Wei, J. & **Bolton, M. L.**(2017). Compression Rates and Spatial Judgment Biases Made from Synthetic Vision Perspective Displays. The *Journal of Aerospace Information Science*, 14(10), 523-532.

Compression Rates and Spatial Judgment Biases Made from Synthetic Vision Perspective Displays

Jiajun Wei* and Matthew L. Bolton†

University at Buffalo, State University of New York, NY 14260

In this work we investigated the relationship between error in human spatial judgments made from perspective displays with the compression rates used to represent the judged spatial quantities. Two-dimensional perspective displays are often used to represent 3D information to humans. Such displays can use different compression rates (actual distance conveyed per unit of screen distance) for identical spatial quantities. We used an existing data set in which spatial judgments (relative distance, angle, and elevation) were made about the relative location of a terrain point shown on a simulated aircraft synthetic vision systems display. We then measured the correlations between compression rates and associated judgment error. Correlations were computed for average participant judgments as well as for each participant from which an average correlation was computed. Significant negative correlations were found between compression rates and judgment error for all of the analyzed spatial judgments. These were particularly strong for relative distance judgment error. There were noticeable differences between the correlations of averages and the associated average correlations. Our results indicate that compression rates do bias the analyzed spatial judgments. However, there are significant individual differences in how the compression rates appear to impact judgment error.

I. Introduction

Human operators performing tasks that rely on spatial judgments will often use perspective displays: 2D displays that convey 3D information. For example, doctors use perspective displays to perform surgical procedures [1, 2]; unmanned vehicles convey real-time environmental images to remote controllers [3]; and synthetic vision systems (SVS) employ perspective displays to give pilots a first-person view of the terrain in front of their aircraft [4]. In all cases, the displays help human operators build spatial awareness [5] about the relative positions and trajectories of objects in the environments and facilitate the completion of their tasks.

SVS displays are particularly important to the aviation community because they are perspective displays designed to prevent controlled flight into terrain (CFIT). CFIT is one of the leading causes of death in commercial and general aviation [6]. SVS combat this problem by depicting a computer generated version of the terrain directly in front of the aircraft [4, 7–11]. These displays allow pilots to determine where their aircraft is in relation to the terrain regardless of outside visibility conditions. In this way, SVS facilitate pilot spatial awareness [5]: a pilot's ability to identify

---

[*] PhD Student, Department of Industrial and Systems Engineering, `jwei7@buffalo.edu`

[†] Assistant Professor, Department of Industrial and Systems Engineering, `mbolton@buffalo.edu`

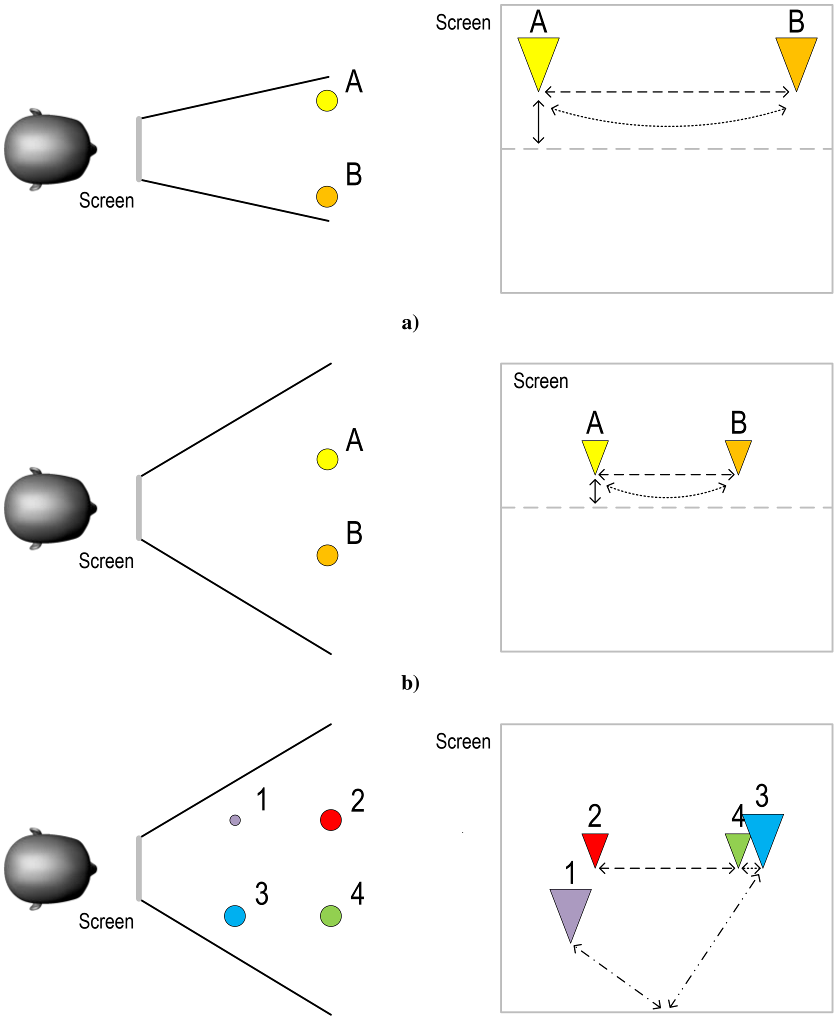
important terrain features near the aircraft, their relative locations, and the projection of those locations into the future based on the aircraft's current trajectory. SVS are of interest in this work because they represent an application domain in which spatial information derived from perspective displays can have profound safety implications.

SVS displays, like all perspective displays, represent a 3D volume of space projected onto a 2D surface. Perspective displays express spatial information (such as distance, angle, and orientation) in a way that has many human factors benefits. Specifically, operators perceive spatial information directly from the display without having to synthesize information from multiple displays that convey information for different dimensions. This is potentially advantageous because it can help reduce workload and accelerate the processing of information [12]. Such displays can also have a high ecological validity in that they represent 3D information as if it was being observed naturally. Further, SVS displays can use different terrain textures (images superimposed onto the displayed terrain) and symbology to facilitate depth cues and help people perform specific flight tasks.

However, there are biases in perspective display like SVS that exist independently of the terrain texture and other symbology. Because of the 3D-to-2D transformation associated with information presented on perspective displays (that is, 3D information transformed into 2D information), identical spatial quantities can be represented differently [5]. "Between-map" or between-display scale differences occur when the represented volume of space is different between two displays. "Within-map" or within-display scale differences occur when identical spatial quantities shown on a given display have different distances and/or orientations to the display's camera (the presumed location of the observer). Both of these are illustrated in Figure 1; Figure 1a and Figure 1b for between-display differences and Figure 1c for within-display differences. In these figures, the left side of the figure represents the actual relative distance between points in a 3D environment and the right hand side shows how they are displayed. Points in the environment are indicated by the tip of identically-sized inverted cones. A bent line is meant to represent visual angle between the two points for the observer.

