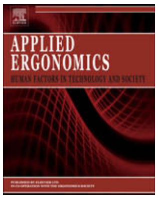
Applied Ergonomics 40 (2009) 597–607

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo



Comparing perceptual judgment and subjective measures of spatial awareness

Matthew L. Bolton, Ellen J. Bass*

Department of Systems and Information Engineering, University of Virginia, P.O. Box 400747, 151 Engineer's Way, Charlottesville, VA 22904, USA

ARTICLE INFO

Article history:
Received 12 June 2007
Accepted 22 April 2008

Keywords:
Spatial awareness
Subjective measures
Perspective judgments

ABSTRACT

Spatial awareness is important in domains where safety hinges on human operators keeping track of the relative locations of objects in the environment. While a variety of subjective and judgment-based measures have been used to evaluate spatial awareness, none have probed all three of its levels: (1) identification of environmental objects, (2) their current locations relative to the operator, and (3) their relative positions over time. This work compares new judgment-based measures of spatial awareness that probe all three levels of spatial awareness to conventional subjective measures. In the evaluation of 14 configurations of Synthetic Vision Systems head down displays (seven terrain textures and two Geometric Fields of View (GFOVs)), 18 pilots made four types of judgments (relative angle, distance, height, and abeam time) regarding the location of terrain points displayed in 112 5-s, non-interactive simulations. They also provided subjective demand, awareness, clutter, SA-SWORD, and preferred GFOV measures. Correlation analyses revealed that displays that received higher awareness and SA-SWORD subjective ratings were associated with smaller errors in abeam time judgments and, for SA-SWORD, smaller errors in relative distance judgments. Thus SA-SWORD provides insight into level 2 spatial awareness and both SA-SWORD and awareness provide insight into level 3 spatial awareness. ANOVA and χ^2 analyses revealed comparable results between display configurations that produced the minimum error in judgments and those recommended by the awareness, SA-SWORD, and preferred GFOV measures.

1. Introduction

Spatial awareness is an aspect of situation awareness (Wickens, 2002a) that encompasses the extent to which a person notices objects in the surrounding environment (level 1), the person's understanding of where these objects are (level 2), and the person's understanding of where these objects will be in the future (level 3) (Endsley, 1995b; Wickens, 2002b). Spatial awareness is important in domains where safety hinges on operators keeping track of the relative location of objects in the environment. Such domains include air traffic control, where spatial awareness helps controllers maintain separation between aircraft; combat operations, where spatial awareness allows soldiers to identify, locate, and predict the location of enemies; driving, where spatial awareness helps drivers avoid contact with other vehicles and environmental obstacles; and aircraft piloting, where spatial awareness helps pilots avoid terrain, other aircraft, and environmental obstacles.

1.1. Subjective measures

Subjective measures have frequently been used to evaluate awareness (Taylor, 1990; Vidulich and Hughes, 1991), workload (Vidulich et al., 1991), clutter (Bailey et al., 2002, 2006), alertness (Dorrian et al., 2007), and many other dimensions of cognition. These types of measures are popular because of the relative ease with which they can be collected. The use of subjective measures typically involves exposing participants to a particular design under varying operational conditions and having them rate their experience (in the case of awareness, how aware they thought they were) using Likert scales.

A common subjective awareness measure is the Situation Awareness Rating Technique (SART) (Hughes and Takallu, 2002; Stark et al., 2001; Takallu et al., 2004). SART was developed by interviewing experienced aircrew and identifying 10 SA constructs: instability of situation, variability of situation, complexity of situation, arousal, spare mental capacity, concentration, division of attention, information quantity, information quality, and familiarity (Selcon and Taylor, 1990; Taylor, 1990). These 10 constructs were found to cluster into three broad categories: attentional demand, attentional supply, and understanding. Attentional demand represents the amount of demand placed on attentional resources. It encompasses the instability, variability, and complexity of the situation. Attentional supply denotes the amount of attentional resources afforded by a design or situation. It includes arousal, spare mental capacity, concentration, and division of attention. Understanding of the situation consists of information quantity, information quality, and familiarity.

---

* Corresponding author. Tel.: +1 434 243 5531.
E-mail address: ejb4n@virginia.edu (E.J. Bass).

0003-6870/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apergo.2008.04.020

Thus, SART is often measured using these three dimensions. Participants rate designs based on attentional demand (*Demand*), attentional supply (*Supply*), and understanding of the situation (*Understanding*) (between "low" and "high") each on 100-point Likert scales (Selcon and Taylor, 1990; Taylor, 1990). The three values are then combined into a single situation awareness score (SA) using the formula SA = Understanding − (Demand − Supply).

Subjective rating scales have also been used to measure spatial awareness (Bailey et al., 2002; Glaab and Hughes, 2003) as well as display clutter (Bailey et al., 2006). Like a SART dimension, these measures use Likert scales to allow participants to rate displays.

Situation Awareness-Subjective WORkload Dominance (SA-SWORD) is another type of subjective measure that has been used to evaluate SA with different displays (Arthur et al., 2004; Hughes and Takallu, 2002; Vidulich and Hughes, 1991). It has also been used to specifically measure spatial awareness (Bailey et al., 2002). SA-SWORD, which was adapted to measure awareness from the Subjective Workload Dominance technique, is an Analytic Hierarchy Process (AHP) based metric that allows participants to make pairwise comparisons between experimental conditions on a 17-point scale concerning the relative amount of SA provided by each (Vidulich and Hughes, 1991). For n displays, SA-SWORD requires that $\binom{n}{2}$ comparisons be made in order to populate an $n \times n$ comparisons matrix. The matrix is then used to calculate scores for each of the displays by normalizing the geometric means for each row of the matrix (Turner, 1996; Vidulich et al., 1991).

Researchers have also used discrete choice to evaluate awareness. For such a metric, participants specify which display conditions they prefer during display operation (through a selectable choice) or through a post run questionnaire. Arthur et al. (2004), Comstock et al. (2001), and Glaab and Hughes (2003) used such methods to evaluate Geometric Fields of View (GFOVs) (the angular boundaries of the volume of space represented in the display) in perspective cockpit displays.

Despite their ease of collection (and thus popularity), it is not clear what these subjective metrics measure. While SART and similar rating techniques have been correlated with performance measures (Selcon and Taylor, 1990), they have also been shown to be correlated with operator confidence (Endsley et al., 1998). Thus, SART-like awareness metrics may be measuring confidence, or some other cognitive attribute, rather than awareness. Additionally, studies utilizing SA-SWORD have found strong correlations between pilot display preference and SA-SWORD scores (Endsley, 1995a). It is possible that preference influences SA-SWORD rankings or vice versa. Similar concerns can also be levied at discrete choice subjective measures. Further, Pew (2000) has expressed concern that awareness assessment found using SART and SA-SWORD may be confounded by workload.

1.2. Judgment-based measures

Judgment-based metrics provide another approach to measuring awareness. A widely utilized SA measurement technique that uses judgments to assess awareness is the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988). In SAGAT, participants answer a battery of questions probing different categories of operator SA knowledge during pauses or breaks in experimental activity. Participants are scored based on the percentage of the questions they answer correctly.

Judgment-based metrics have also been used to measure spatial awareness specifically. Yeh (1992) used ordinal distance judgments to assess spatial perception for stereoscopic and perspective displays. Several studies have utilized azimuth and elevation angle judgments of the relative position of two objects over synthetic terrain (Adelstein and Ellis, 2000; Barfield and Rosenberg, 1995; McGreevy and Ellis, 1986). A variation on this procedure measured the error in azimuth and elevation angle in gaze during simulated flight when pilots were asked to stare in the direction of memorized targets during pauses in simulated flight (Dorighi et al., 1992). Wells et al. (1988) used the ability of participants to accurately replicate, from memory, the position of objects presented in helmet-mounted displays as a metric of spatial awareness. Fracker (1990) had participants replicate the display position of enemy aircraft during pauses in simulated flight. In studies by Marshak et al. (1987), participants made judgments about the location of targets shown on map displays during pauses in simulated flight. While each of these techniques measured different aspects of spatial awareness, none measured all three of the levels specified by Wickens (2002b).

1.3. Objective

Quantifiable measures of spatial awareness are necessary to inform the development of decision support tools in many domains. This research compares subjective awareness measures to new judgment-based measures that directly probe all three levels of spatial awareness: identification of an object in the environment probed level 1 spatial awareness; estimates of the object's relative position probed level 2 spatial awareness; and a prediction of the time it would take to reach that object probed level 3 spatial awareness. Both directional and absolute error terms for each judgment assess spatial awareness accuracy at each level. Since these measures are theoretically grounded (based on the spatial awareness literature), their comparison to subjective measures could reveal what aspects of spatial awareness the subjective measures assess.

