Faculty and Researchers   Faculty and Researchers Collection

2013-09

Pilot Perception and Confidence of Location During a Simulated Helicopter Navigation Task

Yang, Ji Hyun; Cowden, Bradley T.; Kennedy, Quinn; Schramm, Harrison; Sullivan, Joseph

Pilot Perception and Confidence of Location During a Simulated Helicopter Navigation Task

Ji Hyun Yang, Bradley T. Cowden, Quinn Kennedy, Harrison Schramm, and Joseph Sullivan

Yang JH, Cowden BT, Kennedy Q, Schramm H, Sullivan J. *Pilot perception and confidence of location during a simulated helicopter navigation task.* Aviat Space Environ Med 2013; 84:952–60.

Introduction: This paper aims to provide insights into human perception, navigation performance, and confidence in helicopter overland navigation. Helicopter overland navigation is a challenging mission area because it is a complex cognitive task, and failing to recognize when the aircraft is off-course can lead to operational failures and mishaps. **Methods:** A human-in-the-loop experiment to investigate pilot perception during simulated overland navigation by analyzing actual navigation trajectory, pilots' perceived location, and corresponding confidence levels was designed. There were 15 military officers with prior overland navigation experience who completed 4 simulated low-level navigation routes, 2 of which entailed auto-navigation. This route was paused roughly every 30 s for the subject to mark their perceived location on the map and their confidence level using a customized program. **Results:** Analysis shows that there is no correlation between perceived and actual location of the aircraft, nor between confidence level and actual location. There is, however, some evidence that there is a correlation ($\rho= -0.60 \sim -0.65$) between perceived location and intended route of flight, suggesting that there is a bias toward believing one is on the intended flight route. **Discussion:** If aviation personnel can proactively identify the circumstances in which usual misperceptions occur in navigation, they may reduce mission failure and accident rate. Fleet squadrons and instructional commands can benefit from this study to improve operations that require low-level flight while also improving crew resource management.

Keywords: navigation, perception, confidence, bias, terrain association.

OVERLAND VISUAL navigation at low altitudes, which we define as flight at or below 200 ft (~61 m) above ground level (AGL) is an increasingly important task for rotary wing aviators. Surprisingly, the factors that determine success in low level navigation are not well understood. Our research seeks to further the study of aviation by quantitatively studying pilot performance in a controlled experimental environment.

Airborne navigation—the act of understanding where the aircraft is and which direction it should travel next—is important both for mission accomplishment and hazard avoidance. If the aircraft is not where it is supposed to be, it cannot accomplish its mission. Hazard avoidance encompasses both point hazards, such as power lines, and area hazards, such as active ranges. Military settings include hazards of enemy action, which may be of either type. At higher altitude, navigation may be performed by various means including: dead reckoning, visual navigation, radio aids to navigation, global positioning system (GPS), and inertial navigation systems. GPS and inertial navigation systems are frequently combined, and are referred to as hybrid navigation or simply G/INS. A summary of several common methods follows; for details, see Eschenbach and Stanski-Pacis and de Voogt (6,10). Adam et al. (1) address issues arising around the usability and potential pitfalls with current cockpit GPS systems and Casner (2) discusses training requirements for GPS usage.

The low-level navigation environment is different from navigation at altitude for several reasons. Radio aids to navigation may be unreliable. This increases the relative importance of other methods, particularly visual navigation. Visual navigation also is of increased importance at low altitudes for hazard avoidance. Although training is a part of all navigation tasks, it is most critical for visual navigation. A look at the Naval Safety Center's statistics page (8) points to the importance and risk of helicopter overland navigation. For example, on 21 December 2011, an MH-60S struck trees and crashed in an open area during a day mountain flight.

Crew coordination at low altitudes requires division of duties between the flying pilot, who we henceforth refer to as the 'pilot at the controls' (PAC) and the nonflying pilot, who we henceforth refer to as the 'pilot not at controls' (PNAC). The PAC is typically responsible for the tasks required to safely pilot the aircraft and for critical responses during emergencies. The PNAC is responsible for communication, planning, and navigation. Both pilots are responsible for the identification and avoidance of obstacles, as appropriate. We are reminded in the work of de Voogt et al. (5) that the notion of 'crew' frequently includes those who are not physically present in the aircraft, including other aircraft in a formation, controllers, and ground crews.

Broadly speaking, a pilot may be on-course or off-course, and he may perceive himself to be on-course or off-course. Sullivan (11) summarizes this, as does **Table I**.

Assessing Navigation Performance	Low Confidence	High Confidence
Low Correctness	Struggling. No accurate fix, aware that aircraft is off-track	Dangerous. Lost and does not realize it. Positively misidentified correlating features.
High Correctness	On course and lucky. Accurate fix, but not confident in navigation solution.	Skilled performer. On track and certain.

Of the four possible combinations, the one of greatest concern is when the crew is off-course, but self-assesses that they are on-course (labeled "Dangerous" in Table I). This is because pilots who are lost and are aware of such are most likely to take corrective action; these corrective actions frequently involve help from other crewmembers or controlling agencies and were outside of our experiment.