For example, perspective displays have a geometric field of view (GFOV) that defines the angle between the lateral boundaries of space depicted on the display. As the GFOV is increased while the size of the display remains constant, the amount of spatial information on the display increases and the size of constant spatial quantities on the display decreases (appears smaller). Figures 1a and 1b illustrate how identical spatial quantities are represented between displays with different GFOVs. Specifically, points A and B are at identical relative locations to the viewer in both displays. As the GFOV increases from Figure 1a to Figure 1b, the same spatial quantities are depicted as being smaller. Note that between-display scale differences can also occur when the volume of space represented by the display doesn't change, but the size of the display is decreased. In this situation, quantities displayed on the original display would appear smaller on the smaller display.

Examples of within-display differences are shown in Figure 1c. In this, the points indicated by cones 2, 3, and 4 are all at the same relative height on the display's artificial horizon and point 1 is below the relative height of the other three points. For objects that are the same size (like the cones for points 1 and 2), the objects that are further away



a)

b)

c)

Figure 1. Figures illustrating compression rate differences of displayed spatial quantities between- (a and b) and within-displays (c).

will appear smaller than those that are closer. Similarly, distances that run parallel to the display's line of sight will be displayed with less screen real estate than identical distances that run perpendicular to it (this is sometimes called foreshortening). In Figure 1c, the distance between points 2 and 4 and between points 3 and 4 are identical, but the distance between 3 and 4 is displayed with much less screen space. Further, the distances of points 1 and 3 from the observer are identical, but because of the differences in the height of the points, the distance along the ground from the observer to the points on the display is smaller for point 1 than point 3.

Both between- and within-display differences can result in biased judgments about the spatial quantities that are depicted. These biases are theorized to be based on the relative compression rates of displayed spatial information [5]. Specifically, as the amount of screen space (inches) used to represent a spatial quantity goes down, the displayed compression rate (actual spatial quantity divided by the length of the quantity on the display) goes up. Thus, as the compression rate of a given spatial quantity increases, the spatial quality appears to be smaller on the display, and humans tend to perceive that the magnitude of the spatial quantity decreases.

This phenomenon has been used to explain biases in perspective displays. For example, McGreevy and Ellis [13] measured participants' ability to judge the elevation and azimuth angles between two objects on a perspective display during target tracking tasks. They found not only a significant effect of the GFOV on the magnitude of spatial judgment, but significant effects of azimuth angle within a given GFOV. They hypothesized, post-hoc, that the compression rates used to display information (caused by between-display differences for the different GFOVs and within-display differences for the azimuth angles) were responsible for these variations. Smallman et al. [14] suggested that variations in compression rates (due to within-display differences in orientation) were responsible for over and underestimations of different post lengths participants were expected to reproduce under different visual angles. Schreiber et al. [15] observed a linear degradation in humans' ability to match maps to perspective displays as a function of the amount of foreshortening in the display. Muthard and Wickens [16] found that display sizes (and thus compression rates between displays) could influence participant relative distance judgments for difficult judgment tasks. Stelzer and Wickens [17] evaluated aircraft pilots' abilities to accurately perform "tracking" tasks based on aircraft position information displayed on large and small displays. Pilots exhibited greater tracking error for the small display than with the large display, a phenomena that Stelzer and Wickens thought could be attributed to the relative increased compression in that display.

Bolton et al. [18, 19] used measures of spatial awareness that were based on the accuracy of human judgments made about the location (angle, distance, elevation, and abeam time) of a point, relative to the human observer, shown on terrain in a SVS display. In particular, Bolton and Bass [19] investigated a number of spatial biases theorized to exist in perspective displays used to convey terrain location information to pilots [5]. Most importantly to the discussed work, they found multiple instances of biases associated with within- and between-display differences in scale. These biases manifested between coded experimental levels (30°and 60°GFOVs, near and far relative distances, large and small angles, and elevations of points on the terrain both above and below the aircraft; these levels are described in more detail in the Methods section) for judgments about the point's relative locations (distance, angle, and elevation) that were consistent with reductions in the magnitude of spatial judgments with increases in compression. Specifically, between-display scale difference resulted in participants making significantly smaller relative angle and height judgments with the 60°GFOV (the more compressed display) than with the 30°GFOV. Biases due to within-display scale differences appeared to manifest in a number of situations: participants significantly underestimated point elevations more for small angles than for large angles; significantly underestimated relative angle judgements more for far distances than for near distances; underestimated in their relative distance judgments for far distances and overestimated them for near distances; and underestimated in relative elevation judgments more for far distances than for near distances.

The literature reviewed above shows that judgment biases manifest between display conditions associated with both between-display and within-display scale differences. The findings are consistent with the hypothesis that judgment magnitudes of a spatial quantity decrease with increases in the compression rate used to display that quantity for both between- and within-display differences. However, because the discussed literature has focused on the individual between-display and within-display biases, none have directly evaluated the relationship between compression rates and associated judgment biases. That is, if compression rates are indeed responsible for the observed biases, then we would expect to see a strong negative correlation between compression rates for a given displayed spatial quantity and the error in human judgments about that quantity. This work sought to test this hypothesis using the data set collected by Bolton and Bass [19]. Below we provide information for understanding this data set. We then describe our analysis procedure and our results. Finally, we interpret our findings and discuss future research directions.

II. The Data Set

The data set used in this experiment was collected by Bolton et al. [18–20]. The original study used novel, judgment-based measures for evaluating spatial awareness. Specifically, general aviation pilot participants were asked to make four relative judgments (angle, distance, elevation, and abeam time) about the relative location of a point on SVS displayed terrain. Error in these spatial judgments was used to assess pilot spatial awareness. Results from this effort were published over three papers. Bolton et al. [18] compared judgment accuracy for all four judgments with seven different textures and two GFOVs. Bolton and Bass [20] compared results between the judgment-based measures of spatial awareness and more conventional subjective measures that were commonly employed in SVS studies. The work reported by Bolton and Bass [19], which was covered in the introduction, used an analysis of variance to identify conditions where perspective display biases could result in differences in the human spatial judgments.

Because of the two GFOVs used in the experiment, between-display scale differences occur between trials. Because the location of each terrain point (relative angle, distance, and elevation) also varied between trials, there was variation in the compression rate used to display the location of the terrain point due to within-display scale differences. For these reasons, this was an appropriate data set for directly exploring the relationship between compression rates due to within- and between-display factors and judgment bias.

III. Methods

Below we describe the methods we used to collect and analyze the experimental data. As stated previously, the described experiment was conducted to facilitate the analyses described in [18–20]. As such, there are elements of the experimental design that are not relevant to work we present here. For this reason, there are several places in this section where we explicitly refer the readers to [18–20] for details on these elements.

