Seeing the Big Picture: Pilot Assessments of Cockpit System Interactions Contribution to Situation Awareness

David R. Meyer, Christina F. Rusnock(✉), and Michael E. Miller

Department of Systems Engineering and Management, Air Force Institute of Technology, Wright-Patterson AFB, OH 45433, USA
christina.rusnock@afit.edu

Abstract. Pilots build and maintain situation awareness based on their interaction with the world around them. This interaction includes a complex and dynamic series of tasks including running checklists, reading instruments and displays, looking out of the window, or listening to the radios. However, since the quality and quantity of information derived from each interaction is not well known, cockpit designers can only have an abstract understanding of how much situation awareness their system imparts to the human operators. This paper examines the opinions of pilots regarding how performing cockpit tasks contribute towards gaining situation awareness. Twenty-one military aviators were asked to rate 19 generic cockpit tasks based on how they contribute to or degrade situation awareness. This research shows that modern avionics, such as the Heads-Up Display, Multi-Function Display, and other sensors can provide strong positive situation awareness, but depending on the mission phase and other factors, they may not be significantly more advantageous than their analog counterparts.

Keywords: Situation awareness · Aircrew · Pilot · Cockpit design · Human-machine interaction

1 Introduction

The ability to conceive of the aircraft's whereabouts, status, weather, fuel state, terrain, and, in combat, enemy disposition is critical to effective aircraft operation. In critical phases of flight, in poor weather, or in the face of systems malfunctions, this ability can mean the difference between mission success and failure or even aircrew survivability [1]. Operator performance in complex or dynamic environments is often a function of situation awareness (SA) [2]. Aircrews spend considerable portions of time and effort developing and maintaining SA, especially in evolving environments [3]. Indeed the military has had a keen interest in SA in the cockpit dating back to World War I and especially more so in the information age of the late 20th century [4–6]. In tactical situations, this includes knowledge about the locations, actions, and capabilities of both friendly and enemy forces [3]. A leading cause of military and 88 % of commercial aviation accidents have been attributed to poor SA [7, 8]. If SA is essential to mission

D. Harris (Ed.): EPCE 2016, LNAI 9736, pp. 419–428, 2016.
DOI: 10.1007/978-3-319-40030-3_41

accomplishment, then harnessing the pilot interaction with the cockpit controls and displays which aid the collection of key information is a driving goal in both engineering design and operator training. This paper explores the operator's perspective on how cockpit tasks affect SA.

2 Background

Situation awareness (SA) is the knowledge of environmental factors that influence decisions and depends on the operator's internal perceptual model of the world [9]. More specifically, SA has been defined as the operator's perception (or mental model) of elements in the environment around them within a volume of space and time, the comprehension of their meaning, and the projection of their status into the future [3, 10]. Non-routine and unpredictable situations, such as those that can occur during emergency situations or combat, demand effective integration of large quantities of information to be processed with limited cognitive capacity [2, 4]. Situation awareness drives decision making based on the operator's mental model of the environment. The process of gaining situation awareness, while independent from the process of decision making, forms a basis for this mental model [4, 11]. The performance of a pilot is highly related to their ability to fuse information to augment SA and drive decision-making [12]. The resulting decision leads to an action, which alters the environment and situation and the cycle begins again.

No current criteria exist to determine the required level of SA needed to obtain a desired level of performance [13]. Less than perfect SA causes the operator to assume a level of risk for error as they must make a decision without complete information. Good SA increases the probability of good performance, but does not guarantee high performance as factors, such as excessive workload, poor decision making, or poor task execution can degrade performance in the presence of good SA [11]. Because of the critical effects situation awareness has on mission outcome and survivability, designers have a vested interest in providing the operator with appropriate resources, and operators have a vested interest in maintaining a high level of SA [1].

Designers can incorporate heads-up displays (HUD), multi-function displays (MFD), automation, expert systems, advanced avionics, and sensors to provide more information in a more useful manner [10]. In this regard, for a given task, all cockpits are not created equal. Cockpit design affects the number of required tasks the operator must perform, the workload of those tasks, and the information provided to the operator during their completion [4]. Automation and intelligent cockpits can tailor information to match the operator's needs or perform support functions based on the required situation [2]. Operators develop procedures to scan information sources and train to maximize the use of all available tools and observations to increase SA.

Traditional methods for measuring SA often probe the operator's awareness of critical facts within an environment or require the operator or an observer to rate the operator's SA when using a system or simulator [14]. However, these methods do not readily provide insight into the utility of any single device in the cockpit for improving SA. Evaluations of the impact of individual display or display formats on SA are limited within the human factors research, but often include evaluating a system in which a single display is modified within the cockpit [15].