Sullivan et al. (12) led us to wonder about perception biases in navigation. For example, subject 5 in Sullivan et al. (12) missed a waypoint and started to track north of the intended route (12,13). The subject missed the waypoint and took a left turn into a valley located to the north of the waypoint. Then, he flew north of the intended trajectory, believing he was on-track. Instead of using available visual cues on the flight simulation screen or out-the-window (OTW) view to realize that he was off course, he perceived that he was still on course. This pattern suggests the pilot was using some biased visual cueing in which he overweighed OTW cues that fit into his perception that he was on course, and disregarded OTW cues that did not fit with the hypothesis. Correct navigation procedures consist of observing elapsed time, noting the expected position of the aircraft, and then observing the outside world to see if this is correct, and adjusting as necessary (this technique is referred to as "Clock to Map to Ground"). An incorrect navigation procedure, or bias, occurs when the pilot reconciles the difference between the expected position and observed ground features by forcing his expected position to be his observed position in his mental model. The forcing is accomplished by adjusting the (internal) weighting parameters for the various navigation queues to give high weight to those that reinforce the pilot's hypothesized position and giving low weight to or simply discarding those cues that disagree with his hypothesis.

Not only did the subject show misperception, but he also was very confident in his misperception. This somewhat unexpected phenomenon was observed in several pilots during the experiment and we started to question whether one's confidence is indeed correlated to their navigation performance or not, or that one's confidence may even increase with greater navigation errors. Next, we noted that subjects' bias did not seem arbitrary, rather it showed a pattern or consistency. The bias was toward the planned route trajectory. Connecting with their observed high confidence, we started to question whether pilots' high confidence is related to this bias toward the intended route rather than their actual current location.

Our study follows previous work (4,12,13) and proposes to support previous experimental observations, with particular attention paid to the common belief that high confidence is a good indicator of good performance (4,9). The goal of our experiment was to place subjects with various levels of navigation experience with limited visual terrain cues. To test our notions about perception, performance, and confidence, we test several hypotheses, listed below:

> Hypothesis 1: There is a negative correlation between confidence and distance from subject's perceived location and actual helicopter position.
> Hypothesis 2: There is a negative correlation between confidence and distance from subject's perceived location and the intended route of flight.
> Hypothesis 3: The longer a subject navigates through an intended route, the greater the distance between the perceived location and the actual helicopter position.
> Hypothesis 4: Confidence decreases the further into the navigation route.

METHODS

Subjects

There were a total of 15 subjects ranging from 27 to 41 yr of age, with an average age of 36 yr (SD = 4.8). Total flight hours (TFH) ranged from 0 to 2500 with an average of 1431 (SD = 803.5). One subject had 0 TFH, three subjects had TFH between 0 and 1000 h, five subjects had TFH between 1000 and 2000 h, and six subjects had TFH between 2000 and 3000 h. Total overland hours (OFH) ranged from 0 to 2000 with an average of 870 h (SD = 634.2). One subject had 0 h, six subjects had OFH between 0 and 1000 h, seven subjects had OFH between 1000 and 2000 h, and one subject had OFH between 2000 and 3000 h. For those who had flight experiences, their last flight was within a year for five subjects, within 3 yr for six subjects, within 5 yr for one subject, and within 10 yr for two subjects. Their last overland flight was within a year for five subjects, within 3 yr for six subjects, within 10 yr for two subjects, and more than 10 yr for one subject. There were 11 U.S. Navy, 4 U.S. Army subjects, and 1 Hellenic Air Force subject. To participate in this experiment the subject needed to have overland navigation training. Subjects for this study were recruited from the Naval Postgraduate School student body and faculty. Recruitment was completed through an IRB approved e-mail sent to Operations Research Department and Modeling Virtual Environments and Simulation students.

Equipment

The software used to run the simulation was Image Generator, Terrain & Map D8, Data Logger by Delta3D,

and OpenSceneGraph. These programs used inputs through X-Plane 9.21rc2, a commercially used flight simulator. The software converts the X-Plane data into the OTW and map views based on the subjects inputs. A 43" by 24" screen was used to present the OTW view and 33" by 33" display for the map and cockpit view. The X-Plane and the Image Generator were set with a modernized autopilot flying at 150 ft (~46 m) AGL at 65 kn. This altitude remained fixed throughout the route and maintained obstacle clearance in the mountainous terrain. Moving the joystick up and down did not affect the pitch of the aircraft, but did allow the subject to look up and down. The roll of the aircraft was completed with left or right joystick inputs. This put the aircraft in coordinated turns. The software also updated the instruments to correspond with the current flight profile.

A confidence app software was created in order to gain useful confidence output data. This program allowed the subjects to click where they perceived themselves to be on the map display. After the subjects right-clicked on the map, a red dot showed on the screen and a confidence scroll bar appeared. This confidence bar allowed the subjects to rate how confident they were of their perceived location. This bar ranged from 100, very confident, to 0, very lost. After the route was complete, the software also created a text file that contained the elapsed time of when the subject made his location estimate, the actual helicopter latitude and longitude, the subject's estimated latitude and longitude, and the subject's confidence on his perceived location.