A. Hypothesis

As compression rates increase, the screen distance used to represent the associated spatial quantity goes down. Thus, we expect the magnitude of human judgments to decrease relative to the actual value. As such, we hypothesized that there should be a negative correlation between the compression rate of a given spatial quantity (relative distance, relative angle, and relative elevation) and the respective human judgment error (relative distance error, relative angle error, and relative elevation error). The section below describes the experiment and data analysis procedures used to test this hypothesis.

B. Participants

The data set we used here was from a study that used eighteen general aviation pilot participants [18–20]. All had fewer than 400 hours of flight experience (M = 157, SD = 75). They had no previous experience with the experimental procedures used in the study or with SVS displays. They were familiar with the out-the-window view of general aviation aircraft.

C. Materials and apparatus

The experiment used to collect the data was conducted in a controlled, windowless laboratory with stable lighting. Each simulation was displayed by a computer workstation, and participants' judgments were collected automatically. Simulations, as shown in Figure 2, depicted SVS head-down displays. This display was 9.25 inches wide and 8 inches high. It was presented with an air speed indicator on the top left, an altimeter on the top right, and the GFOV on the bottom right. The SVS display itself had an artificial horizon running laterally across the center of the display, a radar altimeter in the upper right corner, a roll indicator in the upper center, pitch reference to either side of the center, and heading indicators in the dead center of the display and below the roll indicator. In the displayed simulations, a yellow inverted cone (d = 500 ft, h = 500 ft) was used to indicate the location of a terrain point. All simulations were stored as videos and showed SVS displays in straight, level flight at 127 knots for 5 seconds with a spatial resolution of 836×728 pixels and a refresh rate of 30 frames/second. Simulations were displayed using custom software [21].

D. Independent variables

The original experiment had five within-subject variables. They were texture, GFOV, and three scenario geometry variables representing the relative (azimuth) angle, distance, and elevation of the terrain point to ownship. The orders used to display textures and GFOVs to participants were between-subject variables.

There were seven textures used in the experiment and two GFOVs (30° and 60°). It is important to note that different terrain textures can contain different depth cues and can thus have a significant impact on the perception of spatial quantities [18, 22]. However, the purpose of the work presented here was to investigate the impact of compression rates on human spatial judgments. Thus, readers should refer to [18] for more details about the included textures and their impact on judgment accuracy.

The relative location of the terrain point displayed in a given scenario (the scenario geometry) was controlled based on its relative angle, distance, and elevation to the pilot in the simulated environment. Each of these three scenario geometry variables had two levels where, across trials, the value from each level was distributed as shown in Table 1.

Figure 2. The SVS display and symbology used to conduct experiments used as the source of the evaluated data set.

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

Independent Variable	Level	Range	Distribution
Relative Angle (*a*)	Small	[0°, 6.5°]	N (μ=3.75, σ=1.25)
	Large	[8.5°, 15°]	N (μ=11.25, σ=1.25)
Relative Distance (*d*)	Near	[1, 3.25 nmi]	N (μ=2.25, σ=0.417)
	Far	[3.75, 6 nmi]	N (μ=4.75, σ=0.417)
Relative Elevation (*e*)	Below	[-1000, -100 ft]	U (-1000, -100)
	Above	[100, 1000 ft]	U (100, 1000)

Scenario geometry and GFOVs were counterbalanced within each texture. During experiments, participants either saw all of the 30°GFOV trials first or all of the 60°GFOV trials first. Thus, GFOV order had two levels. Textures were also introduced to participants in one of three orders, where the orders textures were introduced in was counterbalanced between participants (see [18] for more details).

E. Dependent Measures

There were four dependent measures representing human judgments about the relative spatial location of the terrain point in each scenario: its relative distance (d_j), angle (a_j), elevation (e_j), and abeam time (t_j). All of these judgments were made using the interface shown in Figure 3. In the leftmost frame, a participant would make relative angle (in degrees) and distance (in nmi) judgments by using a mouse to position an 'x' on a top-down view similar to an aircraft's navigation display. In this, the relative angle and distance were computed relative to ownship, shown at the bottom of the display. Relative distance and angle were displayed next to the ‘x’ as it was being placed. All relative distances were positive and increased as the ‘x’ was moved away from ownship. Angles were made relative to ownship’s vector of displacement (the magenta line down the center of the display), where angle increased from zero as the ‘x’ moved to the left of the vector and decreased from zero as the ‘x’ moved to the right. The relative elevation judgments (in ft) were made using the frame in the upper-right of the display. In this, participants positioned an ‘x’ vertically on a side view of the aircraft. The relative elevation was displayed next to the ‘x’ as it was being placed. Relative elevations increased from zero as the ‘x’ moved up from the current elevation of the aircraft (shown by the horizontal magenta line) and decreased from zero as the ‘x’ moved down below it. Abeam time judgments (the time the participant thought it would take the aircraft to fly to the point of closest approach to the terrain point) were recorded in minutes and seconds. Between experimental blocks, subjective measures were also collected from participants related to the spatial awareness of the presented SVS display concepts. Because they do not factor into the hypotheses tested in this paper, neither the abeam time judgments nor the subjective measures are considered in the presented results. See Bolton et al. [18–20] for more information on these measures.