Synthetic Vision Systems (SVS) were used in this study because subjective measures have frequently been used to evaluate SVS designs with respect to both spatial awareness and general situation awareness. Controlled Flight Into Terrain (CFIT), where a fully functional aircraft is inadvertently flown into the ground, water, or other terrain obstacle, has been the cause of more than 22% of all fatal accidents in worldwide commercial aviation since 1987 (Boeing, 2007). Such accidents are often characterized by a loss of situation awareness in low-level flight and low visibility conditions (Khatwa and Roelen, 1998). SVS are technologies that address this problem by using onboard terrain databases and Global Positioning System (GPS) data to create a computerized picture of the world in front of the aircraft, regardless of the actual visibility conditions.

Two aspects of SVS that can affect spatial awareness are terrain texture (the imagery drawn on the synthetic terrain) and Geometric Field of View (GFOV; the angular boundaries of the volume of space represented in the display). An experiment was conducted to investigate the new measures of spatial awareness. Bolton et al. (2007) report the results (and corresponding insights) of using these measures to evaluate which terrain textures and GFOVs best facilitate spatial awareness for SVS. Bolton and Bass (2008) use these measures to identify spatial awareness biases in SVS. The work reported in this paper compares the results obtained with these new measures to traditional subjective measures. Thus, this works attempts to determine to what degree the subjective measures reflect the objective performance encompassed by the judgment-based measures.

The new spatial awareness measures were evaluated with respect to a terrain point indicated on the terrain of a SVS head down display. Identifying the terrain point probed level 1 spatial awareness; judgments of the relative azimuth angle, distance, and height of the terrain point to ownship probed level 2 spatial awareness (providing a three-dimensional perspective of the pilot's perception of the terrain's location); and an abeam time judgment (the time it would take the pilot to fly to the point of closest approach to the terrain point) probed level 3 spatial awareness. For this analysis, judgment performance (as measured by absolute and directional errors) was compared with demand, awareness, clutter, SA-SWORD, and discrete choice (preferred GFOV) subjective measures. Since no previous studies have evaluated spatial awareness using these new judgment-based measures, it was not clear if the subjective measures would exhibit correlation with the judgment-based measures.

2. Methods

2.1. Participants

Eighteen general aviation pilots volunteered for the study. All participants had less than 400 h of flight experience ($M = 157$, $S = 75$). They were familiar with the out-the-window view from a cockpit but not with SVS displays. They were paid $100 for their participation.

2.2. Apparatus

Experiments were run in a windowless constantly lighted laboratory. Workstations displayed each simulation and collected participant judgments. Simulations depicted SVS head down displays with the symbology shown in Fig. 1. The location of the terrain point was indicated using a yellow inverted cone ($d = 500$ ft, $h = 500$ ft) rendered as part of the SVS environment. The tip of the cone intersected the terrain at the terrain point. All simulations depicted SVS displays in straight, level flight (no pitch or roll) at 127 knots with no additional influences on motion. They were displayed as 5 s, 836 × 728 pixel, 30 frames per second, Windows Media Video (WMV) files. Custom software played the WMV files and collected participant responses (Bolton et al., 2006).

2.3. Independent variables

There were five within subject independent variables. These included texture, GFOV, and three scenario geometry variables: the relative (azimuth) angle, distance, and height of the terrain point to ownship.

Seven textures (Fig. 2) were used in the experiment: three base textures (F, E, and P), and four derivatives of them (EF, PF, PE, PEF). The three base textures were chosen because each had been used in SVS (Glaab and Hughes, 2003; Schnell and Lemos, 2002), and each facilitated different depth cues that persisted under combination (see Bolton et al., 2007).

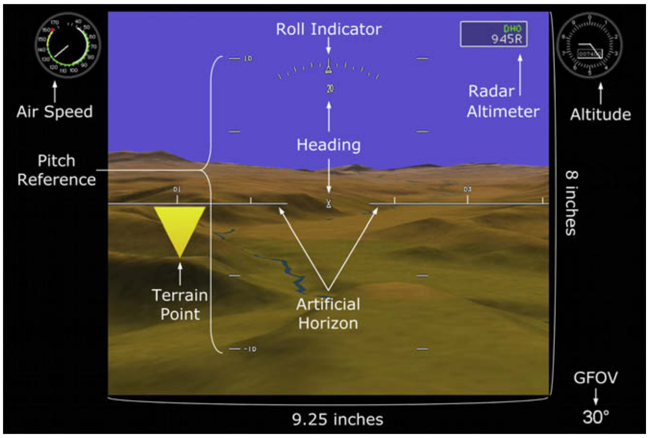


Fig. 1. The SVS display and symbology used in the experiment (labels added). SVS displays were presented to participants with an eye distance of approximately 30 inches and a horizontal visual angle of approximately 18°. This figure is reproduced with permission from *Human Factors*. Copyright 2007 by the Human Factors and Ergonomics Society. All rights reserved.

Two GFOVs (30° and 60°) were used in this experiment. These were selected because they were shown to have pilot preference by Arthur et al. (2004), Comstock et al. (2001), Glaab and Hughes (2003).

The location of the terrain point varied based on its relative position to ownship at the end (last frame) of a simulation by changing the three scenario geometry parameters: the relative angle, distance, and height of the terrain point with respect to ownship. Each of the variables had two levels (Table 1).

2.4. Dependent measures

Eight dependent measures were calculated from the four judgment values (relative angle (°), relative distance (nmi), relative height (ft), and abeam time (s)) from the three judgment tasks (Table 2). There were two dependent measures associated with each of these judgment values: one for directional error and one for absolute error. Each directional error term represented both the direction and magnitude of the error in the judgment value. When a participant overestimated a judgment, the corresponding directional error term was positive. When the participant underestimated a judgment, it was negative. Absolute error terms represented the magnitude of the error judgment and were calculated as the absolute value of their corresponding directional error term. Only absolute error terms are discussed in this manuscript.

There were six types of subjective measures in this study. For each texture and GFOV combination, demand, awareness, and clutter ratings were collected on a 100-point scale. Demand and awareness represented the attentional demand and situational understanding dimensions of SART, respectively, with the awareness scale also being comparable to the spatial awareness measure used by Bailey et al. (2002) and Glaab and Hughes (2003). The clutter dependent variable was comparable to that used by Bailey et al. (2006) and required participants to rate the amount of display clutter they observed. To avoid potential confounds with workload, the experimental task did not require aircraft piloting and therefore the supply dimension was not collected.

Because collecting the $\binom{14}{2} = 91$ pair-wise comparisons associated with a SA-SWORD probe for each level of the texture × GFOV interaction was considered to be too difficult for participants, SA-SWORD pair-wise comparisons were collected for each texture within a given GFOV. Thus, there were two SA-SWORD dependent measures, SA-SWORD 30° and SA-SWORD 60°, based on $\binom{7}{2} = 21$ comparisons.

The sixth type of subjective measure was called preferred GFOV. Participants specified which GFOV they thought provided better spatial awareness for each texture (30°, 60°, or neither).

2.5. Procedure

Each experimental session lasted between 3 and 4 h. The participants completed consent forms and were briefed about the experiment. For each trial, participants viewed 5-s simulations of an SVS heads down display in flight (Fig. 1) in which participants could observe global optic flow (Gibson, 1986). At the end of the 5 s, the simulation paused for 1 s, and the screen was cleared. Each simulation (representing a unique combination of within subject variable levels) depicted a unique terrain configuration.