Visual Navigation Task

The route needed to be in a location that did not favor any particular pilot's previous fleet experience and covered an area that had challenging terrain so that there was great possibility of misperceiving the surroundings. Finally, it needed to be an area adequately mapped in Falconview to use in our analysis. The mountainous area of Twentynine Palms was selected for this experiment for several reasons. The first was that the area includes some landmarks and there are multitudes of executable routes. Secondly, most of the subjects of the study had not operated in this area. Finally, we consider the high altitude desert terrain to be comparable to the current operating environments in Iraq and Afghanistan.

After choosing the operating area, routes were selected to support hypothesis testing. To collect sufficient data, four routes were generated, along with a practice route. The subject had to navigate through the route using a joystick which controlled heading (roll) only; subjects had no control over yaw, power, or airspeed. Subjects were able to control pitch, but the simulator was set up such that pitch changes did not have accompanying altitude and airspeed changes. The pilots did not have to control attitude, airspeed, rotor speed, or turn coordination (commonly referred to as "ball" due to the setup of the turn-and-slip indicator in real aircraft), therefore greatly simplifying the navigational

task. For the last two routes, a scripted "autopilot" guided the subjects along a set course without their control.

The auto-navigation routes were added to normalize the experiment in the following manner: if all of the subjects were able to control the helicopter through the routes, each pilot would see different terrain features because the probability of two independent pilots flying the same course is practically zero. Controlling the route with the autopilot allowed the experiment proctor to pause the route at the exact same points, so that each subject saw the same terrain. Moreover, it ensured that each pilot was presented the terrain identically—that he approached it from the same azimuth and roll angle. Four routes designed in this study can be found in Cowden et al. (4).

With the number and types of routes chosen, the waypoints comprising the route were selected. The routes needed to be fair, yet challenging enough for the pilots to get off-course. We subjectively created routes that were appropriate for a late first tour aviator's level of experience, approximately 750–1000 total hours. The practice route was designed to get the subject familiar with the control of the helicopter, feel comfortable using the confidence program, establish a scan pattern, and gain familiarity with the interfaces. The practice route was a short, four-waypoint route. This route was based off prominent landmarks, yet still required the pilot to make large heading changes.

Procedure

Subjects were introduced to the experimentation lab with an IRB approved welcome script that notified the subjects of the focus of the study, a brief overview of what would be expected out of them, rules of the lab, and the voluntary nature of the study. The subjects were given an informed consent form to read and sign. After the informed consent, subjects were given a questionnaire relating to their flying experience and background. The background questionnaire included basic demographics, familiarity of the simulation-operating environment, experience with overland navigation, flight hours, and time since last flight. These data were collected to help group the subjects for analysis.

Once the subjects completed the background questionnaire, they were familiarized with the experiment, including the flight parameters of the helicopter, what was contained on the video screens, joystick control, and how to use the confidence feature of the simulation. Once they seemed comfortable with how the simulation would run, they were given a map of the practice route. This was an 8×11 map printed from Falconview. The map was a 1:50K TLM, just like the one that they would see on the monitor. This map was annotated, or "doghoused," with the waypoint number, distance in nautical miles, time to fly the leg at 60 kn, and total elapsed time. This paper map was only allowed during the map study and not during the flight portion of the simulation. The subject could only use the map on the monitor,

which included numbered waypoints, during the simulation. The subjects were given unlimited time to review the practice map before flying the simulation. The practice simulation was four waypoints long on an easy route. One of the main objectives of the practice route was to make the subject comfortable with the flight profile and monitor views, along with getting a solid grasp of using the confidence program. This route was paused roughly every 30 s for the subjects to point out their perceived location on the map and their confidence level. The subjects were given some navigation assistance from the proctor if they were lost. Once the subjects completed the route, they were asked if they were comfortable with the simulation and programs. They were allowed to have extra practice flying the helicopter if they were not comfortable.

After the completion of the practice route, the subjects were tasked with completing four navigation routes where data were collected. In the first two routes, the subjects were providing roll inputs while flying, whereas the last two routes were flown on autopilot. The first two routes began with a map study period of 3 min; in contrast the last two map study times were 2 min. Map study times were limited to provide increased difficulty by limiting the amount of headings and timings on the route, and to keep the experiment under an hour in duration. The last two map study times were less because the routes were shorter and the helicopter was on auto-navigation, reducing the task load on the subject. Before executing the auto-navigation route, the subjects were also given a scenario. In this scenario, the subject simulates flying with a new pilot in the squadron who is responsible for the navigating and flying. Both the new pilot and the subject must fly in an area where they have never been. The new pilot is supposed to follow the route, but there is a chance that he can get off-track. The scenario informs the subject that the intended route is not necessarily what the new pilot will fly. Once the map study was complete, the subject conducted the navigation portion of the simulation. During the first two simulations, the route was paused about every 40 s, which was not a hard number because the evaluator wanted to minimize pausing during turns. Pausing during turns can be disorienting to subjects and it was hard to remember the amount of bank they had after they finished the pause. During the second two simulations, the pause points were in the same location for each subject and happened between 20–40 s. Again, these pauses occurred during level flight. After the completion of each of the navigation routes, the subjects were given a post-task questionnaire. It questioned whether the subject felt they strayed off-course, misperceived terrain, and asked what they could have done differently to remain on-course.