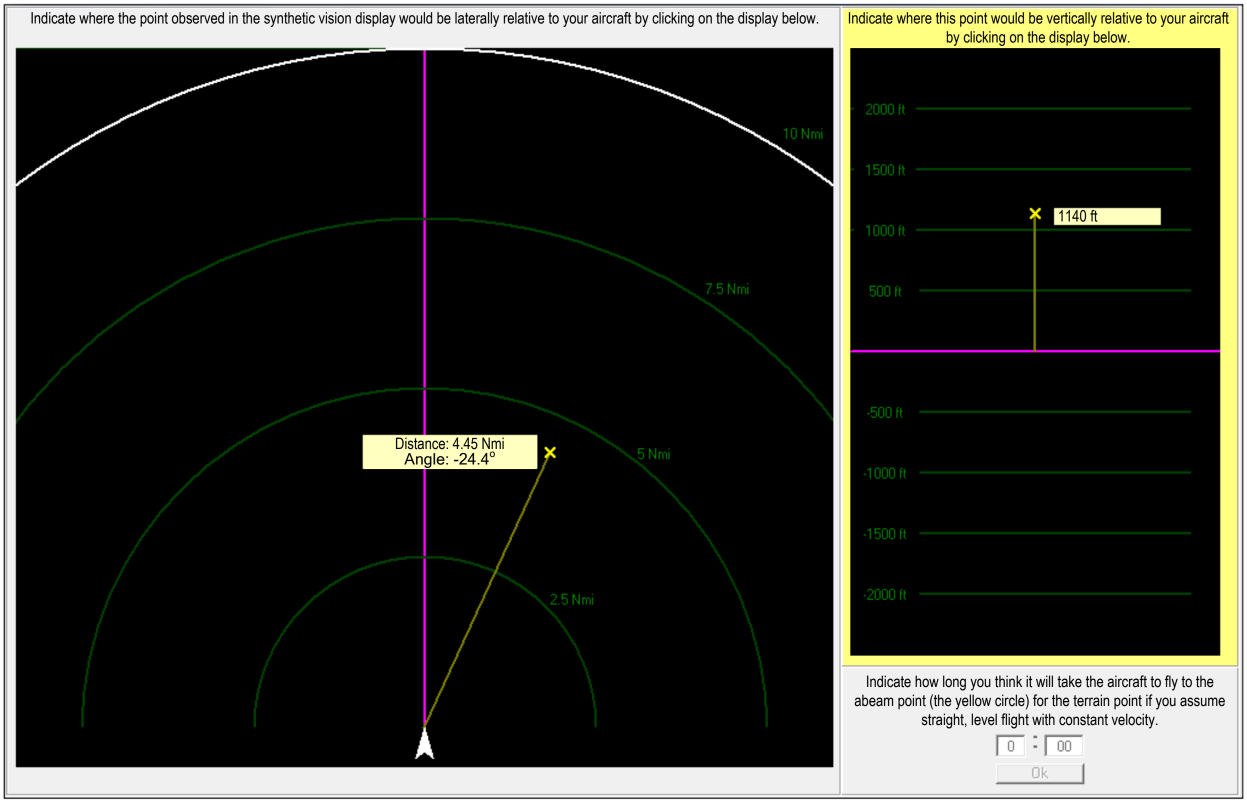
F. Procedure

Experiments were run in four-hour sessions. At the beginning of the experiment, participants completed consent forms and were then given a presentation introducing them to the experiment. They were then presented with trials. Each trial showed a participant a five-second simulation using the display discussed in Figure 2. This would pause for one second and the participant would be presented with the judgment display (Figure 3). This collected the spatial judgments about the scenario shown in the simulation. Note that the five-second simulation was used because it constituted a realistic upper bound on the amount of time a pilot might view a given display in flight while giving him or her enough time to observe the global optic flow of the scenario.

For each GFOV, the first twelve training trails were used to introduce the first terrain texture; for the other six textures, there were four training trials per texture. Training trials were meant to introduce pilots to new display concepts without significantly influencing judgment performance over the course of the experiment. During training trials, participants were given accuracy information about their judgments on the judgment display. No such accuracy feedback was given for non-training trials (see [18, 19] for more details). Subjective measures were collected between trial blocks of like textures and GFOVs (see [19, 20]). In total, participants saw 112 counterbalanced experimental trials (7 textures × 2 GFOVs × 2 relative angle levels × 2 relative distance levels × 2 relative elevation levels = 112) and 72 [2 GFOVs × (12 trials for the first texture + 6 × 4 trials per remaining texture) = 72] interspersed training trials.

Interface for making relative distance (d_j) and relative angle (a_j) judgments. This represent an above view of the aircraft similar to a nav display.

Interface for making the relative elevation judgment (e_j). This represents a side vew of the aircraft.



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.

Interface for making the abeam time judgment (t_j).

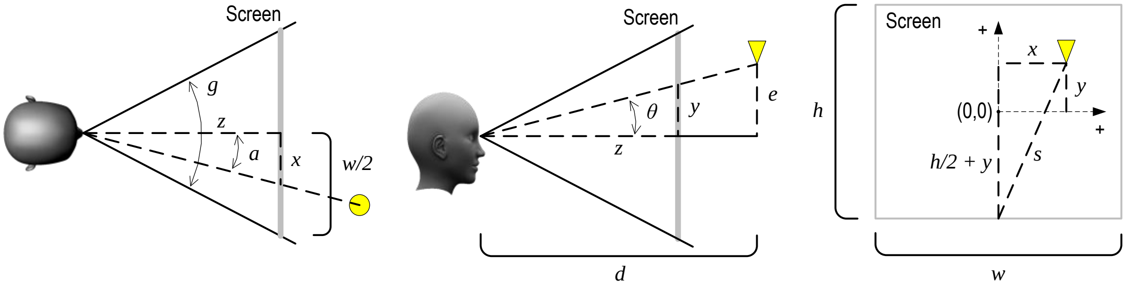
Figure 3. The judgment collection interface.

G. Experimental Design

This experiment used a mixed design with 18 participants. Participants were evenly and randomly distributed between the six combinations of the GFOV and texture presentation orders (see [18, 19] for more details).

H. Data Analysis

To be able to evaluate the connection between the compression rates of spatial quantities (spatial quantity per inch) and human spatial judgments about those quantities, we needed to compute the compression rates associated with each trial. To do this, we made several assumptions about what screen spatial qualities humans used to make their judgments independently of the terrain texture or other depth cues used in the display. For the relative angle judgment, we assumed that people used the information shown from the center of the display laterally across to the terrain point (x from Figure 4c; note that Figure 4c shows the SVS display as if a viewer was looking directly at it). For the relative elevation judgment, we assumed that people used the information shown from the center of the display vertical to the terrain point (y from Figure 4c). For the relative distance judgment, we assumed that people used the information shown from the bottom center of the display along the terrain leading up to the terrain point (s from Figure 4c). These are reasonable assumptions because the display quantities represent the ecological equivalents of what people would use if making judgments from a comparably sized window.



a) b) c)

$\tan(g/2) = \frac{w/2}{z}$ and thus $z = \frac{w/2}{\tan(g/2)}$

$\tan(a) = \frac{x}{z}$ and thus

$x = z \cdot \tan(a) = \frac{w \cdot \tan(a)}{2 \cdot \tan(g/2)}$

d)

$\tan(\theta) = \frac{e}{d} = \frac{y}{z}$ and thus

$y = \frac{e \cdot z}{d} = \frac{e \cdot w}{2 \cdot d \cdot \tan(g/2)}$

e)

$s^2 = x^2 + (h/2 + y)^2$ and thus

$s = \sqrt{x^2 + (h/2 + y)^2}$

f)

Figure 4. The relationships between the scenario geometry and screen coordinate systems used to calculate compression rates.