For each trial, participants made four judgments based on the relative position of the terrain point: relative angle, relative distance, relative height, and abeam time using the interface in Fig. 3. For the relative distance and angle judgments, participants placed a yellow × in the upper left section of the display corresponding to

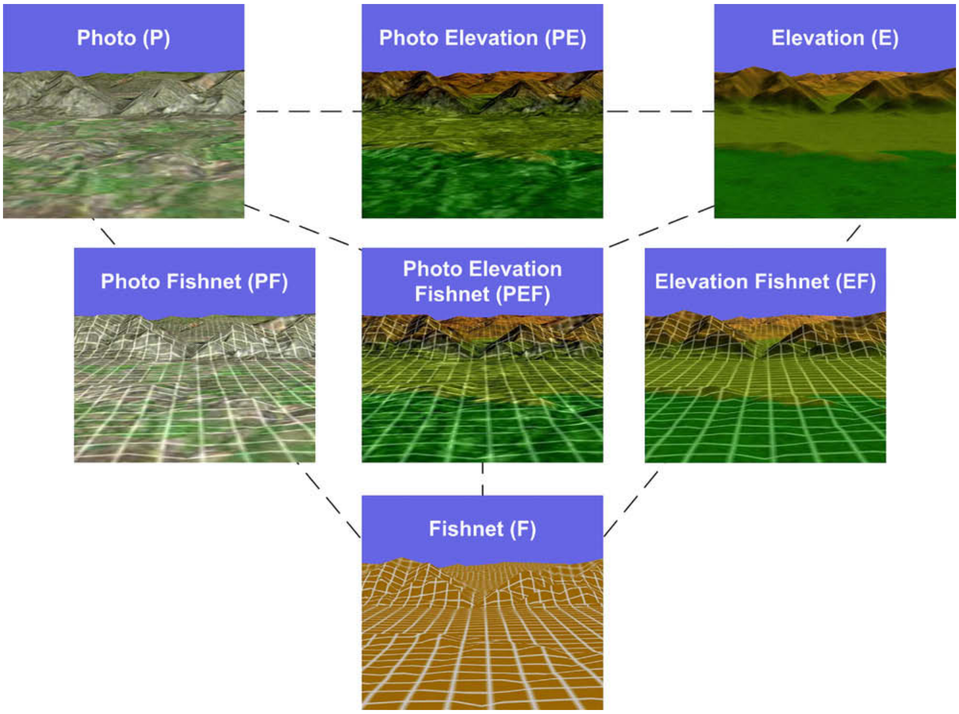


Fig. 2. The terrain textures evaluated in the experiment. This figure is reproduced with permission from *Human Factors*. Copyright 2007 by the Human Factors and Ergonomics Society. All rights reserved.

the lateral location of the terrain point relative to the aircraft. Values for relative angle (°) and distance (nmi) were displayed next to the ×. For the relative height judgment, the participant placed a yellow × on a vertical scale in the upper right of the display corresponding to the relative height of the terrain point. The relative height was displayed in feet next to the × as it was moved. For the abeam time judgment, participants entered the time judgments in minutes and seconds using the keyboard. To support this time judgment, a yellow dot on the relative distance and angle judgment collection interface indicated the location of the abeam point based on the relative distance and angle judgment. Participants were asked to perform these tasks as quickly and accurately as possible. For training trials, participants were given feedback relating to the accuracy of their judgments (see Bolton et al., 2006 for more experimental apparatus details).

Each participant experienced 112 counterbalanced experimental trials (7 textures × 2 GFOVs × 2 angles × 2 distances × 2 heights = 112) and 72 training trials. Participants saw all of the trials with one GFOV before seeing any trials with the other. GFOV presentation order was counterbalanced between participants. Textures used to derive other textures always appeared before their derivatives in order to avoid complications associated with presenting a derivative texture before participants had seen its bases. Each participant saw two of the base textures, the combination of them, the third texture, and the rest of the combinations. Three texture orders were created so that no base texture was introduced in more than one ordered slot: {P, E, PE, F, PF, EF, PEF}, {E, F, EF, P, PE, PF, PEF}, and {F, P, PF, E, EF, PE, PEF}. Texture orders were counterbalanced between participants.

For the first texture seen for the first GFOV, there were 12 training trials. For the other six textures, there were four training trials per texture (4 × 6 = 24). This pattern was repeated for the second GFOV. Thus, there were 12 + (4 × 6) = 36 training trials for each GFOV for a total of 2 × 36 = 72 training trials. Participants received judgment accuracy feedback after each training trial (see Fig. 3).

The order in which the eight scenario geometry levels were presented was unique for each texture and GFOV combination. Thus there were 14 scenario geometry presentation orders. Scenario geometry variable levels were counterbalanced between presentation orders so that each combination of variable levels appeared in each ordered slot twice and directly followed every other combination twice.

Immediately following the completion of all the trials for each texture and GFOV combination, subjective demand (Taylor, 1990),

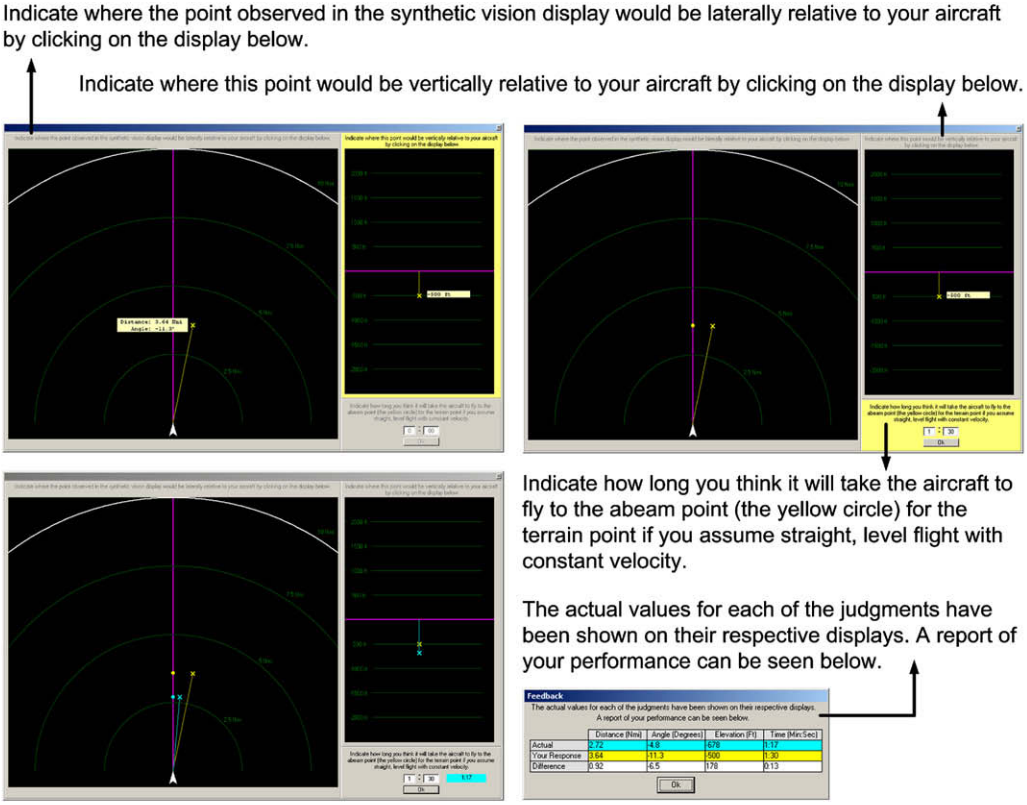
Table 1
Terrain point relative position to ownship (scenario geometry) level encoding.

Independent variable	Range	Distribution	Level
Angle	$[0°, 6.5°]$	$N(\mu = 3.75, \sigma = 1.25)$	Small
	$[8.5°, 15°]$	$N(\mu = 11.25, \sigma = 1.25)$	Large
Distance	$[1, 3.25$ nmi$]$	$N(\mu = 2.25, \sigma = 0.417)$	Near
	$[3.75, 6$ nmi$]$	$N(\mu = 4.75, \sigma = 0.417)$	Far
Height	$[-1000, -100$ ft$]$	$U(-1000, -100)$	Below
	$[100, 1000$ ft$]$	$U(100, 1000)$	Above

Table 2
Dependent measure formulations.

Terrain point position measure	Actual value	Judgment value	Directional error dependent measure	Absolute error dependent measure
Azimuth angle	A_a	A_j	$A_e = \begin{cases} A_j - A_a & \text{if } A_a > 0 \\ -A_j + A_a & \text{otherwise} \end{cases}$	$\lvert A_e \rvert$
Distance	D_a	D_j	$D_e = D_j - D_a$	$\lvert D_e \rvert$
Height	H_a	H_j	$H_e = \begin{cases} H_j - H_a & \text{if } H_a > 0 \\ -H_j + H_a & \text{otherwise} \end{cases}$	$\lvert H_e \rvert$
Abeam time	τ_a	τ_j	$\tau_e = \tau_j - \tau_a$	$\lvert \tau_e \rvert$

Note: all terrain point position measures were made relative to ownship. A_a and A_j were measured relative to the aircraft's vector of displacement with angles in the clockwise direction being positive and angles in the counterclockwise direction being negative. H_a and H_j were measured relative to the aircraft's height with positive heights above the aircraft and negative heights below.



Indicate where the point observed in the synthetic vision display would be laterally relative to your aircraft by clicking on the display below.

Indicate where this point would be vertically relative to your aircraft by clicking on the display below.

Indicate how long you think it will take the aircraft to fly to the abeam point (the yellow circle) for the terrain point if you assume straight, level flight with constant velocity.

The actual values for each of the judgments have been shown on their respective displays. A report of your performance can be seen below.

Feedback

The actual values for each of the judgments have been shown on their respective displays.
A report of your performance can be seen below.

