots without a GWIS [17]. When used in a GA flight test, accompanied by terminal forecasts and surface observations, and integrated with a traffic information service, subjects commented enthusiastically on the utility of a GWIS [18]. More than 82% of subjects had positive responses to the utility of precipitation maps, surface observations, and terminal forecasts individually [19]. All subjects had a positive overall impression of the system; 88% indicating that it would be important to make available to GA operations [19]. Pilots using a GWIS in two-person crews in a commercial flight deck environment also enthusiastically embrace this technology [17, 20].

FAA FISDL and NASA AWIN jointly funded a simulation experiment at Research Triangle Institute (RTI)
to evaluate pilot weather flying with and without a GWIS similar to the prototype evaluated in this paper. In this study, subjects were in IMC, had access to an autopilot, and were not given a present position symbol on the GWIS display. Results indicated that while this GWIS increased awareness of the general location of convective weather, it did not improve pilot diversion decision-making (subjects did not understand the location of weather with respect to their position), increased workload for at least half the subjects, and reduced reliance on ground-based weather professionals [21, 22]. So, while demonstrating positive effects in a simulation environment, and being well received by pilots in a demonstration, it appears that GWISs may not universally improve flight decisions.

Prior research evaluating pilot performance with GWISs indicates some specific design decisions that will influence the usability and utility of these systems. Participants in a two-crew simulation experiment suggested that while radar summary and lightning strike information was critical, they only needed either an IFR/VFR category map or a page that showed symbols coding surface observation ceiling and visibility for stations on a map [17]. There was no clear determination which of these two redundant information sources was preferred [17]. The RTI study that failed to show a significant advantage of a GWIS' use on aviation decision-making attributed this result to pilots' failure to understand weather location with respect to their aircraft's position [21, 22]. Pilot complaints corroborated this interpretation [23]. A subsequent experiment, while failing to demonstrate a performance advantage, indicated that an own-ship position symbol reduced pilot workload [23]. This experiment also demonstrated that NEXRAD data resolution impacted a decision that required estimation of distance to a storm cell; that is, pilots with larger resolution NEXRAD data (8km square) made safer decisions than those with smaller resolution data (4km square) [23]. One early implementation of an uplinked radar mosaic GWIS, developed at MIT Lincoln Labs with funding from the FAA Datalink Operational Requirements Team (DLORT), had a 15 minute update rate, 6km-square resolution and employed a "lossy" algorithm (resulting in less well-defined precipitation areas) to compensate for lower available bandwidth (250bps) [24, 25]. In desktop usability assessments, all subjects found the high level of lossy compression unacceptable, and some found that the medium level lacked the functional equivalence of the uncompressed image [14, 26]. Prior research has also suggested that improvements are not limited to the presentation of graphical weather information, suggesting that METAR information should be presented in plain English translation, rather than in the standard ICAO teletype encoding [23, 27]. Prior research suggests that early GWIS prototypes reduce pilots' interactions with FSS and ATC, and thereby may reduce the overall view they have on the weather [21, 22]. A human factors evaluation of several multi-functional cockpit systems with GPS moving map overlays, demonstrates the variety of input devices available for cockpit information systems, and indicates that benchmark tasks are more or less easily accomplished, in large part, due to the usability of these devices and menu navigation [28]. This investigation also showed that some of these systems provide automatic brightness control to improve readability, while others allow for no opportunity to control brightness [28], a serious detriment to use in an actual aviation environment.

In summary, there are many aspects of GWIS design that, if improperly implemented, could result in pilots having an incorrect and/or incomplete understanding of the weather, as well as increased workload and an excessive opportunity cost to other flight deck duties. These design considerations include, but are not limited to: the information to be displayed (type of weather information, geographical references, etc.), the manner in which this information is presented (color coding, resolution, level of abstraction, symbology, etc.), and the characteristics of the physical interface that houses this system (input devices, brightness control, etc.). Rather than suspecting that GWISs in general are not appropriate technologies, we suggest that more work must be done on designing the information provided by these systems to be reliable, action-oriented [29] and decision-centered, and on designing the interface to these systems to be more easily and effectively used.

THE COWS EXPERIMENT - The AWIN Convective Weather Sources (CoWS) experiment investigates how GA pilots use weather information available from aural, “out-the-window” visual, and GWIS-displayed cues, to support in-flight decisions related to convective weather systems. While the focus of this research is to better understand how GWISs are used in in-flight decision-making, the experimental protocol also allowed us to conduct a usability assessment of the prototype GWIS implementation we used. Two earlier publications described preliminary CoWS results based on partial data collection. The first [30] discussed pilots' relative confidence, information sufficiency, and workload ratings when using aural, out-the-window visual, and graphically represented weather information cues in flight near convective weather. The second discussed the accuracy and consistency of the test subjects’ ability to identify convective weather relative to their aircraft location and flight track [31]. This paper discusses the results of the usability assessment and participants' comments on usability and utility of this GWIS.

OBJECTIVE

The International Standards Organization has defined the “usability of a product (as) the extent to which the product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction, in a specified context of use” [32]. In the CoWS experiment,
GA pilots were asked to assess the usability of a prototype graphical weather information system in the context of a flight in the vicinity of convective weather. While participants were not flying the aircraft, participants rated flight scenarios' validity as very high and they considered these experimental flights to be fairly representative, in terms of information available and workload, of their typical flights [30]. The objective of this paper is to present the usability data obtained for this prototype GWIS, by this user population and in this context, and discuss the implications of these results for improving the design of GWISs.

METHODS

APPARATUS - Apparatus for the CoWS experiment included supporting ground infrastructure, test aircraft, and the tethered GWIS. Four prototype AWIN/Honeywell broadcast VHF data link (VDL) transmitters were located in Virginia and provided a broadcast link of packaged weather data files to the test aircraft along four routes of flight (Figure 1, rings indicate 40nm broadcast range).

Figure 1. CoWS Experimental Test Range.

NASA Langley’s Raytheon B-200 Super King Air, a nine-passenger, pressurized twin-turboprop airplane, was operated at speeds and altitudes consistent with those of the smaller, piston-engine GA aircraft used by the participant population. The onboard GWIS included a VDL receiver, Global Positioning System (GPS) receiver, and two laptop PC’s with tether cables to two small handheld display units. The display unit screens were approximately 7.5cm tall by 10.5cm wide. Five bezel buttons (12mm x 6mm, each) on the right side of the unit actuated soft menu fields, and a rate-controlled joystick controls pan, zoom, and crosshairs for symbol selection (Figure 2). The unit presented lossless, nationwide radar mosaic imagery at 4-square-km resolution with a 6 minute nominal update rate assuming adequate broadcast reception, and surface weather observations (METARs) in text and symbolic form for reporting stations in the mid-Atlantic region. The display also presented contextual features (rivers, interstates, and state boundaries), airport identifiers, present position and track symbol, creation time stamp for the radar product (upper left corner), a scale legend (upper right corner), and indicates missing data (horizontal yellow stripe where data was missing). This graphical information could be viewed alone, or augmented using three other modes of information presentation: NEXRAD Mosaic, METAR, or both NEXRAD and METAR (NEX/MTR). When the METAR information was available, the text for selected METAR icons was available. When in NEXRAD-only or Graphics modes, the identifiers for airports or NAVAIDs were available. The user could select one of three GPS modes: GPS Off (no aircraft position/track symbol), GPS Lock (display centered on the aircraft's position), and GPS Free (aircraft symbol provided, but the display was not locked to this position). The joystick could be used in three modes: Scroll (to view that which is off the edges of the displayed map; not usable in GPS Lock mode), Crosshairs (to select METAR, Airport, or NAVAID symbols), and Zoom (to change map scales: 500, 200, 100, 50, 25, 10 nautical miles per 1.5cm, or 1/7th, of display width and 1/5th of display height). At the 500nm scale, the user could see the map for the entire continental United States. The interaction and informational elements of this GWIS are described in more detail with reference to results in a following section.