Once all four routes were completed, the subjects were given one final questionnaire. This questionnaire covered topics on why they believe pilots get lost, what they do if they sense they are not on-course, and what they think their confidence level was during the experiment. This questionnaire allowed for subject grouping based on similar responses. The subjects were asked to add any additional comments and the evaluator asked other pertinent questions to give insights on why they misperceived terrain on the route and confidence levels.

Statistical Analyses

We used Spearman's rank correlation coefficient (ρ) to determine statistical dependence between variables for the first two hypothesis tests. Paired-sample *t*-tests were used. The significance level, α, was set at 0.05. Demographic variables were collected from the background questionnaire that included TFH, overland flight hours, participation in similar past experiments, and experience with low-level and desert low-level navigation.

There were two major dependent variables used for analysis; 1) confidence (CONF); and 2) the distance from the actual helicopter position to the perceived position. Pilots' confidence was self-reported using the confidence app; i.e., subjects rated their navigation confidence from 0 to 1 for each pause point. The CONF is defined as a confidence measurement between 0 and 1, where 0 indicates the lowest confidence and 1 the highest confidence. CONF_BIN is a variation of CONF constructed into a binary variable. The CONF_BIN is defined "high" if CONF $\geq$ 0.5 and "low" otherwise. The threshold of 0.5 was chosen for the CONF_BIN variable because it was the numerical midpoint of the CONF range. This midpoint was easily defined on the confidence app, making it a likely division between high confidence and low confidence. If a subject believed there was a good chance his perceived location was not close to the actual location, he would not choose a confidence level over 0.5.

The error in perceived location was derived from the great circle distance (7) between the actual latitude and longitude position of the aircraft and the subject's perceived latitude and longitude:

$$ERROR1 = R \arccos(\sin\varphi_a \ \sin\varphi_p +$$
$$\cos\varphi_a \ \cos\varphi_p \ \cos(\chi_a - \chi_p) \qquad \text{Eq. 1}$$

where ERROR1 = the great circle distance between perceived and actual location (km), R = Earth's radius at the Twentynine Palms, CA, area = 6372.8 km, ϕ_a = latitude of the actual aircraft position (radian), ϕ_p = latitude of the perceived aircraft position (radian), χ_a = longitude of the actual aircraft position (radian), and χ_p = longitude of the perceived aircraft position (radian).

Similarly, the second type of error that was calculated was the distance between where the subject perceived he was compared to the intended route of flight:

$$ERROR2 = R \arccos(\sin\varphi_i \ \sin\varphi_p +$$
$$\cos\varphi_i \ \cos\varphi_p \ \cos(\chi_i - \chi_p) \qquad \text{Eq. 2}$$

where ERROR2 = great circle distance between perceived and intended location (km), ϕ_i = latitude of the

intended aircraft position (radian), and χ_i = longitude of the intended aircraft position (radian).

The NAV is a variation of ERROR1 constructed into a binary variable, indicating whether the subject stayed within a certain threshold/boundary. Pilots were instructed to stay within 0.5 km of the route; we buffered this to be 0.75 km. The NAV was defined "On-track" if ERROR1 < 0.75 km and "Off-track" otherwise. The 0.75-km distance for obtaining the NAV variable was used because subjects were told prior to their navigation tasks that they should be confident in their perceived location if they were within 0.5 km of their actual location. The 0.75 km gave the subjects an additional 0.25 km error distance because it is difficult for pilots to recognize if they fell within the 0.5 km distance exactly while navigating. This additional error distance also helped to affirm, without any doubt, that the subject had the wrong perception of his location.

RESULTS

Hypothesis 1

We failed to reject the first null hypothesis, i.e., confidence and distance from subjects' perceived location to actual helicopter position (ERROR1) are not correlated for all four routes. Note that failing to reject the null hypothesis 1 is compatible with what we observed from the previous experiment. Failing to reject the null hypothesis for this case could be a causal factor for pilots getting off-track, along with the associated mishaps and mission failures.

Hypothesis 2

We can reject the second null hypothesis, i.e., confidence and distance from subject's perceived location to intended route of flight (ERROR2) are correlated for route 1 ($\rho = -0.65$, $P < 0.05$) and route 4 ($\rho = -0.60$, $P < 0.05$). This result means that the subject has high confidence when he believes he is close to the intended route for routes 1 and 4 regardless of his actual closeness. This result shows that there is evidence of biased visual perception favoring their intended location.

Hypothesis 3

There is a statistically significant difference in ERROR1 at the beginning of the route and ERROR1 at the end of the route for all four routes [$t(14) = -2.07$, $P < 0.05$; $t(14) = -3.15$, $P < 0.005$; $t(14) = -7.71$, $P < 0.001$;

$t(14) = -2.82$, $P < 0.01$]; the longer the subject flies, the more his perception error increases. This result can help pilots realize that they might want to reevaluate their perceived location the further along the route they are and reduce their corresponding confidence in their location.