	Distance (Nmi)	Angle (Degrees)	Elevation (Ft)	Time (Min:Sec)
Actual	2.72	-4.8	-678	1:17
Your Response	3.64	-11.3	-500	1:30
Difference	0.92	-6.5	178	0:13

Fig. 3. The judgment collection interface in multiple modes of operation (clockwise from the upper left): the relative distance, angle, and height judgments; the abeam time judgment; training feedback on the judgment collection interface; numerical training feedback shown concurrently with feedback on the judgment collection interface. This figure is reproduced with permission from *Human Factors*. Copyright 2007 by the Human Factors and Ergonomics Society. All rights reserved.

awareness (Bailey et al., 2002; Glaab and Hughes, 2003), and clutter (Bailey et al., 2006) ratings were collected (Fig. 4a). After all of the trials for a GFOV were completed, participants made SA-SWORD pair-wise comparisons (Vidulich and Hughes, 1991) between each texture with that GFOV (Fig. 4c). After all of the trials were completed, participants specified which GFOV (preferred GFOV) if any provided superior spatial awareness (Fig. 4b). For the SA-SWORD and preferred GFOV measures, participants were instructed (during the experiment briefing) to comparatively assess how well they were able to determine where the aircraft was relative to the surrounding terrain using the displays (this is comparable to the instructions presented for the awareness measure (Fig. 4a)).

Participants were given a short break at three points in the experiment: the first occurred after the participant completed all of the trials for the third texture for the first GFOV, the second occurred after the participant provided the SA-SWORD pair-wise comparisons for the textures with the first GFOV, and the third occurred after the participant completed all of the trials for the third texture for the second GFOV.

2.6. Experimental design and data analysis

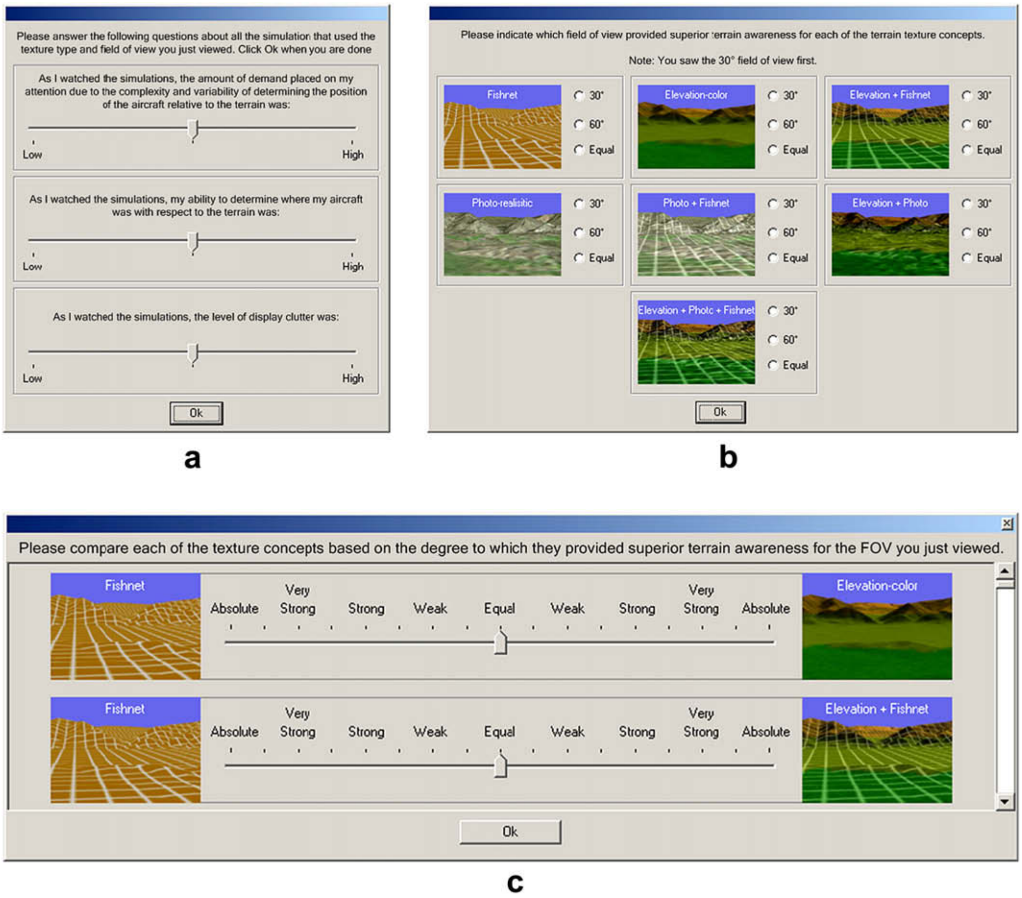
The experiment employed a repeated measures design. Three participants were randomly assigned to each of the six combinations of the GFOV and texture orders (2 GFOV orders × 3 texture orders = 6).

The demand, awareness, clutter, and SA-SWORD dependent measures were subjected to post-processing before being evaluated in the subsequent analyses. Each participant's demand, awareness, and clutter measures were converted into *z*-scores based on the mean and standard deviation of the participant's ratings for the respective measure. SA-SWORD 30° and SA-SWORD 60° pair-wise comparisons were converted to numerical scores using the aforementioned method described by Turner (1996) and Vidulich et al. (1991).

In order to determine what dimension of spatial awareness was being assessed by the subjective measures, a series of correlation analyses between the subjective and absolute error judgment-based measures was conducted. Correlations were also calculated between subjective measures in order to determine how they interrelate. This was done across the mean error for independent variable levels of the texture × GFOV interaction. Correlations were also evaluated across levels of the texture main effect in order to determine to what degree texture was responsible for associations between dependent variables. Correlations could not be calculated across GFOVs because there were only two.

In order to compare the results obtained from the subjective measures to those found for the judgment-based measures, the main and two-way interaction effects of the within and between subject factors on the subjective dependent measures (with the exception of preferred GFOV's nominal data) were assessed using a univariate repeated measures' analyses of variance (ANOVA) with a Type III sum of squares (Brace et al., 2003). A Mauchly's Test of Sphericity was performed in order to ensure that the assumptions for the repeated measure analysis were not violated. When sphericity was violated ($p < 0.05$) a Greenhouse–Geisser ε correction factor was used (Brace et al., 2003). Post-hoc analyses were used to evaluate differences between textures. When sphericity was violated, a Bonferroni post-hoc analysis was used (Stevens, 2002). A Tukey's HSD was used otherwise.

A cross-tabulation χ^2 analysis was used to test for significant differences between the preferred GFOV options selected (both in general and for each texture).



Please answer the following questions about all the simulation that used the texture type and field of view you just viewed. Click Ok when you are done

As I watched the simulations, the amount of demand placed on my attention due to the complexity and variability of determining the position of the aircraft relative to the terrain was:

As I watched the simulations, my ability to determine where my aircraft was with respect to the terrain was:

As I watched the simulations, the level of display clutter was:

Please indicate which field of view provided superior terrain awareness for each of the terrain texture concepts.

Note: You saw the 30° field of view first.

Please compare each of the texture concepts based on the degree to which they provided superior terrain awareness for the FOV you just viewed.

Fig. 4. The subjective measures data collection interfaces: (a) demand, awareness, and clutter; (b) preferred GFOV; (c) SA-SWORD.

3. Results

Correlation results are shown in Tables 3 and 4 with corresponding scatter plots in Figs. 5 and 6. The number of correlations performed would necessitated a correction of $\alpha = 0.05/29 = 0.001724$. Thus, trends are presented with $\alpha = 0.01$. ANOVA results for the subjective measures are presented using $\alpha = 0.05$ for significance (Table 5). ANOVA results for the judgment-based measures appear in Bolton et al. (2007). Differences observed between levels of significant main effects appear in Fig. 7. Results for preferred GFOV appear in Fig. 8.

3.1. *The relationship between the subjective and judgment-based measures*

Correlation analyses results across texture × GFOV levels (Table 3 and Fig. 5) revealed negative correlation trends between SA-SWORD and $|D_e|$ as well as $|\tau_e|$. Awareness also exhibited a negative trend with $|\tau_e|$. A positive trend was observed between SA-SWORD and awareness. Neither $|H_e|$ nor $|A_e|$ were correlated with any of the subjective measures. Similarly, the demand and clutter subjective measures were not correlated with any other measures.

The correlation analysis across texture (Table 4 and Fig. 6) revealed a negative trend between SA-SWORD 30° and $|\tau_e|$. Negative trends were nearly indicated ($p < 0.05$) between SA-SWORD 60° and both $|D_e|$ and $|\tau_e|$ and between SA-SWORD 30° and $|D_e|$, indicating that texture contributed to the associations found in the correlations across texture × GFOV levels. Similarly, neither $|H_e|$ nor $|A_e|$ correlated with any of the subjective measures and neither demand, awareness, nor clutter, showed any correlations with any of the other measures.