The values of these quantities (x, y, and s) were impacted by the size of the display, its GFOV, and the elements of scenario geometry. Figure 4 shows how x, y, and s were computed for each scenario. In this, Figures 4a and 4b show views of the display from above and in profile, respectively. They also illustrate the relationship between the position of the observer and the terrain point as if it were the actual point being viewed through a window of the same size of the display. Figure 4c shows the view of the display from the viewer's perspective. Figures 4d to 4f, show how x, y, and s are computed respectively. In all of these, g is the GFOV; h is the height of the display in inches; w is the width of the display in inches; z is the distance from the observer to the screen; a is the actual relative angle of the terrain point; d is the actual relative distance of the terrain point; e is the actual relative elevation of the terrain point; and θ is the vertical visual angle formed by the terrain point in the viewer's eye. From each scenario, g, h, w, a, d, and e were known.

The compression rates associated with the relative distance, angle, and elevation were then calculated using the variables from Figure 4. The compression rate of relative distance was calculated as

$$d_{comp} = \frac{d}{s} \text{ (nmi/inch).} \tag{1}$$

The relative angles compression rate was computed by

$$a_{comp} = \left| \frac{a}{x} \right| \text{ (degrees/inch).} \tag{2}$$

Finally, the compression rate associated with relative elevation was computed as

$$e_{comp} = \left|\frac{e}{y}\right| \text{ (ft/inch)}. \tag{3}$$

Note that Equation (1) is not computed with an absolute value because both d and s will always be positive.

In our analyses, we planned to correlate these compression rates with judgment error. Thus, judgment error values were computed for each of the considered human judgment types for each scenario. Each of these was calculated so that an overestimation of the value produced a positive judgment error and an underestimation produced a negative one. This means that, if judgment error decreased, then a participant either overestimated in a judgment less or underestimated in a judgment more. It does not mean that the overall magnitude of the error decreased. If we let d_j, a_j, and e_j represent a participant's relative distance, angle, and elevation judgments respectively (as defined above), distance error is

$$d_{error} = d_j - d, \tag{4}$$

angle error is

$$a_{error} = \begin{cases} a_j - a & \text{if } a \geq 0 \\ a - a_j & \text{otherwise} \end{cases}, \tag{5}$$

and elevation error is

$$e_{error} = \begin{cases} e_j - e & \text{if } e \geq 0 \\ e - e_j & \text{otherwise} \end{cases}. \tag{6}$$

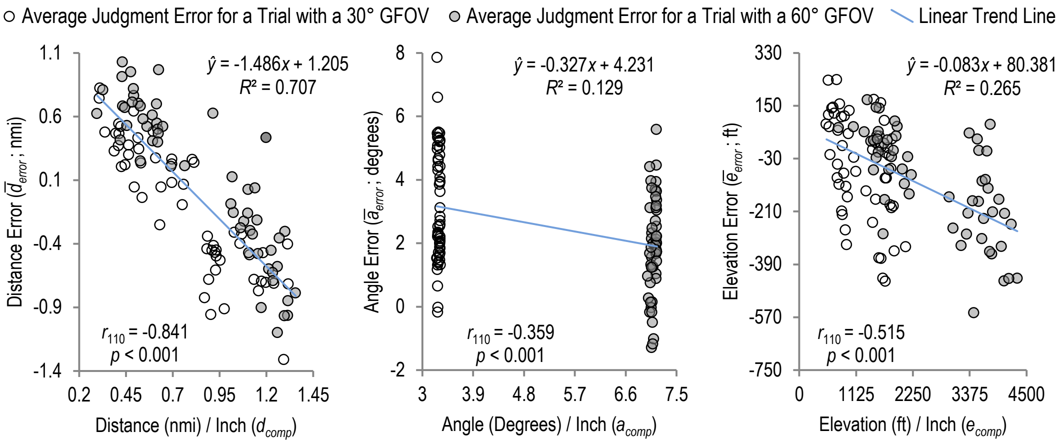
To determine how the compression rates correlated with the judgment error generally [23], the mean judgment error across participants was computed ($\overline{d}_{error}$, $\overline{a}_{error}$, and $\overline{e}_{error}$ for distance, angle, and elevation error, respectively) and Pearson's product-moment correlations (correlations of averages) were computed to measure the relationship between compression rates and the associated average judgment error over all of the trials. We tested whether the produced correlation coefficients were less than zero using a one-tailed t-test with an α level of 0.05. Linear models were also fit to the data using linear regression to evaluate how well compression rate predicted average human judgment error. Further, because we wished to see how well these values corresponded to the individual participants' judgments and their average correlations [23], we computed compression rate and judgment error correlations for each individual participant for all three spatial judgments across all of the trials and computed the average of these values (average correlations) using a Fisher's Z-transformation [24].

IV. Results

A total of 2016 valid trials were collected. $\overline{d}_{error}$, $\overline{a}_{error}$, and $\overline{e}_{error}$ were calculated and correlations of these averages and regression models were computed. Significant negative correlations of averages were found between compression rates and judgment error for all three of the human judgments considered. Statistics associated with these analyses are reported in Figure 5. Note that data points are distinguished in these plots based on their GFOV to make it clear how much of the variance in compression rates was due to between (GFOV) and within-display (scenario geometry) factors. Trend lines shown in each of these plots represent the linear-regression-produced equation over the range of observed compression values. These are presented with their associated linear model ($\hat{y}$) and R^2 value that indicates the goodness of the model's fit.

An examination of the regression analyses revealed that linear models appeared to do a reasonable job of describing how average judgment error changed as a function of compression rate for relative distance and elevation judgments (Figure 5a and c). However, only the linear model for relative distance (Figure 5a) explained a significant amount of the variance in the judgment error, as indicated by the R^2 values of the models (Figure 5). For the relative distance and elevation judgments, the models show average judgments move from overestimation (positive judgment error) to underestimation (negative judgment error) over the range of compression rates from our data. From the models, we can calculate that average human judgments will be accurate (have an average judgment error of zero) when the compression rates are 0.811 nmi/inch and 971.250 ft/inch for relative distances and elevations, respectively (note that this assumes the screen viewing distance used in the experiment). We can tell from the data in Figure 5b that, on average, participants tended to overestimate relative angle judgments. Thus, our data do not allow us to extrapolate out to find a compression rate that best facilitates relative angle judgment (the regression line does not cross the x-axis for the range of compression rates; Figure 5b).