Hypothesis 4

There is a statistically significant difference in CONF at the beginning of the route and CONF at the end of the route for all four routes [$t(14) = 3.11$, $P < 0.005$; $t(14) = 2.40$, $P < 0.05$; $t(14) = 2.31$, $P < 0.05$; $t(14) = 2.90$, $P < 0.005$]. This result suggests that the longer the subject navigates along a route, the lower his corresponding confidence becomes. This result follows along with hypothesis 3, that the perceived error appears to increase the longer the subject navigates. Pilots' CONF is reducing with an increasing perception error (ERROR1). Although there is no correlation between confidence and perception error, there is a trending effect of confidence getting lower further into the route while perception error is increasing.

The comprehensive analysis of the data related to the experiment output to the modified SDT matrix (10) for assessing navigation skills. **Table II** shows experimental data of the confidence versus navigational error using the CONF_BIN and NAV variables. Table II shows that 58.3% of the time, subjects were on-track and had a correspondingly high confidence level. This table also shows that only 7.0% of time, pilots had low confidence yet were actually on-track. These percentages reflect that the subjects were highly unlikely to misperceive their location when on-track, but the problem arose when the subjects were off-track. Subjects were off-track, yet still highly confident 27.0% of the time during the navigation. This indicates subjects were highly confident about their navigation performance 77.9% of the time when they were off-track. The misperception error is about 3.5 times greater than correct perception when a pilot is off track. This result relates to the dangerous section of the matrix where pilots are lost and do not know it, and this is the second largest navigational state of the experiment among four navigational states. It is also in this area where mission failure and mishaps occur due to incorrect navigation.

Table III shows that the confidence and correctness for each route align with the overall breakout. Most notable is that the auto navigation routes (3,4) had a lower

---

TABLE II. MATRIX OF EXPERIMENTAL NAVIGATION PERFORMANCE RELATING CONF_BIN AND NAV.

Assessing Navigation Performance	CONF_BIN: Low	CONF_BIN: High	Grand Total
NAV: Off-track	7.7% (22.1%)	**27.0% (77.9%)**	34.7% (100%)
NAV: On-track	7.0% (10.7%)	58.3% (89.3%)	65.3% (100%)
Grand Total	14.7%	85.3%	100%

CONF_BIN: a variation of confidence constructed into a binary variable; NAV: a variation of ERROR1 constructed into a binary variable. CONF_BIN = "high" if CONF ≥ 0.5 and "low" otherwise. NAV = "On-track" if ERROR1 < 0.75 km and "Off-track" otherwise. Percentages in parentheses are calculated based on Off-track or On-track NAV, respectively. Of interest is the bolded area, corresponding to off-track and high confidence, or the "dangerous" quadrant

TABLE III. MATRIX OF EXPERIMENTAL NAVIGATION PERFORMANCE FOR EACH ROUTE COMPARING NAVIGATION PERFORMANCE AND CONFIDENCE.

Assessing Navigation Performance		Route				
		1	2	3	4	Average
NAV and CONF_BIN	Off-track	26.6%	28.4%	35.6%	34.8%	30.9%
	High Confidence	19.5% (73.3%)	21.3% (75.0%)	31.7% (88.9%)	25.8% (74.2%)	24.2% (78.2%)
	Low Confidence	7.1% (26.7%)	7.1% (25.0%)	4.0% (11.1%)	9.0% (25.8%)	6.7% (21.8%)
	On-track	73.5%	71.7%	64.4%	65.2%	69.1%
	High Confidence	65.5% (89.2%)	65.4% (91.2%)	55.5% (86.2%)	56.2% (86.2%)	61.2% (88.5%)
	Low Confidence	8.0% (10.8%)	6.3% (8.8%)	8.9% (13.8%)	9.0% (13.8%)	7.9% (11.4%)
	Total	100.0%	100.0%	100.0%	100.0%	100.0%

CONF_BIN: a variation of confidence constructed into a binary variable; NAV: a variation of ERROR1 constructed into a binary variable.
CONF_BIN = "high" if CONF ≥ 0.5 and "low" otherwise. NAV = "On-track" if ERROR1 < 0.75 km and "Off-track" otherwise.

percentage of route correctness, 73.5% and 71.7% for routes 1 and 2 versus 64.4% and 65.2% for routes 3 and 4, respectively. The subjects misperceive their location more frequently when control inputs were not required from them. Some explanations for this could be due to complacency, and/or experiment fatigue. During the auto navigation routes, subjects seemed to be more relaxed during the navigation and map study. Subjects were less likely to be actively tracking the course, which lead them to believe that the aircraft was heading on course. This type of complacency is common in multipiloted aircraft and can be attributed to mishaps. Also noteworthy is the fact that route 3 had the highest percentage of time in the "dangerous" quadrant. Order effects may explain why route 3 took more time than the other routes. The "dangerous" quadrant of route 3 could be larger than that of route 1 and 2 because route 3 was the first time the subject dealt with auto-navigation. Additionally, route 3 could also have taken more time than route 4 because the pilot realized at the end of route 3 that the auto-navigation did not follow the intended route of flight, making the confidence on route 4 less than 3. This would correspond to a lower amount of time in the "dangerous" quadrant. The auto-navigation segments are important because they represent the situation of a junior pilot flying a route with an experienced, senior (and presumably, better) navigator.