3.2. *Effect of texture and GFOV on the subjective measures*

Texture produced a significant main effect for the awareness dependent measure (Table 5). While a Tukey's post-hoc analysis indicated that there were no significant differences between the

Table 3
Correlations between average dependent measure values for each level of texture × GFOV.

	Demand	Awareness	Clutter	SA-SWORD
Awareness	$r(14) = -0.17$ $p = 0.572$			
Clutter	$r(14) = 0.34$ $p = 0.234$	$r(14) = -0.29$ $p = 0.309$		
SA-SWORD	$r(14) = -0.27$ $p = 0.356$	$r(14) = 0.67$ $p = 0.009^*$	$r(14) = 0.03$ $p = 0.923$	
$\lvert D_e \rvert$	$r(14) = -0.26$ $p = 0.379$	$r(14) = -0.46$ $p = 0.098$	$r(14) = -0.29$ $p = 0.322$	$r(14) = -0.73$ $p = 0.003^*$
$\lvert A_e \rvert$	$r(14) = 0.31$ $p = 0.285$	$r(14) = -0.19$ $p = 0.514$	$r(14) = -0.51$ $p = 0.065$	$r(14) = -0.38$ $p = 0.182$
$\lvert H_e \rvert$	$r(14) = 0.09$ $p = 0.766$	$r(14) = -0.37$ $p = 0.195$	$r(14) = 0.29$ $p = 0.318$	$r(14) = -0.53$ $p = 0.051$
$\lvert \tau_e \rvert$	$r(14) = 0.15$ $p = 0.609$	$r(14) = -0.68$ $p = 0.007^*$	$r(14) = -0.10$ $p = 0.740$	$r(14) = -0.79$ $p = 0.001^*$

$^*p < 0.01$.

Table 4
Correlations between average dependent measure values for each texture.

	Demand	Awareness	Clutter	SA-SWORD 30°	SA-SWORD 60°
Awareness	$r(7) = -0.43$ $p = 0.334$				
Clutter	$r(7) = 0.57$ $p = 0.186$	$r(7) = 0.02$ $p = 0.958$			
SA-SWORD 30°	$r(7) = -0.44$ $p = 0.323$	$r(7) = 0.70$ $p = 0.079$	$r(7) = 0.20$ $p = 0.660$		
SA-SWORD 60°	$r(7) = -0.48$ $p = 0.273$	$r(7) = 0.68$ $p = 0.091$	$r(7) = 0.15$ $p = 0.741$		
$\lvert D_e \rvert$	$r(7) = 0.07$ $p = 0.89$	$r(7) = -0.49$ $p = 0.269$	$r(7) = -0.58$ $p = 0.176$	$r(7) = -0.86$ $p = 0.014$	$r(7) = -0.79$ $p = 0.037$
$\lvert A_e \rvert$	$r(7) = 0.35$ $p = 0.437$	$r(7) = -0.64$ $p = 0.122$	$r(7) = -0.33$ $p = 0.473$	$r(7) = -0.64$ $p = 0.123$	$r(7) = -0.72$ $p = 0.069$
$\lvert H_e \rvert$	$r(7) = 0.73$ $p = 0.064$	$r(7) = -0.24$ $p = 0.599$	$r(7) = 0.19$ $p = 0.678$	$r(7) = -0.68$ $p = 0.091$	$r(7) = -0.66$ $p = 0.106$
$\lvert \tau_e \rvert$	$r(7) = 0.32$ $p = 0.478$	$r(7) = -0.75$ $p = 0.053$	$r(7) = -0.39$ $p = 0.383$	$r(7) = -0.92$ $p = 0.003^*$	$r(7) = -0.86$ $p = 0.014$

Note: because SA-SWORD 30° and SA-SWORD 60° were collected separately for each of their respective GFOVs, they were not averaged across GFOV in this analysis.
$^*p < 0.01$.

mean awareness ratings for each texture (Fig. 7b), differences were found using a least significant difference post-hoc. This revealed that E, EF, PF, and PEF were amongst the four textures in the homogeneous subset of textures that received the highest scores (Fig. 7b).

Texture was also a significant main effect for both the SA-SWORD 30° and SA-SWORD 60° dependent measures. Bonferroni post-hoc analyses revealed that there were five textures that were in the homogeneous subsets of textures that received the highest scores for SA-SWORD 30°: P, EF, PF, PE, and PEF (Fig. 7c). For

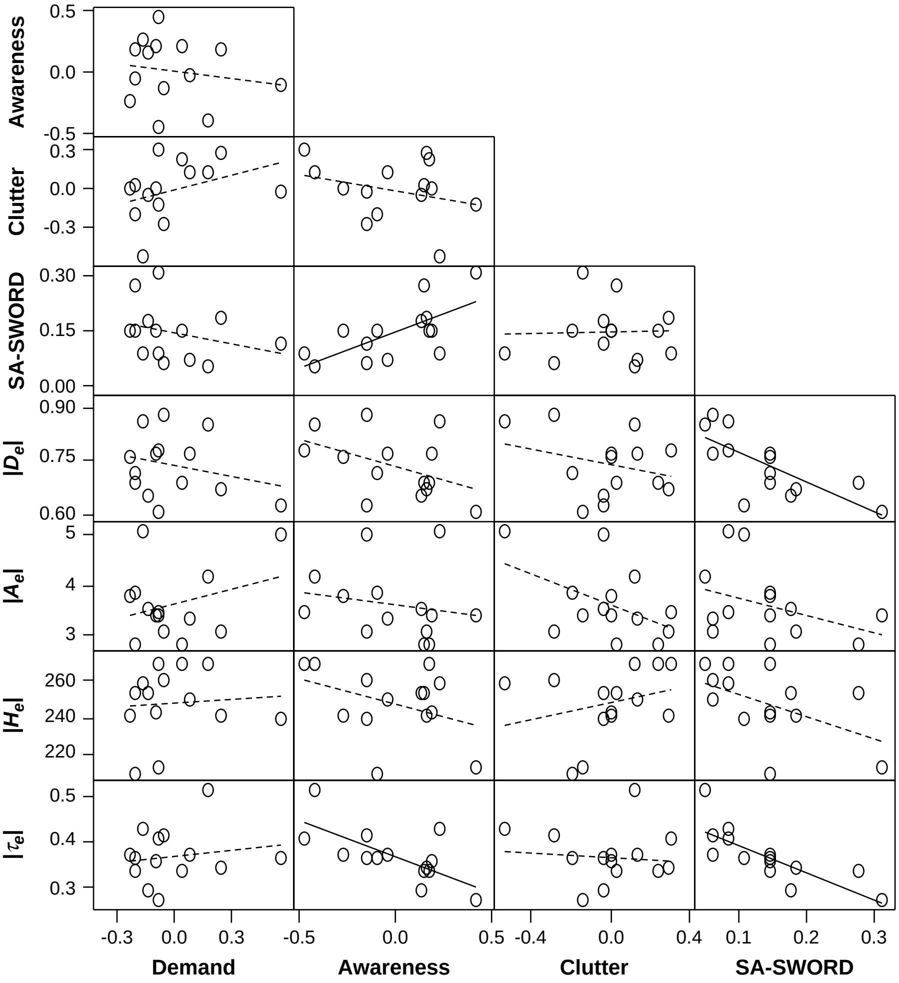


Fig. 5. Scatter plots showing the correlations between average dependent measure values for each level of texture × GFOV. Solid lines indicate a correlation is significant at the $p < 0.01$ level.