Figure 2 - The CoWS Prototype GWIS.

SCENARIOS - The ideal flight scenario operated under Instrument Flight Rules (IFR) but in Visual Meteorological Conditions (VMC). The test aircraft departed from NASA Langley/ Langley Air Force Base (LFI) on a flight path that, if allowed to continue, would obliquely intercept a frontal convective system of at least moderate intensity at approximately 120nm from top-of-climb, and at an altitude above the haze layer (typically 14,000 feet). The location of the GWIS’ ground-based infrastructure and other airspace considerations constrained the region in which these flights could occur. To accommodate this constraint and minimize training and materials, four potential IFR flight plans were developed from LFI to Hickory, NC [HKY]; Charleston, WV [CRW]; Abingdon, VA [VJI]; or Clarksburg, WV [CKB]. One of these four
flight plans was chosen on the morning of each flight based on prevailing weather conditions. Participants did not perform flying duties during these flights; a NASA test pilot served as pilot in command (PIC). The convective weather during the flight tests could generally be described as well-defined, significant lines and areas of cells, with surrounding towering cumulus buildsups and occasional embedded cells. Flight conditions were generally unrestricted in visibility, on top of any lower cloud layers, and laterally clear of towering cumulus and cumulonimbus cells. We attempted to achieve these weather conditions on the outbound leg of all flights. During the inbound portion of the flights, we typically tried to fly as close as 20nm to interesting convective cells.

EXPERIMENTAL CONDITIONS AND DESIGN - The experiment collected data according to a within-subjects experimental design. Three pilots, constituting a team, participated in each flight. The team flew three times, allowing each participant to experience each experimental condition. On each flight, one participant was allowed to use the GWIS, and also received aural weather information during the outbound experimental phase of the flight. This participant completed a usability survey during the inbound phase. One of the other two participants received an out-the-window view as well as the aural information. The other participant received only the aural weather information. Participants who had used the GWIS on a prior flight were allowed to look at the GWIS briefly during the inbound phase. The aural weather information included listening to a HIWAS broadcast station, getting a Flight Watch briefing and getting an ATC report. The aural-only condition represents what is available to most GA pilots during IMC today. The condition with aural information and the out-the-window view represents what GA pilots have for weather information in today's VMC. The condition with aural information and the GWIS represents what GA pilots might have in the near future in IMC.

PARTICIPANTS - Participants were recruited from local regional airports and through advertisement. Applicants reported their flying experience and weather exposure on a Background Questionnaire. Participant selection criteria included: an instrument rating, 10-50 flight hours within the last 90 days, and 50-1000 cross-country or 100-2000 total flight hours. In addition, participants were not selected who had worked for a scheduled air-carrier in the prior year or who had participated in the aforementioned RT/IAWIN experiment. Weather experience has been found to significantly affect weather-related decision-making and information acquisition (Wiggins & O'Hare 1995) so candidate participants were clustered into 3 groups of “exposure experience” using cross-country hours. The midpoints of each cluster are 135 (low), 379 (medium), and 738 (high) cross-country hours respectively ($p < .0001$). Twelve participants were selected to form four three-member teams, each team composed of one participant from each of the clusters to balance exposure experience across flight scenarios. Cross-country experience level ranged within teams and was counter-balanced over teams to mitigate concerns about generalization to the participant population and (experience x flight) interactions. Cue assignment to participant experience levels was counter-balanced to mitigate concern about (cue x experience) level interactions.

PROTOCOL - When a participant team arrived at NASA Langley in the morning, each participant was provided with an introductory briefing, consent form, schedule, and Preliminary Questionnaire. They then received a mission motivation and briefing; a local terrain, NAVAID and airport identifier review; a route briefing for the flight to be taken; and practice on forms and procedures to be used during the outbound experimental phase. Following a short break, participants had 40 minutes in total, to review a standard preflight weather briefing composed of a DUATS text briefing, associated weather graphics, and a pre-recorded briefing from a Flight Services professional. Participants then completed the Preflight Weather Situation Awareness questionnaire. While other participants completed knowledge tests and personality inventories, the participant who was assigned to receive the AWIN GWIS was trained on the display. This training used a scripted Microsoft Office PowerPoint presentation with digital photographs of the actual screens and a scripted aural instruction to introduce the GWIS. This training device was designed to allow participants to interact with a representation of the actual display, albeit along the lines of the script, and to provide a standardized training. The training explained the interface control features and modes; and described the information presented by the system, including symbols and color-coding. Following this standardized training, participants received a comprehension survey. This survey required participants to interpret screen symbols and color conventions, locate information regarding age of weather data, demonstrate knowledge of the menu structure and display modes. The results of this survey were used to indicate requisite compensatory training. Usually only a few items, if any, required this compensatory training. Participants were able to explore the actual system during ground operations and during ascent once aboard the test aircraft, and were given a quick review by an experimenter.

The in-flight portion of the experiment began after the aircraft had climbed to cruising altitude and when the aircraft was approximately 120nm from the first convective weather area of moderate or greater intensity. The outbound leg of the in-flight portion concluded when approximately 20nm from this area, or at approximately 100nm from the initial experiment starting point, whichever occurred first. Throughout the outbound phase of the flight, Weather Situation Awareness (WXSA) questionnaires were given every 8 minutes (approximately every 25nm), and Position Update tasks and aural
weather information were alternately provided between the WXSA questionnaires, such that each was provided approximately every 16 minutes. Each of these events is described below.

The Position Update task was designed to compensate for the loss of positional awareness and workload induced by not piloting. For this task, participants copied scheduled reports from the pilot in command (airspeed, altitude, heading, position, next waypoint, and current time) onto a prepared form; plotted position on an IFR low altitude en route chart; and calculated elapsed time and ground speed. They were also required to note any ATC transmission affecting the flight. All participants received scheduled aural weather information. The first aural cue was obtained from a local automated Hazardous In-Flight Weather Advisory Service (HIWAS) broadcast outlet, the second from querying EFAS personnel, and the third from querying ATC. The WXSA questionnaires were handed to participants at the indicated times. The WXSA items addressed participants' weather situation awareness and flight decisions. In particular, participants were asked to identify the location of the nearest convective cells. Participants were instructed that there would only be enough time to complete the WXSA if they rely on their "mental snapshot" of the weather they have at the time the questionnaire is administered.

At the conclusion of the outbound leg, participants were asked to plot the aircraft's position on their en route IFR chart, draw weather within 50nm of the flight path on the chart, and complete the Inbound Questionnaire. This instrument contained NASA-TLX [34] -derived scales for workload assessment, asked participants to indicate other weather sources that would have been helpful, and about their flight decisions. After completing the Inbound Questionnaire, the participant using the GWIS was asked to complete a Usability Questionnaire, which included some QUIS [35] items, and provide any additional comments he had. Following the flight, participants were provided with a short debriefing questionnaire for that flight. At the conclusion of the third flight for a team, when all participants had been exposed to the display, participants and experimenters more fully discussed issues of experimental validity and display usability. These final debriefing sessions typically lasted approximately two hours. Data used in this paper is extracted principally from the Usability Questionnaire and notes from final debriefing interviews. Selected elements of the WXSA Survey and the Debriefing Questionnaire augment these principal sources where they address related issues.