At the completion of the navigation and debriefing portion of the experiment, the participants were given a post-task questionnaire. This questionnaire was written to answer two questions. The first was to obtain navigation techniques that the more successful pilots used, while the latter was an attempt to normalize confidence levels. In the attempt to normalize the confidence levels, some notable outcomes arose. The first was that only one participant felt that pilots were not over reliant on navigation equipment like GPS, with six neutral responses and eight positive. The second result of the questionnaire is that 12 of the 15 participants thought that it was easy to misinterpret terrain during overland navigation, with the other 3 responses being neutral. The last questionnaire output was the most alarming. Only two participants (13.3%) believed that they were overconfident in their navigation skills. This is surprising considering the percentage of time the participants were in the "dangerous" quadrant of flight. When the participants were off-track, they had a high confidence, or wrong perception. This suggests that pilots are misperceiving their overconfidence during navigation.

Fig. 1 shows three types of errors that can be defined in the corresponding navigation scenario. ERRORpa = ERROR1 as defined in Eq. 1, ERRORpi = ERROR2 as defined in Eq. 2, and ERRORai is a distance between actual and intended location. We ran post hoc analysis on a correlation between pilot confidence and ERRORai, which did not show a statistical significance as between ERROR1 and pilot perception. This result again supported that pilot confidence is not correlated with actual and perceived performance of the pilot.

Fig. 2 shows data for each error type from all four routes. ERRORpa (= ERROR1) is larger than ERRORpi (= ERROR2) for all four routes [$t(14) = 8.64$, $P < 0.001$; $t(14) = 7.93$, $P < 0.001$; $t(14) = 5.60$, $P < 0.001$; $t(14) = 3.95$, $P < 0.001$]. ERRORai was larger than ERROR2 for all four routes too [$t(14) = 4.88$, $P < 0.001$; $t(14) = 5.15$, $P < 0.001$; $t(14) = 6.69$, $P < 0.001$; $t(14) = 3.56$, $P < 0.001$]. ERRORai was smaller than ERROR1 for routes 1, 2, and 4 [$t(14) = 4.94$, $P < 0.001$; $t(14) = 5.38$, $P < 0.001$; $t(14) = 2.82$, $P < 0.005$]. The analysis indicates that pilots tend to overestimate their performance in general, and the bias is toward their expectation.

We can confirm the bias on the navigation map showing all three locations, i.e., the helicopter's actual trajectory, pilots' perceived location, and intended route. **Fig. 3** shows representative misperception that 15 pilots experienced during the navigation. The thick black line indicates the helicopter's actual trajectory, diamonds
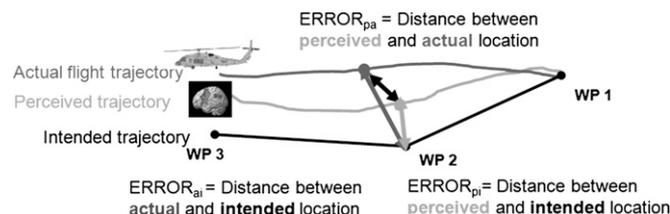
Fig. 1. Three types of errors, i.e., ERRORpa, ERRORpi, and ERRORai. ERRORpa = ERROR1 and ERRORpi = ERROR2.
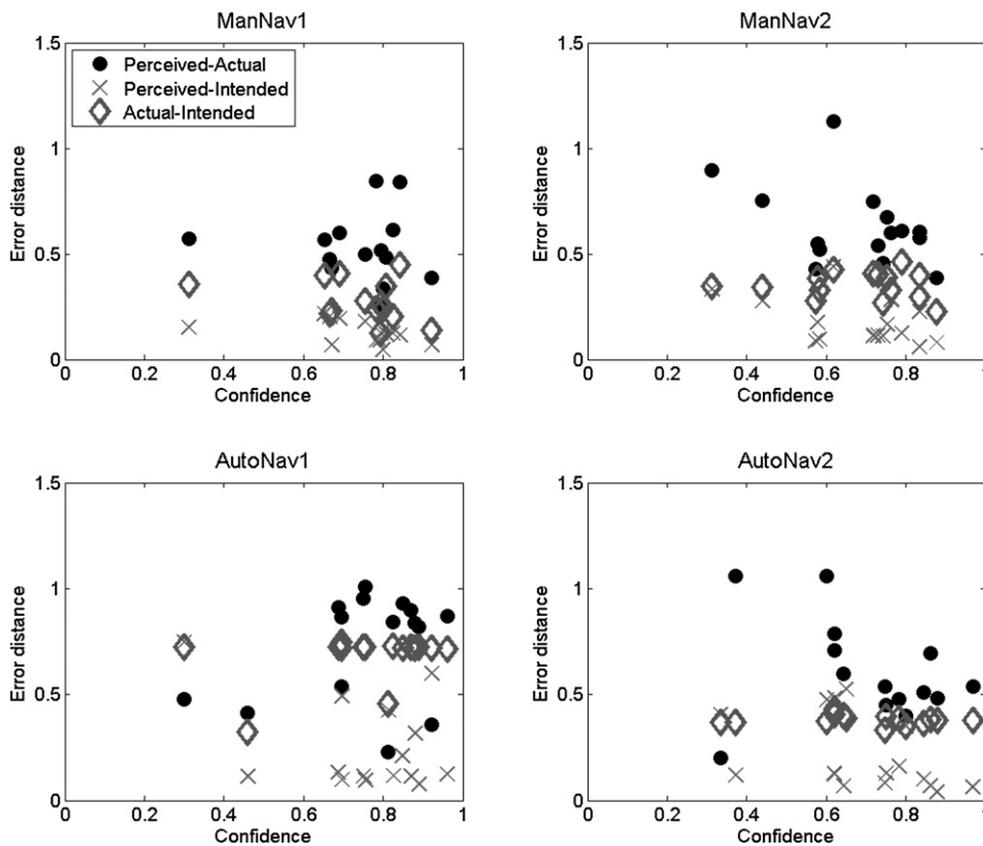
Fig. 2. Comparison of the three errors for each route. Note that ERRORpa (black circle) is the biggest and ERRORpi (cross) is the smallest of all.