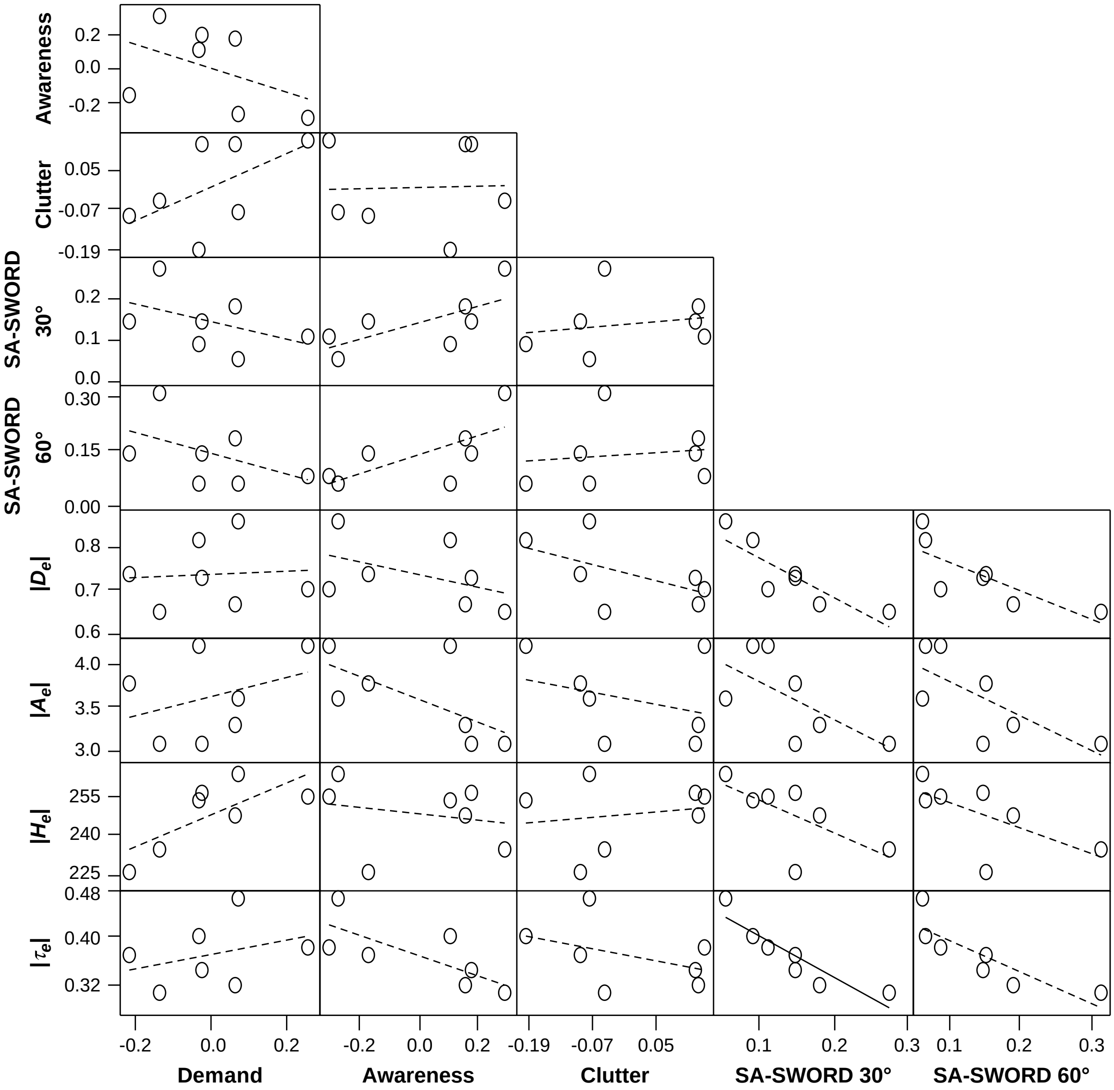


Fig. 6. Scatter plots showing the correlations between average dependent measure values for each texture. Solid lines indicate a correlation is significant at the $p < 0.01$ level.

SA-SWORD 60°, there were four such textures EF, PF, PE, and PEF (Fig. 7d).

GFOV was not a significant main effect for demand, awareness, or clutter. While participants were more likely to prefer the 60° GFOV (Fig. 8a), no significant differences were found between the preferred GFOV option selected in general and the options selected for each texture (Fig. 8b).

The texture × GFOV interaction was not a significant effect for any of the subjective dependent measures.

4. Discussion

By comparing the subjective measures to the theoretically grounded, judgment-based spatial awareness measures, this work provides insight into what dimensions of spatial awareness the subjective measures are evaluating. Given that this evaluation was conducted in the context of an experimental evaluation of design parameters, this work also provides insights into the effectiveness of the subjective measures in distinguishing between levels of within subject variables.

4.1. Demand, awareness, and clutter

The correlation analyses indicate that the awareness subjective measure is capable of assessing level 3 spatial awareness given that

Table 5
Repeated measures ANOVA results.

Dependent measure	Independent variable		
	Texture	GFOV	Texture × GFOV
Demand	$F(6,102) = 0.80$ $\eta_p^2 = 0.05$ $p = 0.57$	$F(1,17) = 0.00$ $\eta_p^2 < 0.01$ $p = 0.99$	$F(3.79,64.40) = 1.43^a$ $\eta_p^2 = 0.08$ $p = 0.24$
Awareness	$F(6,102) = 2.78$ $\eta_p^2 = 0.14$ $p = 0.02^*$	$F(1,17) = 0.03$ $\eta_p^2 < 0.01$ $p = 0.87$	$F(6,102) = 0.63$ $\eta_p^2 = 0.04$ $p = 0.71$
Clutter	$F(6,102) = 0.59$ $\eta_p^2 = 0.03$ $p = 0.74$	$F(1,17) = 1.13$ $\eta_p^2 = 0.06$ $p = 0.30$	$F(6,102) = 1.40$ $\eta_p^2 = 0.08$ $p = 0.22$
SA-SWORD 30°	$F(3.10,52.66) = 6.86^b$ $\eta_p^2 = 0.29$ $p < 0.01^*$	NA	NA
SA-SWORD 60°	$F(3.40,57.88) = 12.62^c$ $\eta_p^2 = 0.43$ $p < 0.01^*$	NA	NA

Note: superscripts indicated that a Greenhouse–Geisser ε correction was applied to the degrees of freedom due to a violation of sphericity: a$W = 0.11$, $\chi^2(20) = 32.34$, $p = 0.04$, $\varepsilon = 0.63$; b$W = 0.04$, $\chi^2(20) = 49.07$, $p < 0.01$, $\varepsilon = 0.52$; c$W = 0.04$, $\chi^2(20) = 46.34$, $p < 0.01$, $\varepsilon = 0.57$.
$^*p < 0.05$.

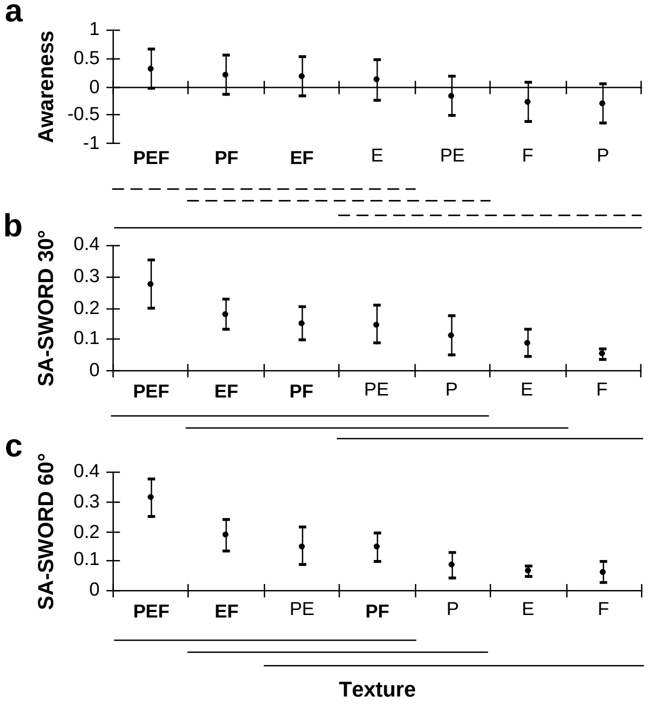


Fig. 7. Post-hoc analyses' results for the texture main effect on the subjective measures. Filled circles indicate means. Bars around means in (a) represent Tukey's intervals as indicated by a Tukey's HSD with $\alpha = 0.05$. Bars around means in (b) and (c) represent 95% confidence intervals. Solid lines under textures indicate homogeneous subsets (no significant differences between textures) as indicated by a Tukey's HSD (a) or a Bonferroni post-hoc (b and c). Dotted lines under the textures in (a) indicate homogeneous subsets as indicated by a least significant difference post-hoc. Bolded textures indicate which textures produced the least judgment error as indicated by Bolton et al. (2007).

it exhibited a negative trend with $|\tau_e|$ across the levels of texture × GFOV. This is interesting given that no correlations were observed between awareness and any of the judgment-based measures designed to probe level 2 spatial awareness ($|D_e|$, $|A_e|$, and $|H_e|$), and level 3 spatial awareness can be viewed as derivative of level 2 spatial awareness. A potential explanation can be found by examining Figs. 5 and 6. Here, even when a linear trend is not observed, we can see that the linear fit indicated a negative relationship between awareness and the level 2 spatial awareness judgment error terms. Thus, this general negative tendency across the three level 2 measures may have ultimately contributed to the trend observed for the $|\tau_e|$.

The fact that demand and clutter were not correlated with any of the judgment-based measures suggests that they are not measuring spatial awareness.