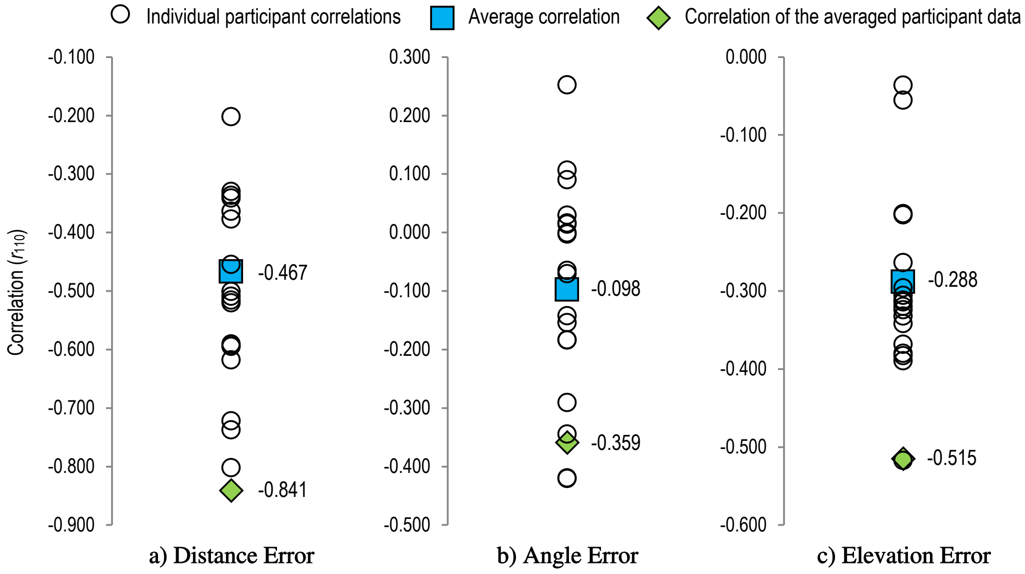
The correlations observed for each participant, the associated average correlation, and their relative magnitude compared to the correlations of averages (Figure 5) are shown in Figure 6. Note that in Figure 6, the average of the participants' correlations is shown as a blue square. The correlation computed from the data averaged across participants for each trial (Figure 5) is shown as a green diamond. These results show that there was both a fair



a) **b)** **c)**

Figure 5. Scatter plots showing the correlations between compression rates and the average judgment error for each trial.

amount of variance in the individual correlations and an average of noticeably less magnitude than the correlation achieved with the data averaged across participants.



a) Distance Error b) Angle Error c) Elevation Error

Figure 6. Dot plots showing the correlations seen between judgment error and compression rates for each participant.

V. Discussion

As we hypothesized, a significantly negative correlation was observed between the compression rate and judgment error for relative distance. Significant negative correlations were also observed between compression rates and judgment error for relative angle and relative height, which was also consistent with our hypotheses. However, these correlations were not as strong as those observed for distance. From this, it is clear that compression rate plays an important role in biasing human judgments about relative spatial quantities depicted on perspective displays, although not necessarily consistently for every spatial quantity judgment a person might make. Previous studies [5, 13–17, 19] have focused on looking for significant differences between experimental levels that could be associated with compression biases due to between-display and within-display differences in scale. By directly correlating display compression rates of spatial quantities with human judgment error about those quantities across between- and within-display factors, this work has provided the most definitive evidence to date that compression rates bias human spatial judgments from perspective displays. Thus, the work reported here constitutes a significant contribution.

The results of this paper also show that compression rates are linearly related to spatial judgment error for relative distances. This is significant because it could have potential design considerations. Specifically, our results suggest that there are compression rates that result in an average judgment error of 0. Designers could manipulate compression rates for judgment tasks to ensure that such an "optimal" compression rate is used. Alternatively, over- or underestimation of spatial quantities may be desirable in certain circumstances to elicit specific behavior from operators. For example, compression rates could be manipulated to make certain environmental threats seem closer and thus more pertinent to help ensure pilots respond to them in a timely manner.

Alternatively, instead of modifying the way perspective displays depict spatial quantities, training could be used to improve pilot judgments with such displays. However, it is not clear from this work how much training would be required to counteract the effect of the compression bias. It is also not clear how to best train pilots to avoid the compression bias.

Ultimately, the utility of such interventions will depend on how errors in spatial judgments impact actual flight performance. Thus, this should also be the subject of future work.

It is important to note that the data set used in the presented work was not originally intended for the analyses we reported here. As such, there are limitations in our results. Below we discuss these limitations and explore avenues of future work.

A. Confounds

Because the data set used in these analyses was originally meant for other purposes, there are some confounds in the design that could impact the results we present here. First, judgments were made with seven different terrain textures, thus allowing the spatial information in the textures to influence human judgment accuracies [18]. Further, training trials were distributed throughout the experiment. These training trials may have improved participant performance over the course of the experiment. However, the counterbalancing used in the experimental design should have mitigated the effects of these confounds on the results we presented here.

B. Range Restriction

A stronger correlation was observed between compression rate and relative distance error than between compression rates and error in the other two judgments. This may be due to the different ranges of compression rates seen for the three judgment spatial quantities. If we convert the range of compression rates for relative distances from nmi/in to ft/in ([1797.84 ft/in, 8243.36 ft/in]), we can see that it covers higher compression rates as well as a greater range of compression rates than the range seen for relative elevation ([537.05 ft/in, 4318.89 ft/in]). Further, both relative distances and relative elevations were represented over a greater range of compression rates than were seen for relative angles, which effectively had only two rates in the experiment (there were only very minor differences in relative angle compression rates observed within a given GFOV; Figure 5b). The more limited ranges of observed compression rates for relative elevations and angles might explain the lower observed correlations. Specifically, by having compression rates vary less than the others, the effect of other sources of variation (such as terrain textures that facilitate different depth cues; see [18]) will reduce the magnitude of the observed correlation. This phenomenon is generally known as range restriction [25]. Thus, stronger correlations would likely have been observed for more complete ranges of compression effects for relative elevation and relative angle. Future work should explore how the effect of these range restrictions could be mitigated.

C. Individual Differences

Our analyses computed both correlations of averages and average correlations. The discrepancies seen between the correlations of the averages (the green diamond in Figure 6) and the average correlation (the blue square in Figure 6) indicate that the biasing effect of compression manifests most strong when data is averaged over the general population, while the actual effect on any given individual will both vary and, on average, be less strong [23]. However, the results are consistent in that observed average correlations were negative and, for distance error and elevation error, of the same order of magnitude as their correlations of averages counterparts. This suggests that compression rates are still important for individual judgments, even if the effect is less than it would be for the population in aggregate.

However, the discrepancies between comparable statistics computed in these two ways and the large ranges of correlations observed for individual participants (see Figure 6) suggest that there are significant individual differences between participants. As such, it may not be a good idea to attempt to compensate for the compression biases based on the average because it could negatively impact individuals who significantly differ from the average. Should designers want to account for the effects of compression biases in their displays, they might need to allow for customization to account for the variance of individuals. For example, tests could be used to capture how compression rates affect specific individuals so that the system could provide customized judgment support or otherwise compensate for individually biased behavior. Future work should investigate why compression rates bias individuals differently and how to best address these biases in display/interface designs.