correspond to each pause point, and matching shaded circles represent pilots' perceived locations. Thin black lines with connecting circles are planned/intended routes. It is clear that pilots' perceived locations mostly do not match with actual locations. Instead, perception matches with the intended route.

As an exploratory analysis, we also examined whether there was any correlation between pilots who had experience with low-level desert navigation and those who had very little to none. There were seven subjects who had low-level desert navigation experience and eight who did not. The Student t-test revealed a statistical significance [$t(7) = 2.68$, $P < 0.05$] in Route 1 on ERROR2 only. The result showed that pilots who were less familiar with the task tended to have more bias toward the intended route. However, we were not able to find the statistical significance in the other routes and expertise effects should be examined further.

This experiment had three participants who had previously conducted a similar experiment. These participants had seen a route in Sullivan et al. (12), so they could have an advantage over other participants during the two auto-navigation routes that were based off the route. We found no learning effect in the participants.

Regarding scenario (route) differences, there was statistically significantly different CONF between some of the routes. The difference in CONF is expected, because the routes were set up to have varying difficulties. Route 1 was supposed to be harder than route 2, causing the

CONF to be lower in route 2 than route 1. With a harder route, there are more chances for the participant to get off-track, thereby reducing their CONF level. Route 3 and 4 were set up to be similar, but there is a large difference in the data. The participants' realization at the end of route 3 that the autopilot did not follow the intended route may have caused the difference. This realization may have caused the participant to be less confident in the location of route 4. The data shows that there was not a trending effect of increased or reduced confidence throughout the experiment. Regarding ERROR1 and ERROR2, there were no scenario differences found.

Power analysis (3) was conducted for the significant correlation coefficients for routes 1 and 4 of hypothesis 2. The power ranges between 0.78 and 0.86 given a sample size = 15, $\alpha = 0.05$, and observed $\rho = 0.60 \sim 0.65$. This power is high considering the small sample size of the experiment, meaning that pilot bias toward the intended route is likely.

DISCUSSION

In this paper we have demonstrated a methodology for determining the relationship between pilot perception and performance and conducted an initial study. No correlation was found between perceived and actual location of the aircraft, nor between confidence level and actual location. There is, however, some evidence that there is a negative correlation between perceived location and intended route of flight, suggesting that

Fig. 3. Obvious pilot misperception showing a consistent bias. Actual helicopter trajectory (thick black line), pause point to let pilots mark their perceived location and confidence (diamonds), and pilots' perceived locations (shaded circles) from this human-in-the-loop experiment.

there is a bias toward believing one is on the intended flight route.

As helicopter missions, both civil and military, are frequently held at low altitude over varying terrain, the importance of visual navigation, even augmented by GPS and other systems, will not be diminished in the near-term. Military aircraft frequently fly over rugged terrain to find targets (or survivors). Civil aircraft, such as "Life Flight" helicopters rely on visual navigation to find and recover their patients.

In our initial exercise of the model, we found that there is no significant relationship between a pilot's confidence and accuracy in navigation. Surprisingly, we found that there is also no significant relationship between a pilot's experience and the accuracy of his navigational self-assessment. It is the second finding that we have the most interest in, and our immediate recommendation across rotary-wing aviation is to include syllabus events for both initial and refresher pilots where they test their navigation skills in a simulator, and then are shown the correctness of their route. This would serve to demonstrate to pilots that their self-assessments may not be accurate, with the intended goal of making their assessments more conservative. It would also demonstrate to leadership that the problem of confidence vs. correctness in navigation is not just a 'nugget' (junior pilot) problem, but is prevalent across all ranks.

Our experiment was conducted with only a PAC given the single task of navigating in an overland environment with control over only the aircraft's heading.

As such, it was impossible to crash the simulator. In future experiments, we propose to give the subjects control over all three axes of flight, as well as incorporating mission elements such as radio traffic or specified "time on target." Finally, it would be informative and important to see how a crew of subjects, consisting of PAC, PNAC, and crewmembers would perform, specifically with the research question of what mix of experience levels results in the best performance.

These next steps are currently being developed in our own laboratory, but a more complete solution will involve a partnership with military and civil training centers using full-motion simulators. Direct implementation of the results from this experiment to new procedures and technologies is difficult because it involves personal confidence. The most important result from the experiment is that pilots need to be trained to recognize that confidence does not correspond to correctness during navigation. A single simulator event, possibly conducted simultaneously with an existing syllabus event could be implemented into Naval Aviation training; we feel that the most appropriate place would be graduate-level training, held at the Fleet Reserve Squadrons along with the corresponding Army and Air Force helicopter training schools. Results from this experiment could also be added to aviation physiology and safety center documents.