RESULTS

Participants' responses on the usability, debriefing, and inbound questionnaires are reported below with annotations from extensive debriefing sessions. These data reveal users' perceptions on the general usability of this GWIS, usability of interaction elements, and usability and utility of information elements. Protocols from debriefing sessions provide the basis for discussions of how these participants would use this GWIS, how use of this system would affect situation awareness, flight deck decision-making, collaboration among NAS users, and pilot workload. One-tailed t-tests were conducted on rating scale data to test if values were significantly ($\alpha = .05$) greater than 50% of the scale, and greater than 75% of the scale. For the remainder of this document, where t-test statistics are significant for means greater than 75% of the scale, the adjective "very" is used to describe the average rating response. For t-test statistics significant for means greater than 50%, but not significant for means greater than 75%, the adjectives "fairly" and "marginally" are used to describe the average rating response. Because debriefing sessions were relatively unstructured, numbers of participants cited as commenting on a particular aspect of the system should be interpreted as indicative of the salience of this issue to participants, but not necessarily a percent agreement of the sample population.

GENERAL ASSESSMENT - Generally, participants found the functionality of the GWIS fairly adequate. In particular, they were enthusiastic about the advantage of having graphical NEXRAD weather in flight, to the point that other interface deficiencies appeared relatively insignificant to them. Participants' ratings, on average, indicate that they were fairly comfortable in describing their interaction with the system as more "wonderful," than "terrible;" and more "satisfying," than "frustrating." They were very much more willing to describe their experience with this system as "easy," than "difficult." Participants' comments regarding the general usability of this GWIS were fairly positive, and generally reflected that the utility of having a graphical weather information system on board outweighed any specific concerns they had with the interface. Rasmussen and Vicente [36] emphasize the importance of designing systems such that they invite and are robust to exploration and learning in the operational environment when possible. The ability to easily learn and explore the system is therefore important aspects of general satisfaction with this system. Participants' ratings associated with the ease of learning and exploring this system were also positive. On average, ratings indicated that the participants thought that learning the system and advanced features was fairly easy, getting started with the system was very easy, and the time to learn to use the system was fairly fast. On average, ratings indicated that the system supported participants' need to explore features; that the system was very
encouraging of this exploration, that it was very safe to explore features, and that discovering new features was fairly easy. Participants generally stated that they had a comfortable familiarity with the system after desktop and preflight training. One participant felt that additional training was still necessary to use the system effectively and commented that he experienced information overload with the system.

As we detail below, however, the enthusiasm with which participants generally regarded the system and their ability to start using it is less evident when they respond to questions about particular aspects of the system. We present their responses according to the aspects of the GWIS that are used to manipulate the information shown, the Interface Control Elements, and the information that is displayed on the GWIS and how this information is displayed, i.e., Presentation and Information Elements.

INTERFACE CONTROL ELEMENTS - The physical interface of this system includes a knob for adjusting brightness, a small rate-controlled joystick, and five bezel keys. Participants were not explicitly asked about use of the brightness knob, but few were observed to use it. Only one participant expressed displeasure with the joystick control, finding it too sensitive. The bezel keys determine the mode of the joystick, the type of weather information displayed, how GPS position information is used, and whether the soft labels associated with these keys are shown. One of these keys also, depending on the type of weather information displayed, allows access to METAR text information and enables labeling of selected airport and NAVAID symbols. Participants did not comment on the physical interface of the bezel keys.

Weather information modes can be either “Graphics” (no weather information, but all contextual features, and access to airport and NAVAID identifiers), “METAR” (contextual features and surface observation symbols, with access to METAR text), “NEXRAD” (contextual features and composite NEXRAD imagery, with access to airport and NAVAID identifiers), and “NEX/MTR” (contextual features, NEXRAD imagery, and surface observation symbols, with access to METAR text). The utility and usability of the weather information displayed in these modes is further discussed as Information Elements in the following section. The last bezel key allows users to turn the displayed menu labels off. When the menu labels are on, the right most approximately 20% of the screen is obscured. Pressing any of the hard keys will then redisplay the displayed menu labels. Two participants noted that the range scale for the NEXRAD data disappears when the menus are turned off.

The joystick has three modes: Zoom, Crosshairs (for selecting), and Scroll. The Zoom mode enables users to, with the joystick, select one of six scales: 10nm, 25nm, 50nm, 100nm, 200nm, and 500nm (showing the continental United States). These distances refer to the label and size of a reference bar that is in the upper right corner of the display. The area displayed is 7 times the width and 5 times height of this scale bar when the menu labels are off. Participants were asked to indicate the scales they used during the experiment. The distribution of these responses (Figure 3) indicates that these data are fairly normally distributed around the most popular 50nm scale. During debriefings, some participants mentioned the scale they would select if the GWIS did not allow them to change scales. Two participants favored the 100nm scale, but one of these noted that he would want that to be smaller if there were weather in the vicinity; one responded 50nm, and one 25nm. The 50nm scale is the largest scale that includes all the contextual and aviation location information. Two participants indicated that also having a scale that shows 2 to 5nm resolution would aid ground movement.

![Figure 3. Reported Frequency of Map Scales Used](image)

The Crosshairs mode allows users, with the joystick, to move a vertical and a horizontal line to select an airport or NAVAID symbol or a surface observation symbol. Depending on the mode of weather information selected, a bezel key then provides either the identifier of the airport or NAVAID, or the text associated with the surface observation symbol. Again, participants’ opinions regarding the usefulness of the airport, NAVAID, and METAR symbols and text are described below as information elements. Participants’ complaints about menu navigation were primarily focused on the awkward method of obtaining METAR text information and NAVAID/airport identifiers. At worst case, this required a 5-step process: selecting a scale that shows symbols of interest (50nm or lower), selecting a weather mode corresponding to the information you want (NEX/MTR or METAR for METAR text, Graphics or NEXRAD for airport or NAVAID identifiers), changing to crosshairs (selecting) mode, orienting the crosshairs over a symbol (surface observation, airport, or NAVAID), and finally selecting the bezel key to acquire the METAR text or switch between airport/NAVAID identifiers. This process was considered onerous enough to motivate several suggestions for
read at normal hand-held viewing distance, and the display was difficult to read by subjects using bifocal lenses. Participants also suggested that the aircraft symbol is distorted when not flying in a cardinal direction and making it difficult to clearly see the track at a glance.

**Scale Legend** - The scale legend is on the top line of the display, in the upper right corner, and provides a line 1.5 cm long with end hashes followed by the number of miles represented by this line on the display and appended with "NM" for "nautical miles." Experimenters observed that participants often used a pencil or their fingers to make a "ruler" for the unit of scale displayed in the legend, and then determined distance from aircraft position to a display element of interest (weather, airport, NAVAID). While no participants explicitly mentioned that they did this, their discomfort with the manner in which distance information was conveyed is obvious in other comments. Four participants volunteered comments indicating that they would have preferred to have range rings around the aircraft position symbol to help determine the distance between weather and the aircraft position. Another participant volunteered that he would have preferred to use the crosshairs to select the display element of interest and to be provided with the bearing and distance from the GPS-derived aircraft position.