D. Judgement Task

Our results were generally consistent with the compression effects observed in the literature [5, 13–17, 19]. However, our results disagree with those reported by Alexander et al. [26]. In the Alexander et al. study, participants reproduced the relative angle of a point shown on a perspective SVS display on an out-the-window display. Alexander et al. found that participants underestimated angle judgments more for displays with smaller GFOVs. A finding that contradicts the results one would expect based on the effects of compression. It is possible that this discrepancy occurs because of the differences in the spatial transformations required for the judgment task. Alexander et al. [26] had participants map SVS spatial information to out-the-window compared to the judgment task we employed that has participants transform SVS information to a navigation-like display. If this is indeed the case, then it suggests that the nature of the judgement tasks and the spatial transformations they entail could impact how compression effects manifest. Given that the Alexander et al. [26] paper is only one study out of many to produce contradictory results, it is not clear whether this is the case. Future research should investigate what impact judgment task has on error in spatial judgments within the aviation domain as well as which spatial judgment task most relates to pilot flight performance.

E. Other Display Considerations

As we discussed in the introduction, there are other properties of perspective displays that can impact compression rates. These include display size, display resolution, other GFOVs, as well as different displayed trajectories and camera orientations [8, 9, 22, 26–29]. Given that the results presented here suggest that compression rates can be used to explain judgment biases that manifest for both between- and within-display differences, they should be able to explain the biases seen in all of these studies. Future work should investigate if any of the above display factors produce correlations consistent with the results presented here.

We know from Bolton and Bass [19] and the extended SVS literature [8, 9, 22, 30] that terrain texture can impact human spatial perception of quantities shown in SVS displays. Further, wireframes can be used to highlight important elements (such as runways or flightpaths) in SVS displays and can be set to convey different spatial information [7]. Various flight instrument data can be depicted on the SVS display to provide pilots with measures pertinent to the location of the aircraft including its altitude, yaw, and pitch. Future work should investigate how these factors influence and/or interact with the compression rate bias seen in this work.

Additionally, new types of 3D displays are becoming increasingly common, where most of these rely on one or more perspective displays to convey their information. Yeh and Silverstein [31] found evidence of within-display scale differences resulting in judgment biases in binocular (3D) displays. Future work should explore how compression rates impact spatial judgements in 3D displays similar to how it was analyzed here.

F. Other Biases

1. Temporal Biases

The original study that used the same data as the work presented here also collected abeam time judgments. Such judgments have been known to be biased in favor of distance; this is sometimes referred to as τ bias [5, 19]. If that is the case, temporal judgments could be impacted by compression rates. In fact, Muthard and Wickens [16] observed similar biases between relative distance and temporal judgments impact by compression rates between displays. However, it is not clear exactly what screen distance would represent the relative distance used in abeam time judgments. Future work should more deeply explore the connection between compression and τ biases.

2. The Virtual Space Effect

The virtual space effect [5, 13, 32] describes a bias that can occur with perspective displays when the field of view (FOV) of the display (the actual angle formed between the eyes of the human observer and the vertical edges of the display) is not in unity for the GFOV. Specifically, the virtual space effect occurs because of the “window assumption,” [13] where a person treats the perspective display as if the spatial information being observed through it are being observed through an equally sized window. When the FOV is smaller than the GFOV, people will interpret spatial quantities as being smaller than they actually are. When the FOV is larger than GFOV, people will interpret spatial quantities as being bigger than they actually are. This is slightly different from the compression-based biases we discuss in this paper because this impacts the magnitude of spatial quantities in the persons eye instead of the way they are depicted on the display. Due to the distance of participants from the computer displays, the FOV used in the reported work were approximately 18°[19] (though this was not strictly controlled for). This suggests that the virtual space effect could have generally biased human judgments (an in-depth discussion of this can be found in [19]). However, because FOV was kept relatively constant throughout the experiment, any impact the virtual space effect could have had can be explained by the between-display differences in GFOV. Future research should investigate what interaction effects could exist between the virtual space effect and the compression-based biases we explored in this paper.

It should be noted that displays presented in this experiment were shown on a desktop computer screen. Thus, the displays were not presented in the same position and orientation that they would be in an actual cockpit. This could lead to differences in spatial perception similar to those associated with the virtual space effect. However, this does not negate the findings of this paper as the compression effects we observed should still be a factor for different display positions and orientations.

G. Additional Work Environments

The judgment task presented here was aviation specific. However, there are many other critical work domains where human judgments about the relative location of object are made from perspective displays. These include medicine (e.g. surgical displays; [1, 2]), unmanned aerial vehicles [3], and driving (e.g. backup cameras). It is not clear how well our results will generalize to these work domains due to differences in display parameters and human operator tasks. This should be the subject of future research.

VI. Conclusion

The results presented here show a clear relationship between relative distance judgment error and the compression rates used to display distances on SVS displays. Similar, though less pronounced, relationships were seen for relative angle and relative elevation judgments. This suggests that compression rates do generally bias spatial judgments, where the judgment magnitudes decrease as compression rates increase. However, differences seen between the correlations of averages and the average correlations for these concepts suggest that there are significant individual differences that will need to be addressed in future research.

References

[1] Aykan, S., Singhal, P., Nguyen, D. P., Yigit, A., Tuken, M., Yakut, E., Colakerol, A., Sulejman, S., and Semercioz, A., "Perioperative, Pathologic, and Early Continence Outcomes Comparing Three-Dimensional and Two-Dimensional Display

Systems for Laparoscopic Radical Prostatectomy–A Retrospective, Single-Surgeon Study," *Journal of Endourology*, Vol. 28, No. 5, 2014, pp. 539–543.

[2] Crosthwaite, G., Chung, T., Dunkley, P., Shimi, S., and Cuschieri, A., "Comparison of Direct Vision and Electronic Two-and Three-dimensional Display Systems on Surgical Task Efficiency in Endoscopic Surgery," *British Journal of Surgery*, Vol. 82, No. 6, 1995, pp. 849–851.

[3] Donovan, S.-L. and Triggs, T., "Investigating the Effects of Display Design on Unmanned Underwater Vehicle Pilot Performance," Tech. Rep. DSTO-TR-1931, Maritime Platforms Division, Defence Science and Technology Organisation, Victoria, Australia, 2006.

[4] Ford, T., "Synthetic Vision Systems," *Aircraft Engineering and Aerospace Technology*, Vol. 66, No. 5, 1994, pp. 2–4.

[5] Wickens, C. D., "Spatial Awareness Biases," Tech. Rep. AHFD-02-6/NASA-02-4, Aviation Human Factors Division Institute of Aviation, University of Illinois at Urbana-Champaign, Savoy, 2002.

[6] Boeing, "Statistical Summary of Commercial Jet Airplane Accidents Worldwide Operations — 1959 – 2014," , 2015. `http://www.boeing.com/assets/pdf/news/techissues/pdf/statsum.pdf`.