Currently GPS is considered a supplemental navigation device and is not required to execute overland low-level navigation. This experiment suggests that it is not uncommon for pilots to misperceive their location just using visual navigation; the amount to which this is alleviated by GPS is an area for future research.

To enhance the results of this experiment a larger sample size spanning different experience, including days since last type of flight, commonly known as "currency" and Type/Model/Series, commonly referred to as communities, could be used. The larger sample size would allow for a better experience grouping of participants (expert, intermediate, and novice). Being able to effectively group the participants could provide insights into an "overconfident" or "dangerous" population. This could pinpoint where dedicated time and technology need to be spent. The experiment could also be conducted under realistic operation environments. These environments could be nighttime, emergencies, and different weather conditions. Again this would enhance the data for real-world operations.

Our experiment places a live subject with some aviation experience into a stationary (i.e., nonmotion) flight simulator with movable cyclic and collective flight controls. It builds on, but is substantially different from the experiment of Adam et al. (1) in two important ways: first, the cognitive task requirement is similar to aviation; and second, navigation performance is the measure of interest. Casner (2) studies navigation error with GPS in the IFR flying regime. Live aircraft operations are possible in IFR flight, which is typically positively controlled and conducted at altitude. We do not believe that an adequate safety margin exists to conduct such a study at a low level.

Our experiment uses a simulated aircraft held at a constant altitude of 200 ft (~61 m) AGL and 60 kn indicated airspeed with calm true winds. The simulated aircraft does not experience any emergencies and this is known to the test subjects. Finally, there is no crew coordination, either within the aircraft between crewmembers, across aircraft in a flight formation, or with external agencies or controllers. Our experiment, therefore, removes three important distractions present in real world low-altitude navigation. Visual navigation is affected by nighttime, where shading and cultural lighting may make the terrain appear differently than expected. Aids to night vision, such as forward-looking infra red or night vision devices ("Goggles"), may make navigation easier or more difficult, depending on lighting, shading, and other effects.

Authors and affiliations: Ji Hyun Yang, Ph.D., Department of Automotive Engineering, Kookmin University, Seoul, Korea; and Bradley T. Cowden, M.S., Quinn Kennedy, Ph.D., Harrison C. Schramm, M.S., and Joseph A. Sullivan, USN CDR, Ph.D., Naval Postgraduate School, Monterey, CA.

REFERENCES

1. Adam H, Deaton J, Hansrote R, Shaikh M. An analysis of GPS usability and human performance in aviation. International Journal of Applied Aviation Studies 2004; 4:107–32.
2. Casner SM. Flying IFR with GPS: how much practice is needed? International Journal of Applied Aviation Studies 2004; 4:81–97.
3. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
4. Cowden B, Yang JH, Kennedy Q, Schramm H, Sullivan J. Helicopter pilot misperception and confidence during overland navigation. AIAA Modeling and Simulation Technologies Conference; August 13-16; Minneapolis, MN. Reston, VA: AIAA; 2012.
5. de Voogt AJ, Uitdewilligen S, Eremenko N. Safety in high-risk helicopter operations: the role of additional crew in accident prevention. Saf Sci 2009; 47:717–21.
6. Eschenbach R. GPS applications in general aviation. In: Parkinson B, Spilker Jr J, Axelrad P, Enge P, eds. Global positioning system: theory and applications, volume II. Washington, DC: The American Institute of Aeronautics and Astronautics, Inc.; 2004:375–95.
7. Gellert W, Gottwald S, Hellwich M, Kästner H, and Küstner H, eds. The VNR concise encyclopedia of mathematics. New York: Van Nostrand Reinhold; 1989.
8. Naval Safety Center. Aviation summaries by date. Norfolk, VA: Naval Safety Center; March 2013. Report No.:AV-201.
9. Simons D, Chabris C. What smart chess players and stupid criminals have in common. In: Chabris C, Simons D. The invisible gorilla: how our intuitions deceive us. New York: Three Rivers Press; 2010:80–115.
10. Stanski-Pacis G, de Voogt A. The role of GPS in aviation incidents and accidents. In: de Voogt A, D'Oliviera T, eds. Mechanisms in the chain of safety: research and operational experiences in aviation psychology. Aldershot, UK: Ashgate; 2012:117–130.
11. Sullivan JA. Assessing neurophysiologic markers for training and simulation to develop expertise in complex cognitive tasks [Dissertation]. Monterey, CA: Naval Postgraduate School; 2010.
12. Sullivan J, Yang JH, Day M, Kennedy Q. Training simulation for helicopter navigation by characterizing visual scan patterns. Aviat Space Environ Med 2011; 82:871–8.
13. Yang JH, Kennedy Q, Sullivan J, Day M. Bayesian modeling of pilot belief and visual misperception in helicopter overland navigation, Proceedings of IEEE International Conference on Systems, Man and Cybernetics; October 9-12; Anchorage, AK. New York: IEEE; 2011.