**Menu Key Labels** - Participants' comments from debriefing did not include any assessment of the menu key labels, and these were not addressed in the usability rating scales. Observations from training suggested that the term "graphics," which refers to the display mode without weather information and only contextual and aviation symbols, was not intuitive for participants. Several initially misinterpreted this to mean the "graphical" weather that was the most salient feature of this new technology. Participants also noted that the menu labels weren't aligned well with the bezel buttons.

**Weather Radar** - NEXRAD returns were considered very helpful, and the colors used to encode intensity levels were also considered very helpful. All ten participants who responded to the question indicated that they considered radar return color codes to represent categorical levels of radar intensity, rather than reflecting specific values of atmospheric phenomena (i.e., VIP levels). Five participants expressed dissatisfaction with the resolution of the NEXRAD data. When using the lowest scale, units of NEXRAD information appeared too big to be useful. Other participants suggested appropriate resolutions: one said it should be about 0.25nm, another suggested that it should be the same as that which is available on onboard weather radar. Resolution was seen as particularly important for understanding the gradient of NEXRAD weather information intensities. Post-experiment debriefing sessions revealed that participants were not all well informed about the construction of

PRESENTATION & INFORMATION ELEMENTS - The GWIS evaluated here provides a plan view of NEXRAD data, METAR symbols and text, contextual features (rivers, interstates, state boundaries, airport symbols and identifiers, NAVAID symbols and identifiers), and aircraft position symbol. Arrangement of information on the screen appeared to be fairly logical to participants. While two participants stated that the resolution of the display (320 x 200 pixels) was adequate, three participants desired higher resolution. Most comments suggesting the need for improved resolution indicated that text labels (e.g., NAVAIDs, airport identifiers) were too small to be read at normal hand-held viewing distance, and the display was difficult to read by subjects using bifocal lenses. Participants also suggested that the aircraft symbol is distorted when not flying in a cardinal direction and making it difficult to clearly see the track at a glance.

There are three modes for the use of GPS position data: "Off," "Lock," and "Free." These modes determine constraints on the viewable window of the available Continental US map, and whether an aircraft symbol is present. The GPS Free and Lock modes present an aircraft symbol. The GPS Lock mode constrains the viewable region of the map to that which is centered on the aircraft's position, thereby providing a 360° view around the aircraft for the distance indicated by the chosen scale. Therefore, in this mode, the scroll mode is inhibited, and one can only select items that are within the viewed region of the display as defined by their position relative to aircraft position and the map scale chosen. Only one participant indicated that he had forgotten that the scroll function was inhibited during the GPS Lock mode. While two participants explicitly preferred the 360° plan view around the aircraft, one mentioned that he used the GPS Free mode to provide more of a forward view of the weather than backward view. Participants' perceived use of and comments on the aircraft symbol are further described below as an information element.

The display presents information in a North-up configuration. One participant mentioned the need for an interface control element to allow one to change to a Track-up display orientation control. Two participants discussed having an interface control or automatic reorienting of the display; to be North-up when used strategically, and Track-up when used tactically. One can effectively convert this tethered display to a Track-up presentation by turning the display to orient the direction of the aircraft symbol to point up. While the text labeling is difficult to read in this orientation, it would provide the desired orientation of graphical weather information, contextual features and aircraft symbol. Most participants did not realize this "feature" of a tethered system, as few were observed to reorient the display during the flight experiment.
The construction of NEXRAD images was not included in the training of the AWIN display.

The NEXRAD product creation date and time is displayed in the upper left corner of the display and is shown as, for the product delivered on July 7th at 17:41 Zulu time (or 1:41PM EDT): “NEXRAD 07/07 17:41Z.” Half of the participants did not recall using the age of the NEXRAD information in their use of this information, even though, on average, participants’ ratings indicate that the age of this information is fairly apparent. Only five participants attested to using the age information. The average rating of how apparent the age information is was about 90%. Only two of the six participants who confessed to not using the age of the NEXRAD data rated the degree to which this information was apparent. One of these rated it about 97% apparent, the other about 18% apparent. Subsequent commentary forms indicated that the participant who confessed to not using this information, but rated it as very apparent chided himself for not using this and later remembered that this information was available. One participant commented that the age of the information should be detectable “at a glance,” implying that it wasn’t. This comment, along with the rating scale results clearly demonstrates the difference between a bit of information being visually available, and being attended to, and therefore available for further processing.

Seven of twelve participants were less than 50% confident that they knew what the longest delay in NEXRAD weather updates was during their flights. Three of these seven participants rated their awareness of the longest delay as 0%. Seven participants rated the degree to which the age of the NEXRAD data was acceptable, and, on average, rated it less than fairly acceptable. In free form discussion, participants’ comments about the age of weather information predominantly referred to that of the NEXRAD information. Generally these comments implied that the current best-case update rate, 5 minutes, is adequate for strategic use; but that faster update rates are required for tactical use; and slower update rates, particularly over 10 minutes old, are insufficient for appropriate use in flight. Several participants indicated that they would appreciate an alert to indicate when this weather information is “old.” When asked to assume 200 knots true airspeed, four participants indicated that the criteria for this alert would be at 10 minutes, one suggested at 15 minutes. While avionics manufacturers and FAA usage guidelines intend for pilots to use these GWISs only for strategic purposes, participants’ comments clearly indicate that they will use these systems to support tactical decisions as well. This was particularly obvious in responses to the question, “how old do you think the weather information was?” Some of the more concerning responses were: “I can’t remember… (I) assumed it was real-time;” “(I’m) so used to considering it real-time;” “(I) didn’t notice… could thread the needle with it;” “good enough to make a divert decision.” This type of response is particularly troubling when considered in conjunction with the finding from previous analyses from this study and others, that pilots using a GWIS are less likely to request additional information from other sources (i.e., Flight Watch). One participant did recognize the danger inherent in using delayed weather information and mentioned that he compensated for this by looking only at weather very close to the airplane symbol. This strategy would not compensate for old weather information. At times during the return trip, after the experiment was over, weather information was not updated for as long as over 40 minutes.

Despite this displeasure with, and failure to appreciate, the age of the NEXRAD information, participants rated this weather information source as fairly reliable. Inspection of individual scores revealed that four of the participants rated this information as less than 75%, where 100% is reliable, and two participants rated it less than 50% reliable. It is important to realize that these ratings were taken for different flights, and that the experiences during these flights may have differed. While the lower scores for age and reliability may be less prevalent than more acceptable ratings, the fact that these reflect more challenging scenarios for the equipment cannot be separated from rater bias.

When NEXRAD data were not available for a section of the map, an opaque yellow bar was presented for the width of the screen in the region for which there was no data. There could be several of these “missing data” blocks on a screen if the up-linked file was incomplete in different places; and, where these lacking areas were adjacent, wider bars formed. Participants did not fail to appreciate the meaning of this display element. However they question the formatting, and conditions for use. Several participants expressed the opinion that at some point there is a tradeoff between a newer image with mostly missing data, and an older image that is complete. Two participants recommended continuing to display older data when newer data is significantly degraded or there is no service, and provide an alert to indicate that the data is aged.

Surface Observations - Participants viewed the surface observation METAR symbols as not particularly useful. Debriefing comments corroborated this rating scale result. While participants were instructed to consider that they would be landing at a given destination in the given scenario, three participants indicated that these symbols were not very useful for the enroute portion of their flight, but would be of more use if they were actually landing or considering an alternate. METARS may also have been more useful in a weather scenario with widespread low ceiling and visibility conditions, rather than frontal convective systems. Color-coding of the METAR symbols (the upper half coded for ceiling, the lower half for visibility) was rated as fairly helpful, and these color codes were always perceived as categorical indications. While participants generally considered color-coding to be
effectively used for individual types of weather information, two participants indicated that the surface observation symbol colors were difficult to determine when combined with NEXRAD data. One of these participants, as well as a different participant, mentioned that they didn’t really use or they did not select the METAR graphical symbols in order to reduce screen clutter. Two participants indicated that they used the color-coded surface observation symbols to indicate a trend in surface conditions over a region, and to indicate regions where conditions require IFR. One of these noted that these symbols could represent old data.