The less robust post-hoc analysis (least significant difference) showed that four textures were in the set that produced the highest awareness ratings (E, EF, PF, and PEF). This is consistent with results obtained by Glaab and Hughes (2003) who found that participants tended to give E, PF, and EF textures higher terrain awareness ratings than the F texture. Further, three of these textures (EF, PF, and PEF; Fig. 7) were among the set of textures that produced the minimum absolute error across all of the texture main and interaction effects (see Bolton et al., 2007). Considering only these results, when the less robust post-hoc analysis is used, awareness does seem to be capable of serving as reasonable proxy for the judgment-based measures.

The fact that no significant differences were observed between the independent variable levels for the demand, awareness, and clutter subjective dependent measures may be due to the nature of the task. As stated previously, Pew (2000) expressed concern that SART-like subjective measures were confounded with workload. Thus, had participants actually been flying the aircraft, they would have been subjected to more sources of workload. In this situation, their attentional resources would have been in higher demand and they might have been able to assess differences in texture's demand on attentional resources and their ability to convey spatial awareness more acutely. Further, the displays used in this experiment utilized a reduced set of instrumentation than is often employed in SVS display. Had participants seen a full set of SVS instrumentation, they may have shown greater differentiation in their clutter ratings.

While the judgment-based measures were able to produce significant results and convey significant differences between independent variable levels with the given procedure, the subjective measures were not. Given that subjective measures similar to those employed in this experiment have produced significant results for flight and simulation tests (Bailey et al., 2002; Glaab and Hughes, 2003), the demand and awareness measures may be more useful in high fidelity task environments.

4.2. SA-SWORD

The negative correlations observed between SA-SWORD and both $|D_e|$ and $|\tau_e|$ (an increase in judgment error corresponding to a decrease in SA-SWORD) suggest that SA-SWORD does provide insight into both level 2 ($|D_e|$) and level 3 ($|\tau_e|$) spatial awareness.

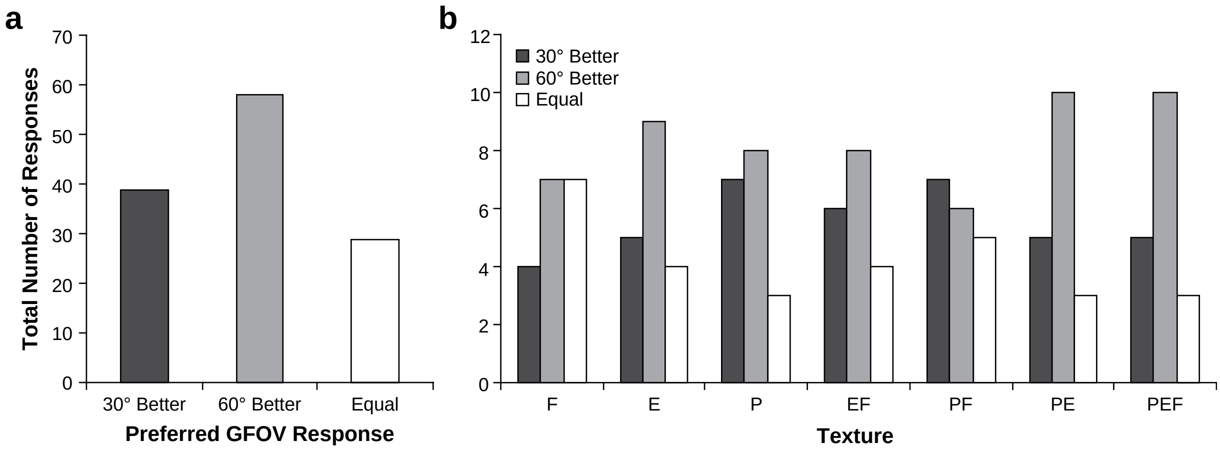


Fig. 8. (a) Aggregate preferred GFOV responses. (b) Aggregate preferred GFOV responses by texture.

Given that both awareness and SA-SWORD were negatively correlated with $|\tau_e|$ across texture × GFOV levels, it is not surprising that a positive correlation was observed between them.

The ANOVA results for SA-SWORD ratings did indicate that participants thought that some textures enhanced spatial awareness more than others. The textures that produced the highest SA-SWORD scores were somewhat consistent with the textures shown to produce the least amount of judgment error. EF, PF, and PEF were amongst the statistically similar textures that produced the highest scores for SA-SWORD 30° and SA-SWORD 60° and were amongst statistically similar textures that produced the minimum judgment error in both the main and interaction effects (see Bolton et al., 2007; Fig. 7).

4.3. Preferred GFOV

The lack of significant differences between GFOV preferences is consistent with the results from the judgment-based measures, where no absolute error terms were significant and there was no clearly dominating GFOV in the directional error terms. This is also consistent with the demand, awareness, and clutter variables for which GFOV was not a significant main effect. However, the results provide no insight into the measurement capabilities of discrete choice awareness metrics.

4.4. Conclusions

This work compared judgment-based measures of spatial awareness designed to probe all three levels of spatial awareness knowledge with subjective measures commonly used to evaluate SVS. Similarities were found between some of the judgment-based measures and subjective measures in both correlation analyses and ANOVA results. Thus it does appear that some of the subjective measures may be capable of providing insight into spatial awareness. The correlations of SA-SWORD and awareness against $|\tau_e|$ indicate that both may provide some insight into level 3 spatial awareness. This is good because accurate level 3 spatial awareness can often be associated with accurate level 1 and level 2 spatial awareness since it builds on the knowledge encompassed at these levels. Since subjective measures do not inherently disambiguate between awareness levels, their measurement of level 3 awareness may imply that they are capturing some spatial awareness information from lower levels.

Since none of the subjective measures exhibited correlations with $|A_e|$ and $|H_e|$, this constitutes a serious shortcoming in the subjective measures' ability to evaluate spatial awareness. Future work may want to investigate ways of measuring these dimensions subjectively.

This effort represents the first study comparing judgment-based measures that probed all three levels of spatial awareness with subjective measures commonly used to evaluate display designs. However, in order to investigate a range of display options in a timely manner, experimental trials were short and non-interactive. In addition, the terrain point was indicated using an artificial object. Given that some of the subjective measures that did not show significant results in this experiment have produced significant results in studies employing flight test and full-flight simulators, future experiments should investigate the judgment-based measures in more realistic flight and simulation tests. As spatial awareness is critical to other domains, future studies should also collect judgment-based and subjective measures. These studies would provide more data useful for exploring what aspects of human awareness the subjective measures are capturing. Such procedures would also facilitate a full SART probe to be included and compared with the other measures.

While the focus of this research was centered on spatial awareness with SVS displays, there are other potential applications. Subjective measures have been used to evaluate awareness for more than just SVS. Examples include, but are not limited to, the approach phase of flight for general aviation pilots under IFR and VFR (Saleem and Kleiner, 2005), weather systems (Bustamante et al., 2005), tunnel in the sky displays (Takallu et al., 2004), air traffic control systems (Adams et al., 2007), unmanned air vehicle management (Denford et al., 2004), and enhanced vision systems (Korn et al., 2004). Thus, given the subjective measures' deficiencies in assessing level 2 awareness in this study, awareness measures in other domains may show similar shortcomings.

When choosing metrics for an experiment, one must consider whether the type of metric will integrate properly into the experimental procedure. For example, one of the criticisms of SAGAT has been that the procedural pauses it, necessitates, have the potential to disrupt operator awareness, affecting the results (Sarter and Woods, 1991). Given that the new spatial awareness measures would require similar pauses if utilized in a higher fidelity simulations, a similar criticism could be levied. Conversely, all of the subjective measures evaluated in this experiment could be collected after a simulation run, subverting this potential confound. Thus, because subjective metrics may be more procedurally convenient in some circumstances, it is critical that researchers understand exactly what they measure so that they can design experiments to collect the data they need in a context that is appropriate. This work has contributed to this understanding and will hopefully direct further studies to advance it.

Acknowledgment

The authors would like to thank James Raymond Comstock, Jr. for his help and guidance during this study. The authors would also like to thank Trey Arthur, Daniel Burdette, Sheila Conway, Stephanie Guerlain, Regina Johns, Lynda Kramer, Kara Latorella, Neil O'Connor, Hayes Press, Valerie Plyler, Laurence Prinzel III, Paul Schutte, Jason Sweeters, Mohammad Takallu, Steve Williams, and the two anonymous reviewers.

References

Adams, C.A., Murdoch, J.L., Consiglio, M.C., Williams, D.M., 2007. Incorporating data link messaging into a multi-function display to support the Small Aircraft Transportation System (SATS) and the self-separation of general aviation aircraft. Applied Ergonomics 38, 465–471.

Adelstein, B.D., Ellis, S.R., 2000. Rotation and direction judgment from visual images head-slaved in two and three degrees-of-freedom. IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans 30 (2), 165–173.