[7] Parrish, R. V., Bailey, R. E., Kramer, L. J., Jones, D. R., Young, S. D., Arthur, J. J., Prinzel, L. J., Glaab, L. J., and Harrah, S. D., "Aspects of Synthetic Vision Display Systems and the Best Practices of the NASA's SVS Project," Tech. Rep. NASA/TP-2008-215130, NASA Langely Research Center, Hamptom, 2008.

[8] Prinzel, L. J. and Wickens, C. D., editors, *Special Issue: Synthetic Vision Systems, The International Journal of Aviation Psychology*, Vol. 19, No. 1, 2009.

[9] Prinzel, L. J. and Wickens, C. D., editors, *Special Issue: Synthetic Vision Systems, Part II, The International Journal of Aviation Psychology*, Vol. 19, No. 2, 2009.

[10] Mulder, J., Van Paassen, M. M., and Mulder, M., "Perspective Guidance Displays Show the Way Ahead," *Journal of Aerospace Computing, Information, and Communication*, Vol. 1, No. 11, 2004, pp. 428–431.

[11] Brychcy, M. L., Shanmugasundaram, R., Allen, D. N., and Harris, K. S., "Flight Simulation Architecture for Tunnel-in-the-sky Guidance and Synthetic Vision," *Journal of Aerospace Computing, Information, and Communication*, Vol. 1, No. 5, 2004, pp. 239–252.

[12] Smallman, H. S., St. John, M., Oonk, H. M., and Cowen, M. B., "Information Availability in 2D and 3D Displays," *IEEE Computer Graphics and Applications*, Vol. 21, No. 5, 2001, pp. 51–57.

[13] McGreevy, M. W. and Ellis, S. R., "The Effect of Perspective Geometry on Judged Direction in Spatial Information Instruments," *Human Factors*, Vol. 28, No. 4, 1986, pp. 439–456.

[14] Smallman, H. S., St. John, M., and Cowen, M. B., "Use and Misuse of Linear Perspective in the Perceptual Reconstruction of 3-D Perspective View Displays," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 46, SAGE Publications, Los Angeles, 2002, pp. 1560–1564.

[15] Schreiber, B. T., Wickens, C. D., Renner, G. J., Alton, J. D., and Hickox, J. C., "Navigational Checking Using 3D Maps: The Influence of Elevation Angle, Azimuth, and Foreshortening," *Human Factors*, Vol. 40, No. 2, 1998, pp. 209–223.

[16] Muthard, E. and Wickens, C. D., "Display Size Contamination of Attentional and Spatial Tasks: An Evaluation of Display Minification and Axis Compression," Tech. Rep. AHFD-05-12/NASA-05-3, Aviation Human Factors Division: Institute of Aviation, University of Illinois at Urbana-Champaign, 2005.

[17] Stelzer, E. M. and Wickens, C. D., "Pilots Strategically Compensate for Display Enlargements in Surveillance and Flight Control Tasks," *Human Factors*, Vol. 48, No. 1, 2006, pp. 166–181.

[18] Bolton, M. L., Bass, E. J., and Comstock, J. R., "Spatial Awareness in Synthetic Vision Systems: Using Spatial and Temporal Judgments to Evaluate Texture and Field of View," *Human Factors*, Vol. 49, No. 6, 2007, pp. 961–974.

[19] Bolton, M. L. and Bass, E. J., "Using Relative Position and Temporal Judgments to Identify Biases in Spatial Awareness for Synthetic Vision Systems," *The International Journal of Aviation Psychology*, Vol. 18, No. 2, 2008, pp. 183–206.

[20] Bolton, M. L. and Bass, E. J., "Comparing Perceptual Judgment and Subjective Measures of Spatial Awareness," *Applied Ergonomics*, Vol. 40, No. 4, 2009, pp. 597–607.

[21] Bolton, M. L., Bass, E. J., and Comstock, J. R., "A Toolset to Support the Development of Spatial and Temporal Judgment Experiments for Synthetic Vision Systems," *Systems and Information Engineering Design Symposium, 2006 IEEE*, IEEE, Piscataway, 2006, pp. 55–60.

[22] Schnell, T., Keller, M., and Etherington, T., "Trade-offs in Synthetic Vision System Display Resolution, Field of Regard, Terrain Data Density, Texture, and Shading During Off-path Operations," *The International Journal of Aviation Psychology*, Vol. 19, No. 1, 2009, pp. 33–48.

[23] Monin, B. and Oppenheimer, D. M., "Correlated Averages vs. Averaged Correlations: Demonstrating the Warm Glow Heuristic Beyond Aggregation," *Social Cognition*, Vol. 23, No. 3, 2005, p. 257.

[24] Fisher, R. A., "On the Probable Error of a Coefficient of Correlation Deduced from a Small Sample," *Metron*, Vol. 1, 1921, pp. 3–32.

[25] Goodwin, L. D. and Leech, N. L., "Understanding Correlation: Factors that Affect the Size of *r*," *The Journal of Experimental Education*, Vol. 74, No. 3, 2006, pp. 249–266.

[26] Alexander, A. L., Wickens, C. D., and Hardy, T. J., "Synthetic Vision Systems: The effects of Guidance Symbology, Display Size, and Field of View," *Human Factors*, Vol. 47, No. 4, 2005, pp. 693–707.

[27] Alexander, A. L., Wickens, C. D., and Hardy, T. J., "Examining the Effects of Guidance Symbology, Display Size, and Field of View on Flight Performance and Situation Awareness," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 47, SAGE Publications, Los Angeles, 2003, pp. 154–158.

[28] Borst, C., Mulder, M., and Van Paassen, M., "Design and Simulator Evaluation of an Ecological Synthetic Vision Display," *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 5, 2010, pp. 1577–1591.

[29] Arents, R., Groeneweg, J., Borst, C., Van Paassen, M., and Mulder, M., “Predictive Landing Guidance in Synthetic Vision Displays,” *Open Aerospace Engineering Journal*, Vol. 4, 2011, pp. 11–25.

[30] Schnell, T. and Lemos, K., “Terrain Sampling Density and Texture Requirements for Synthetic Vision Systems,” Tech. rep., Iowa Space Grant Consortium, Iowa City, 2002.

[31] Yeh, Y.-Y. and Silverstein, L. D., “Spatial Judgments with Monoscopic and Stereoscopic Presentation of Perspective Displays,” *Human Factors*, Vol. 34, No. 5, 1992, pp. 583–600.

[32] McGreevy, M. W., Ratzlaff, C., and Ellis, S., “Virtual space and two-dimensional effects in perspective displays,” *Proceedings of the 21st annual conference on Manual Control*, 1985, pp. 29–1.