Nine of the twelve participants’ ratings, on average, indicate that they found the age of the surface observation data to be fairly apparent, or obvious. Four of these ratings, however, were below 80%, and two below 50%. These same nine participants, on average, found the age of surface observation data to be less than fairly acceptable. Five of these ratings were less than 75% acceptable. Comments regarding the age of the data appeared to focus on the age of the NEXRAD information, most likely due to the priority that they ascribed to the NEXRAD information as previously mentioned. One participant did explicitly comment that he did not consider the age of the METAR information when using it. When asked about their confidence in knowing the longest delay for METAR information, eight of the eleven respondents’ ratings were less than 52% confident, and three of these expressed no confidence in their knowledge of the most aged METAR information.

The only way to determine the age of the METAR information was to access the METAR text page. However five of the twelve participants indicated that they relied principally on the graphical surface observation information symbols, not the METAR text, to understand surface conditions. METAR text in this GWIS was coded and presented, full page, on the screen; thereby obscuring all other information. Participants considered the content of the METAR text screens to be very helpful, and the form very easy to read. Participants did not complain about coded METAR information and all were able to decode all but the special remarks during training. Format of the METAR screens were fairly easy to read. There was wide variance in the number of METARs participants reported accessing (Range = 2, 20; Mean = 8.68; Median = 7.5). On average, the twelve participants’ ratings indicate they considered the graphical surface observation information to be fairly reliable.

Several participants commented that METAR information should be accessible without having to sacrifice the graphical information on the display. Solutions to this problem included spoken METARS, and a dedicated METAR text line at the bottom of the display. One participant expressed the opinion that this is what FSS is for, that it would be easier for him to call and have the METAR text information read to him than to read it on this system. Participants volunteered solutions to this problem which included spoken METARS ("text is inexcusable now!"), a dedicated METAR text line at the bottom of the display, automatic presentation of the METAR text in this dedicated screen area when crosshairs are over a surface observation graphic, and an automatically-generated listing of the surface observations that are relevant to a route - to eliminate picking through the menus and selecting symbols.

NAVAIDS/ Airports/ Contextual Features – Several contextual and aviation location display elements provided users with references to understand where the weather was, and how far they were along their route. The display showed state boundaries as white lines. At the 50nm scale and below, the display showed major rivers as blue lines, and interstates as yellow lines. At the 100nm scale and below, the display could be configured to show grey rectangles, indicating certain airports, and blue circles, indicating certain NAVAIDS. While all twelve participants reported using the airport symbols, their ratings indicated these symbols were not particularly helpful to them. Eight participants reported using the NAVAID symbols, and their ratings indicate that these symbols were, on average, only marginally helpful.

One participant complained that the airport and NAVAID symbols were of limited use because those that are in the system database don’t include those that he would find most useful for orienting himself along the route and estimating the location of weather from aural sources. Specifically, he questioned why TACANs, major VORTACs, and NDBs were not included in the database for display. Further, the display allowed a user to display only one label of a NAVAID or airport at a time. Seeing successive airports or successive NAVAID labels, only requires one to re-orient the crosshairs. However, seeing the label for an airport after a NAVAID (or vice versa) requires not only reorienting the crosshairs, but also selecting the appropriate menu key to change this mode. This problem was mentioned in the same breath as the comments regarding the inaccessibility of the METAR text information, and the solution of a dedicated text line was seen as appropriate for solving this problem as well. One participant suggested that if the screen resolution were better, more labels could be concurrently available. Whatever the solution, the current design appears to not adequately support users’ requirements for orienting themselves with respect to geographical and aviation features and the weather information displayed, “Now it takes too much time, workload, brainpower.”

Airplane Symbol/GPS Modes - This display received input from a GPS to indicate aircraft position. The display has three modes for using this GPS information. The “GPS-Off” mode does not use the GPS information. The “GPS-Lock” mode displays a magenta aircraft symbol in the position indicated by the GPS, and locks the displayed portion of the map to be centered on this aircraft symbol.
The “GPS-Free” mode displays the same magenta aircraft symbol, but the display is not centered and locked on this symbol. The aircraft symbol’s shape indicates direction of movement. The positional information provided by this aircraft symbol was judged to be very helpful. While most of the participants (10) said they switched between GPS modes, all but one preferred the GPS-Free mode.

**USING THE GWIS** - In addition to assessing usability aspects of the individual features of this GWIS prototype system, we were interested in how pilots’ weather situation awareness and decision-making were affected by this system. Rating scale data, but more so debriefing commentary afforded insights into how these systems, as perceived by the participants, would affect general aviation safety.

**Situation Awareness** - Situation awareness (SA) is generally defined as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” [37]. Endsley [37] has further described three levels of SA: Level 1 SA refers to the perception of cues in the environment. Level 2 SA refers to the comprehension of these cues; including the integration of cues, assessment of priority and relevance to mission goals. Level 3 SA refers to an understanding of how these cues will change in the near future, and serves as a prediction of the dynamics of the system in which the human operator is immersed.

Preliminary results from other data obtained in the CoWS study shows that pilots who have access to the GWIS and aural information have much better confidence in their weather picture than those without the GWIS; and have about the same confidence as those who could see out the window and had aural weather information [30]. Pilots using the GWIS and aural information would have requested significantly fewer additional sources of weather information during the scenarios than those with either only the aural information or with aural information and a view out the window [30]. Finally, the GWIS+aural sources condition had the highest average score on the NASA-TLX “Perceived Performance” scale [30]. These preliminary data suggest that, particularly in IMC, participants believe their SA is significantly improved by a GWIS, and may be similar to that obtained in today’s VMC conditions.

Level 1 SA for aviation weather includes the location and intensity of the weather information. Generally speaking, participants commented that the weather and positional information presented by the GWIS was an improvement over that available without it. One participant explicitly stated that this GWIS “gives superior position information and weather situation awareness.” Debriefing comments and results from rating scales associated with contextual, aviation and positional symbols indicate that one of the major problems with using aural weather information today is that it is difficult to develop a spatio-temporal representation of this verbal, transient information. This was a difficult task even for those participants who had access to both the aural and GWIS sources. One source of difficulty was noting/remembering the 3-letter station identifiers used by FSS and automatic services to describe weather locations in the often rapidly delivered weather information transmissions, “... can’t figure out the landmarks and geographical references; and by the time you do, he’s on to something else.” One might expect that this would be easier for those participants with a GWIS display. However participants using this technology still seemed to have difficulty. We believe that this can be attributed to the fact that the aviation location identifiers on the GWIS display were not identical to those used by FSS to define weather boundaries. Further, for large weather systems, one would need to be on a scale larger than the 50nm scale to fully understand the extent of their coverage. This prototype system did not display aviation location information on scales larger than 50nm. Participants’ ratings indicate that the NEXRAD colored graphics helped them determine the location of hazardous weather, but that this was limited by concerns regarding resolution and update rate of this graphical information. Some participants seemed to want more information about surface weather conditions, either for airports not represented or between airports, than was displayed by the system. These participants inferred conditions for these points by extrapolating conditions between surface observation symbols: “(doing so) gives trend information for what’s IFR (between stations).”