Arthur, J.J., Prinzel III, L.J., Kramer, L.J., Parrish, R.V., Bailey, R.E., 2004. Flight Simulator Evaluation of Synthetic Vision Display Concepts to Prevent Controlled Flight into Terrain (CFIT). NASA Technical Report NASA/TP-2004-213008.

Bailey, R.E., Kramer, L.J., Prinzel III, L.J., 2006. Crew and display concepts evaluation for synthetic/enhanced vision systems. In: 2006 Proceedings of SPIE Conference on Enhanced and Synthetic Vision.

Bailey, R.E., Parrish, R.V., Arthur III, J.J., Norman, R., 2002. Flight test evaluation of tactical synthetic vision display concepts in a terrain-challenged operating environment. In: Enhanced and Synthetic Vision 2002, April 1–2, pp. 178–189.

Barfield, W., Rosenberg, C., 1995. Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. Human Factors 37 (1), 173–181.

Boeing Commercial Airplanes, 2007. Statistical Summary of Commercial Jet Airplane Accidents: Worldwide Operations 1959–2006. Retrieved April 15, 2008 from: http://www.boeing.com/news/techissues/pdf/statsum.pdf.

Bolton, M.L., Bass, E.J., 2008. Using relative position and temporal judgments to identify biases in spatial awareness for Synthetic Vision Systems. The International Journal of Aviation Psychology 18 (2), 183–206.

Bolton, M.L., Bass, E.J., Comstock, J.R., 2006. A toolset to support the development of spatial and temporal judgment experiments for synthetic vision systems. In: 2006 IEEE Systems and Information Engineering Design Symposium, April 28, 2006, Charlottesville, VA.

Bolton, M.L., Bass, E.J., Comstock, J.R., 2007. Spatial awareness in synthetic vision systems: using spatial and temporal judgments to evaluate texture and field of view. Human Factors 49 (6), 961–974.

Brace, N., Kemp, R., Snelgar, R., 2003. SPSS for Psychologists: a Guide to Data Analysis using SPSS for Windows. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Bustamante, E.A., Fallon, C.K., Bliss, J.P., Bailey III, W.R., Anderson, B.L., 2005. Pilots' workload, situation awareness, and trust during weather events as a function of time pressure, role assignment, pilots' rank, weather display, and weather system. International Journal of Applied Aviation Studies 5 (2), 347–367.

Comstock, J.R., Glaab, L.J., Prinzel, L.J., Elliott, D.M., 2001. Can effective synthetic vision system displays be implemented on limited size display spaces?. In: Proceedings of the 11th International Symposium on Aviation Psychology. The Ohio State University, Dayton, Ohio.

Denford, M.J., Steele, J.A., Roy, R., Kalantzis, E., 2004. Measurement of air traffic control situational awareness enhancement through radar support toward operating envelope expansion of an unmanned aerial vehicle. In: Proceedings of the 36th Conference on Winter Simulation, Washington, D.C., pp. 1017–1025.

Dorrian, J., Roach, G.D., Fletcher, A., Dawson, D., 2007. Simulated train driving: fatigue, self-awareness and cognitive disengagement. Applied Ergonomics 38 (2), 155–166.

Dorighi, N.S., Grunwald, A.J., Ellis, S.R., 1992. Perspective format for a primary flight display and its effect on pilot spatial awareness. In: Proceedings of the IEEE/AIAA 11th Digital Avionics Systems Conference.

Endsley, M.R., 1988. Situation awareness global assessment technique (SAGAT). In: Proceedings of the IEEE 1988 National Aerospace and Electronics Conference, pp. 789–795.

Endsley, M.R., 1995a. Measurement of situation awareness in dynamic systems. Human Factors 37 (1), 65–84.

Endsley, M.R., 1995b. Toward a theory of situation awareness in dynamic systems. Human Factors 37 (1), 32–64.

Endsley, M.R., Selcon, S.J., Hardiman, T.D., Croft, D.G., 1998. A comparative analysis of SAGAT and SART for evaluations of situation awareness. In: Human Factors and Ergonomics Society 42nd Annual Meeting, Chicago, IL, October 5–9.

Fracker, M.L., 1990. Attention gradients in situation awareness. In: Situational Awareness in Aerospace Operations AGARD-CP-478. Advisory Group for Aerospace Research and Development, Neuilly-Sur-Seine, France.

Gibson, J.J., 1986. The Ecological Approach to Visual Perception. Lawrence Erlbaum Associates, Hillsdale.

Glaab, L.J., Hughes, M.F., 2003. Terrain portrayal for head-down displays flight test. In: Proceedings of the 22nd Digital Avionics Systems Conference. DASC, Indianapolis, Indiana.

Hughes, M.F., Takallu, M.A., 2002. Terrain portrayal for head-down displays experiment. In: Proceedings of the International Advanced Aviation Conference. Anchorage, Alaska.

Khatwa, R., Roelen, A., 1998. An analysis of controlled-flight-into-terrain (CFIT) accidents of commercial operators, 1988 through 1994. Flight Safety Digest 17, 166–212.

Korn, B., Lorenz, B., Doehler, H., Többen, H., Hecker, P., 2004. Passive "Radar-PAPI" landing aids for precision straight-in approach and landing in low-visibility. International Journal of Applied Aviation Studies 4 (2), 13–38.

Marshak, W.P., Kuperman, G., Ramsey, E.G., Wilson, D., 1987. Situational awareness in map displays. In: Proceedings of the Human Factors Society 31st Annual Meeting, New York, NY, pp. 533–535.

McGreevy, M.W., Ellis, S.R., 1986. The effect of perspective geometry on judged direction in spatial information instruments. Human Factors 28 (4), 439–456.

Pew, R.W., 2000. The state of situation awareness measurement: heading toward the next century. In: Endsley, M.R., Garland, D.J. (Eds.), Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates Inc., Mahwah, pp. 33–47.

Saleem, J.J., Kleiner, B.M., 2005. The effect of nighttime and deteriorating visual conditions on pilot performance, workload, and situation awareness, in general aviation for both VFR and IFR approaches. International Journal of Applied Aviation Studies 5 (1), 107–120.

Sarter, N.B., Woods, D.D., 1991. Situation awareness: a critical but ill-defined phenomenon. International Journal of Aviation Psychology 1, 45–57.

Schnell, T., Lemos, K., 2002. Terrain Sampling Density and Texture Requirements for Synthetic Vision Systems. Rockwell Collins Final Report. Rockwell Collins, Cedar Rapids, IA.

Selcon, S.J., Taylor, R.M., 1990. Evaluation of the situational awareness rating technique, SART as a tool for aircrew systems design. In: Situational Awareness in Aerospace Operations 8. AGARD, Neuilly-Sur-Seine, France.

Stark, J.M., Comstock Jr., J.R., Prinzel, L.J., Burdette, D.W., Scerbo, M.W., 2001. A preliminary examination of situation awareness and pilot performance in a synthetic vision environment. In: Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting, October 8–12, pp. 40–43.

Stevens, J.P., 2002. Applied Multivariate Statistics for the Social Sciences. Lawrence Erlbaum Associates, Inc., Mahwah, NJ.

Takallu, M.A., Wong, D.T., Bartolone, A.P., Hughes, M.F., Glaab, L.J., 2004. Interaction between various terrain portrayals and guidance/tunnel symbology concepts for general aviation synthetic vision displays during a low en-route scenario. In: Proceedings of the 2004 Digital Avionics Systems Conference, Salt Lake City, UT.

Taylor, R.M., March 1990. Situational awareness rating technique, SART: the development of a tool for aircrew systems design. In: Situational Awareness in Aerospace Operations 17. AGARD, Neuilly-Sur-Seine, France, pp. 1–17.

Turner, A.D., 1996. Paired-comparison method for interval scaling. Human Factors 38 (2), 362–374.

Vidulich, M.A., Hughes, E.R., 1991. Testing a subjective metric of situation awareness. In: Proceedings of the Human Factors Society 35th Annual Meeting, September 2–6, vol. 2, pp. 1307–1311.

Vidulich, M.A., Ward, G.F., Schueren, J., 1991. Using the subjective workload dominance, SWORD technique for projective workload assessment. Human Factors 33 (6), 677–691.

Wells, M.J., Venturino, M., Osgood, R.K., 1988. Using target replacement performance to measure spatial awareness. In: Proceedings of the Human Factors Society 32nd Annual Meeting, Anaheim, CA.

Wickens, C.D., 2002a. Situation awareness and workload in aviation. Current Directions in Psychological Science 11 (4), 128–133.

Wickens, C.D., 2002b. Spatial Awareness Biases. NASA Technical Report ARL-02-6/NASA-02-4.

Yeh, Y., 1992. Spatial judgments with monoscopic and stereoscopic presentation of perspective displays. Human Factors 34 (5), 583–600.