Level 2 SA for aviation weather information includes an understanding of how relevant existing weather is to the individual pilot, aircraft, and mission. Pilots must estimate the distance and bearing to the weather that is identified, and estimate the ramifications of its intensity for their experience level, aircraft characteristics, and mission. Comparing subjective estimates of weather hazard existence, and distance and bearing; with objective assessments indicates that participants whose aural weather information was augmented with either a window view or the GWIS had a better understanding of weather hazards, and participants using the GWIS had better hazard detection than either of the other two conditions [31]. Generally, the window+aural condition supported the best distance and range estimates, but also resulted in the most false identifications of hazards, i.e., indicating a hazard that did not exist [31]. Results from this comparative analysis indicated that all three sources of weather information provide unique benefits for comprehensively developing good weather SA. Participants’ comments calling for improved scaling indications, e.g., range rings, indicate that they were not comfortable with the level of support the GWIS provided them in achieving accurate distance and bearing estimates.
Another aspect of Level 2 SA is awareness of how relevant the weather information is for flight decisions. One must be aware of the accuracy, and temporal relevance of the information provided to determine its relevance to the current mission goals. Most participants reported feeling less than 50% confident in their awareness of how old the NEXRAD data was, and a few of these indicated they had 0% awareness of how old this information was. Similarly, most participants were less than 52% confident in, and three expressed no confidence in, their knowledge of the most aged surface observation information. Five participants indicated that an alert should be provided to indicate when NEXRAD weather information is "old." Participants also indicated additional information sources that they would use to assess the validity of the information provided by the GWIS. These are discussed with reference to decision-making in the following section.

Level 3 SA for aviation weather information is an understanding of the dynamics of weather systems and the ability to predict changes in its location and intensity in the near future. Eight of the twelve participants reported either that there is a requirement for better understanding weather dynamics, that it is difficult to understand weather dynamics, or that they arrived at a means by which to get predictive information from the GWIS display. One participant suggested that using the GPS free mode enabled him to better estimate weather dynamics by hypothesizing what the weather would look like relative to the aircraft and seeing if, over time, the image matched his expectations. Eight of the twelve participants requested some display feature that shows predicted or trended storm dynamics information.

**Decision-Making** - On the usability survey, and during debriefing sessions, participants were asked about how they would use the GWIS. In addition, we were particularly interested in how they thought use of this technology might affect their determination of whether to and how to avoid weather hazards.

Participants' ratings and comments reflected that they thought the GWIS improved their SA of weather hazard existence and location. When asked how they would use this system, it was clear that some participants felt more comfortable flying in circumstances with the GWIS that they would not fly in without it. Participants mentioned that the information provided by this display would make them more comfortable flying at lower altitudes and in the clouds, where visibility is reduced. One participant stated that, while he knew he was supposed to stay 10nm away from (convective) weather, that with the displayed information, he would get 5nm away. Other participants were more explicit. One stated that he would fly solid IMC and embedded thunderstorms with the display, but not without it. Another stated that he would "feel okay threading the needle with it.” Another, when asked if it could be used for tactical navigation, replied that he thought it could, if he had an opportunity to get used to it. Three participants noted that the time lag in the information makes it unreliable for precisely identifying weather location and intensity, and therefore is not suitable for tactical avoidance of weather. Participants generally felt that 20nm is an appropriate distance to stay away from convective weather in VMC. While one participant considered this distance also safe in IMC, most others preferred to stay 40-60nm away from convective activity. Participants noted that when conditions are VMC with no forecasted enroute weather and the flight is less than 25 miles, or they could visually see and avoid weather, they would not use the GWIS.

Using the GWIS also affects how these pilots thought they would make use of the other weather information available to them. Recall that preliminary results showed that when using this GWIS, pilots desired fewer additional sources of weather information [30]. Results analyzing participants' abilities to objectively estimate weather hazard distance and bearing indicated that the three weather information sources (visual scene, GWIS, and aural information from HIWAS, ATC, and FSS) contribute complimentary information [31]. One participant explicitly described how he would use a number of weather information sources together; stating that he'd use Flight Watch to get storm cell trend information, and onboard weather radar and Strikefinder™ to get more real-time information, and use these to corroborate and supplement GWIS information. Debriefing comments suggest that some participants actively compared information sources to develop higher confidence in their weather SA. Participants who were not allowed access to the GWIS during the experimental phase of the flight, but who had used it on a prior flight, were allowed to see the display during the inbound phase. The coverings on windows that prevented the person with the GWIS from seeing outside during the outbound phase were removed, allowing him a view out a side window. These periods, then, allowed participants to compare visual scene information with what was displayed on the GWIS. Participant comments indicated that they trusted the GWIS more to judge distance of cells, but trusted the visual scene more to judge intensity and cell boundaries.

The preliminary results indicated that participants who had the GWIS and aural weather information were less likely to desire additional information sources using their radio, or additional information from existing or potential onboard equipment or visual cues. A simulation experiment using this same prototype GWIS also found that users of GWISs were less likely to contact ground-based weather professionals [21, 22]. One advantage of this technology is that it allows the pilot to receive weather information without being subjected to the radio congestion problem that can reduce the usefulness of ground-based weather professionals [29]. However, we can also infer that pilots in aircraft equipped with GWISs are less likely to engage in collaborative decision-making.
about the weather with these professionals; and that they might therefore fail to acquire the broader weather picture, and interpretive expertise available from these persons. Responses to the question that asked what additional information participants would desire showed constant individual predilections towards one radio-accessed weather information source that were unaffected by the experimental condition they received. Two participants always wanted to contact Flight Watch. One participant always wanted to contact ATC approach. One participant would always use ATIS, and another would always use AWOS. While it might be surprising that more of these GA participants did not, as a matter of course, contact Flight Watch, the utility of this service is limited by radio congestion as well as availability when flying at low altitudes. Further, contacting Flight Watch, particularly in challenging weather conditions, is workload intensive, requires one to leave ATC frequency, and can be very time-consuming. Debriefing comments emphasized the advantages of radio-accessed weather information, to include ability to ask questions and ensure that the information you receive is relevant to you, and the ease of acquiring aural information vs. information requiring visual redirection when in flight. Participants had varying experience talking to ATC, and one suggested that ATC would not generally be able to provide useful information because they de-clutter their displays so to better serve their primary goal of separating traffic without distraction. Three participants shared the same constant predilections for instrument or visual weather information, always desiring out the window front and side views as well as onboard weather radar. One of these participants also always wanted to have Strikefinder™ (lightning information).

Workload - Only one participant commented that interaction with the GWIS was too loading to use it comfortably. This participant stated that he would have benefited from additional training on the system, but that he also felt there was too much information provided by the system. Preliminary results from a prior analysis indicated no strongly significant differences among any of the NASA-TLX workload scales (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration Level) or an Overall Workload scale. Only (perceived) performance was marginally different among the experimental conditions, where means suggested that perceived performance was noticeably disadvantaged during current IFR flight conditions; that is, without augmenting aural information with either a window view or a GWIS. It is important to recall that while all experimental conditions as having high validity, this view or a GWIS. It is important to recall that while all

DISCUSSION

During debriefing discussions and on open-ended usability questionnaire items, participants described characteristics and features that they desire in an aviation weather information system. These characteristics and features are classified below according to how they would support situation awareness and flight deck decision-making.

SITUATION AWARENESS RECOMMENDATIONS - Participants explicitly indicated additional information and presentation methods that they felt would improve GWIS utility and usability. Table 1 summarizes participants’ recommendations to improve situation awareness according to the following categories: location and intensity of weather (SA Level 1), proximity and hazard level of the weather, and reliability of this information (SA Level 2), and projection of where the weather will be in the future, its intensity level and relevance to the mission at that time (SA Level 3). It is significant to note that ten of the participants made at least one comment indicating the necessity of indicating the relevance of the weather with respect to the aircraft’s position and/or actual flightpath and/or intended flight path. Six of these participants recommended a route overlay similar to that on GPS displays. Three of these also desired a representation of the flightplan. Two other suggestions included providing waypoint icons that could be connected to form a route and an icon for the destination with an indication of “distance from field.” Eight of the twelve participants indicated that predictive/ trending information is insufficient in the system they used. The most frequently mentioned remedy for this insufficiency was an animated loop of NEXRAD imagery. The Appendix holds a table summarizing participant recommendations arranged by the SA level they would support.

Other recommendations stem not directly from explicit participants’ suggestions, but were inferred based on their rating scale data and observations of GWIS use. Recommendations based on these results and observations are listed below:

- Legible text requires careful selection of font size and type, and screen resolution. Off-cardinal orientations of text may be particularly difficult to read.
- Distances must be indicated in a more direct manner than providing a reference unit of measurement (e.g., range rings, point-to-point calculation, gridded background).
- Posting the product creation time/date is insufficient for communicating the age of weather information. Minimally, the elapsed time since the product was created should be used to communicate the age of the presented information.
• Users’ attention must be drawn to important changes in the weather information provided. Alerts to these changes should include increases or decreases in hazard level that affect the flightpath of the aircraft, and the level of reliability attributed to the weather information provided (e.g., whether the information is too old to be considered useful.)

• Participants disliked the implementation of METARS in this prototype GWIS because it obscured their primary weather information interest, the NEXRAD picture, and because it required focused attention to extract the information from it.

• Contextual references (geographical and aviation) must be meaningful to pilots and provide common references to support integration among flight information sources, both in document form (e.g., Low Altitude charts, Approach Plates) and communication among NAS operators (e.g., geographical markers used by FSS to provide weather briefs).

• Participants preferred the GPS Free mode, perhaps because, absent any predictive information in this system, they were able to see more of the weather that was ahead of them on their route.

• Participants require trend information for location, speed, and intensity of weather system changes. We see participants attempting to derive this trend information by slewing the GPS aircraft position to see more of the route in front, using surface observation symbols as a map of an area, and developing conjectures of weather movement for the purpose of checking them with the next picture, and thereby assessing their own internal model of the weather's dynamics. More direct support for estimating these dynamics is required.

• Workload complaints associated with the system are associated with perceptual integration (determining distances, determining geographical referents, switching between NEXRAD and METAR information), and manual control (getting to the METAR text for a station). Particularly in single pilot operations, pilots and manual control (getting to the METAR text for a station). Participants also desired smaller map scales on the GWIS such that it could help them with ground operations. It is interesting to note that participants did not suggest more computationally intensive decision aiding features with respect to using weather information to support flight deck decisions or collaboration with other NAS users. These more elaborate decision aiding features may not have been mentioned simply because participants have not considered the possibility of extending this prototype GWIS to include this form of support. However, further investigation is required to determine if improving SA directly will effectively improve the safety of aviation operations in difficult weather conditions without more computationally-intensive, decision-specific aiding and response selection features.

IMPROVED AVIATION DECISION MAKING - It is important to emphasize that the goal is to not simply improve weather situation awareness as an end, but to encourage safer aviation operations by improving situation awareness and decision-making. While poor situation awareness is often the most constraining aspect of performance in highly dynamic environments with skilled operators [39], good weather situation awareness does not guarantee appropriate performance. One must support the decisions pilots make with aviation weather information, and other information as needed. This need to focus on how pilots use weather information in making decisions is reflected in some comments and questionnaire responses. Participants desired additional information in the GWIS that is not weather related but must be used in conjunction with weather information to make flight decisions. Some of the information desired by participants includes: airport facility handbook information, approach plates, and airport and runway information. One of AWIN's industry partners is actively considering integration of their GWIS into an electronic flightbag which also contains airport facility handbook information and approach plates [40]. Participants also desired smaller map scales on the GWIS such that it could help them with ground operations. It is interesting to note that participants did not suggest more computationally intensive decision aiding features with respect to using weather information to support flight deck decisions or collaboration with other NAS users. These more elaborate decision aiding features may not have been mentioned simply because participants have not considered the possibility of extending this prototype GWIS to include this form of support. However, further investigation is required to determine if improving SA directly will effectively improve the safety of aviation operations in difficult weather conditions without more computationally-intensive, decision-specific aiding and response selection features.

DESIGN GUIDANCE & TRAINING - It is important to consider that these results were obtained from participants in a fairly constrained GA scenario that did not involve using the GWIS while flying. Further, although we endeavored to constrain the character of flight scenarios, by conducting this study as a flight experiment, participants did evaluate the display under somewhat different flight conditions. Therefore, the recommendations provided by participants and derived from their performance and rating data should serve as a list of considerations for further design evaluation. In addition, as shown above, the means by which these recommendations are hypothesized to improve performance are a matter for empirical investigation.

Designing the GWIS proper is only one aspect of ensuring improved safety with this system. We must also be concerned with how quickly and pervasively this technology is adopted, and that training needs are identified to ensure appropriate use of this technology. Kauffmann and his colleagues [9, 41] discuss technology adoption of GWISs by various aviation user groups. Comments from the participants in this study indicate that many of them consider weather information to be an overlay function for a GPS system: “If weather came with GPS features, I’d buy it. Otherwise I’d buy GPS first.” “I’ll first buy a GPS, then I’ll overlay weather, then a weather loop; then “wayout error” course deviation warning; then METARS, (etc.).” A second issue, with
implications for both technology design and training, concerns how GWISs (as well as portable GPS systems) are appropriately used. The FAA Aeronautical Information Manual [42] states that FISDL products, such as ground-based radar precipitation maps, are not appropriate for use in tactical severe weather avoidance (that rather) FISDL supports strategic weather decision-making. However we see clearly from the commentary in this study that not only are participants willing to use these tools in this manner, but they say they would be less likely to acquire weather information from other sources when they do so. Further study must determine whether design formatting can influence how and when pilots use GWISs such that their decisions benefit from efficient use of good weather information, and aren't influenced by misleading information they may present.

Understanding the limitations of GWISs is a training issue. Most participants indicated that a short, one-hour standardized training session, remedial instruction, and about 10 minutes hands-on exploration was sufficient to use the system. However it is clear that, even though participants were shown how to find the age of weather information, and explained the delayed nature of the NEXRAD weather, many did not acquire and/or use this information. It is a matter for further exploration to determine if training on aspects of system reliability might also help counteract reliance on visually compelling, but not necessarily timely, weather information. Training for aviation weather decision-making should not only teach pilots how to effectively interpret visual cues and the limitations of what they can see, but must also instruct them on how to effectively interpret displayed graphical and symbolic weather information, and GWIS limitations.

CONCLUSIONS

This and prior simulation and flight experiment research find that pilots enthusiastically await the opportunity to use a GWIS in flight. Prior research and preliminary results from the CoWS study indicate that GWISs have improved aviation weather situation awareness and decision-making. This usability evaluation shows that pilots particularly appreciate contextual references, the aircraft symbol, and NEXRAD graphical data, and would like to see more integration of flightpath and weather data. However, from this and other GWIS research we also start to understand the content, formatting, and interaction issues that may lead to improper use of a GWIS and may inadvertently result in reduced aviation safety, rather than improved aviation safety.

Future work, therefore, must focus on empirically evaluating new GWIS designs that incorporate recommended user interface improvements for supporting situation awareness and aviation decision-making. GWIS designers must attend to the fact that it seems undeniable that pilots will use these systems to support tactical decisions, not only the strategic decisions for which they are intended. GWISs may require alerts that indicate when relevant hazards exist and when information is less than typically reliable. In particular, pilots in this and other studies fail to notice and/or appreciate the age of weather information when using this information. Additional research should be conducted to determine the cost/benefit trade-offs associated with the age and completeness of weather depictions. As a result of the CoWS and RTI studies, the Minimum Aviation System Performance Standards (MASPS) for FISDL now includes a note that these systems should indicate age or currency of the weather data, in contrast to simply providing the product creation time [43], and the next generation of the GWIS used in these studies provides better indications of product age along these lines. Further research is required to assess the efficacy of providing this information and the presentation methods for conveying it.

While participants in this study did not express recommendations for more computationally intensive aiding functions, these certainly should be considered for their benefit in augmenting pilot's ability to acquire, filter, and assimilate weather and other flight information, and to select and evaluate action plans in response to weather hazards. Along these lines, it is imperative that GWISs communicate, through their interfaces, and that pilots understand, perhaps also through training, the limitations of these systems and how other weather information sources might best be used in conjunction with these systems to best improve weather flying safety.

ACKNOWLEDGMENTS

We gratefully acknowledge the many people at NASA Langley Research Center and at Honeywell who helped with this experiment. We especially thank our project pilot, Charles Cope, and our Crew Chief, Leo McHenry. We also thank Dave Mcluer, Lee Joyce, Ed Radwanski, and Bob Kendall for support of equipment, and Barry Golembiewski for his excellent weather forecasting. We thank Regina Johns for procuring and scheduling participants, Tracy Hunter for short-notice boarding authorities, and Barbara Trippe for aircraft and pilot scheduling. Finally, we are indebted to Jim Joyce at Honeywell and Tim Sobolewski at CLH for ongoing support of the AWIN system, and to Brian Haynes of United Airlines for his vision and enthusiasm in bringing this AWIN system from concept to reality.
REFERENCES


CONTACT

Kara A. Latorella
Aviation Weather Information (AWIN)
Aviation Safety Program - Weather Accident Prevention Project
NASA Langley Research Center
Mail Stop 152
Hampton, VA 23681-2199
k.a.Latorella@larc.nasa.gov
Voice: 757-864-2030
Fax: 757-864-7793

James P. Chamberlain
Aviation Weather Information (AWIN)
Aviation Safety Program - Weather Accident Prevention Project
NASA Langley Research Center
Mail Stop 152
Hampton, VA 23681-2199
j.p.chamberlain@larc.nasa.gov
Voice: 757-864-2147
Fax: 757-864-8588
DEFINITIONS, ACRONYMS, ABBREVIATIONS

AIRMET  Meteorological Advisory
AOPA   Aircraft Owners and Pilots Association
ASOS  Automated Surface Observing System
ATC  Air Traffic Control
ATIS Automatic Terminal Information Service
AvSP  (NASA) Aviation Safety Program
AWOS Automated Weather Observing System
AWIN  (NASA) Aviation Weather Information AvSP element
CoWS Convective Weather Sources study
DLORT (FAA) Data Link Operational Requirements Team
DUATS Direct User Access Terminal System
EDT Eastern Daylight Time
EFAS  Enroute Flight Advisory Service (also known as Flight Watch)
FAA  Federal Aviation Administration
FISDL Flight Information Services Data Link
FSS Flight Service Station
GA  General Aviation
GPS  Global Positioning System
GWIS Graphical Weather Information System
HIWAS Hazardous Inflight Weather Advisory Service
ICAO International Civil Aviation Organization
IFR  Instrument Flight Rules
IMC  Instrument Meteorological Conditions
MIT Massachusetts Institute of Technology
METAR Aviation Routine Weather Report
NAS  National Airspace System
NASA National Aeronautics & Space Administration
NASA-TLX NASA Task Load Index
NAVAID Navigational Aid
NDB Non-Directional Beacon (a NAVAID)
NEXRAD Next Generation RADAR
nm Nautical Miles
PC  Personal Computer
PIC  Pilot in Command
PIREP Pilot Report
QUIS Questionnaire for User Interaction Satisfaction v5.5, Center for Automation Research, University of MD
RADAR Radio Detection and Ranging
SA  Situation Awareness
SIGMET Significant Meteorological Advisory
SPECI Special METAR weather observation
TACAN Tactical Air Navigation (a NAVAID)
TAF Terminal Aerodrome Forecast
VDL  VHF Data Link
VFR  Visual Flight Rules
VHF Very High Frequency radio
VIP Video Integrator Processor (precipitation intensity levels 1-6)
VMC Visual Meteorological Conditions
VOR  VHF Omnidirectional Range (a NAVAID)
VORTAC Combined VOR/TACAN (a NAVAID)
## APPENDIX

### Appendix 1. Participant Recommendations for Improved Situation Awareness.

<table>
<thead>
<tr>
<th>SA Level</th>
<th>Category</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Cloud tops (e.g., highest echo altitude product)</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Pilot reports</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Where is it VFR/IFR</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Turbulence</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Thunderstorms</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Windshear</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Icing</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Complete NAVAID/Airport Database</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Vertical perspective</td>
</tr>
<tr>
<td>1</td>
<td>Additional Information</td>
<td>Higher resolution NEXRAD data</td>
</tr>
<tr>
<td>1</td>
<td>Presentation</td>
<td>Dedicated METAR row at bottom of display</td>
</tr>
<tr>
<td>1</td>
<td>Presentation</td>
<td>Aural presentation of METAR information</td>
</tr>
<tr>
<td>1</td>
<td>Presentation</td>
<td>Display old data when it is more complete than newer data.</td>
</tr>
<tr>
<td>1</td>
<td>Display Views</td>
<td>Larger display</td>
</tr>
<tr>
<td>1</td>
<td>Display Views</td>
<td>Higher display resolution</td>
</tr>
<tr>
<td>2</td>
<td>Display Views</td>
<td>Autozoom and autoscroll function based on mission</td>
</tr>
<tr>
<td>2</td>
<td>Presentation/Interaction</td>
<td>Crosshair function to determine bearing &amp; distance between selected points (e.g., present position &amp; weather, destination &amp; weather).</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Range rings and azimuth from present position</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Course line, Flight path</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Destination Icon</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Airways</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Aircraft heading and other GPS functions</td>
</tr>
<tr>
<td>2</td>
<td>Presentation</td>
<td>Moving map &quot;Overlay weather on current GPS technology&quot;</td>
</tr>
<tr>
<td>3</td>
<td>Additional Information</td>
<td>Terminal Area Forecasts</td>
</tr>
<tr>
<td>3</td>
<td>Presentation</td>
<td>Animation of past NEXRAD images to indicate trend</td>
</tr>
<tr>
<td>3</td>
<td>Presentation</td>
<td>Direction and speed indication arrows (e.g., radar summary charts)</td>
</tr>
<tr>
<td>3</td>
<td>Presentation</td>
<td>Prognostications of NEXRAD movement (e.g., trend arrows)</td>
</tr>
</tbody>
</table>