It might be reasonable to assume that by increasing the availability of information in the cockpit, SA will be improved. However, the literature and experience demonstrate that this assumption is not true. For example:

- Cues in the normal field of view can be missed if they are too subtle or if the operator is not paying sufficient attention [16], yet most pilots would immediately perceive a flashing "Master Warning" light, especially if it was accompanied by an audible warning tone. However, frequent or nuisance warnings desensitize operators, thus negating the effectiveness of those warnings [17, 18].
- As cockpits evolve, new "black boxes" are added: autopilot, radar, weapons systems, mission sensors, navigational aids, computers, etc. While many of the additions undeniably present more information to the aircrew and have the potential to solve specific problems, they each present another interface to the human operator, increasing workload and decreasing the time available to scan each device.
- Information overload is a significant concern in modern cockpits [2, 19]. SA is inherently a function of the quantity and quality of data, however the format of image presentation determines the availability and the likelihood of use of that data. Poorly designed and integrated interfaces increase workload and time needed to synthesize information and open opportunities for confusion and errors [2]. Conversely, well-designed interfaces can reduce workload. However, even the best cockpit displays and controls cannot overcome human performance limits [20], including the human operator's capacity for attention of 2−40 bits per second; far lower than the presentation rate of the environment and modern cockpit systems [21].
- Because glass cockpits rely much more on integrating display layers and inputs, primarily through a Flight Management System (FMS). These systems provide more data in a richer format. However, they also require paging through multiple display layers. As such, interaction now requires a keyboard for instrument manipulation, therefore, accessing a function that once required twisting a dial or a couple of switches is more complex—with an associated increase in workload and time delay.
- By reducing the number of places a pilot is required to scan within the cockpit, attention resources available to the remaining scan areas can be increased. However, if the pilot neglects to maintain a scan elsewhere and becomes dependent or fixated on certain information sources (e.g., the HUD), his or her SA will remain limited [22–24]. Dependency on automation elements may cause fixation to the neglect of other duties.
- Automation can increase feedback or provide new methods of communicating feedback, which may alter how an operator must assimilate that data [25]. If performed well, this can aid the operator in comprehending the meaning of items in the environment and projecting future states; however, if performed poorly, critical information may be obscured.

Each of these examples illustrates that each information source in the cockpit may have a significant influence (positive or negative) on SA. Therefore, there is a need for a method to allow one to understand the effect of each of these information sources on the operator's SA, allowing one to assess the importance of each display and perhaps each information representation. As indicated before, however, the presence of a display does not aid SA. Instead SA is gained as the operator actively interacts with the system, gaining information through this interaction. Since the cockpit and situation awareness are inextricably linked, this paper seeks to answer the question “how do various cockpit tasks contribute to aircrew SA?”

3 Methodology

Meyer [1] distilled over 1200 different cockpit tasks found in typical C-130 airlift missions into 19 generic cockpit tasks representing human-machine interaction and non-machine interaction tasks. This list of tasks is provided in Table 1. These generic tasks can easily be applied as an analogy to the majority of cockpit tasks found in any aircraft. These generic tasks can then be classified as machine or non-machine interaction tasks; where the human either interacts with the machine or with other humans. Some generic tasks, specifically talking and listening tasks were difficult to classify. For instance, if these tasks were conducted via radio or aircraft but the communication was human-human interaction with electronic transmission means, the tasks were classified as non-machine interaction tasks. However, aircraft generated advisories (listening, simple) and monitoring the Radar Warning Receiver (RWR), which permit interaction between the human and the machine could be classified as machine interaction tasks. However, they generally require natural interfaces as would be employed when interacting with another human and are therefore classified as non-machine interaction tasks.

Table 1. Generic cockpit tasks

Machine interaction tasks	Non-machine interaction tasks
Reading instrument or gauge	Looking out of window
Reading MFD/moving map/digital display	Writing (data cards, kneeboard, etc.)
Viewing Head’s Up Display (HUD)	Reading charts, “sticks,” approach plates
Reading raw computer data	Manual computations (flight computer, TOLD, etc.)
Radar/sensor interpretation	Talking, simple (advisory calls, responses)
Keyboard/data entry	Talking, complex (briefings, radio calls, etc.)
Simple maneuvering (maintaining parameters)	Listening, simple (alerts, advisory call)
Complex maneuvering (defensive reactions)	Listening, complex (radio, crew feedback)
Simple button/switch actuation	Background listening (monitoring RWR, radio)
Cumbersome button/switch actuation	

In order to quantify the relative contributions of each individual task to tactical SA, a survey was given to 21 aviators representing pilots and non-pilot operators of various military aircraft. The number of respondents in each aircraft type and crew position are shown in Table 2. Several of the participants had flown on multiple aircraft (hence aircraft count is greater than 21). C-130 aircrew and pilots were intentionally over-represented because Meyer [1] originally focused on the C-130 aircraft.

Table 2. Survey representation

Aircraft	*n*	Aircraft	*n*	Crew	*n*
AC-130 W	1	MC-130H	11	Pilot	12
C-130H	12	MC-130 J	1	Navigator	3
C-130 J	4	MH-65D	1	Electronic Warfare Officer	1
CV-22A	1	T-1	1	Flight Engineer	4
F-15E	2	U-28A	1	Special Mission Aviator	1

Respondents were asked to rate the relative impact of each task on SA. When performing this rating, participants were asked to provide a rating between −3 and 3. Participants were to provide this rating against a hypothetical baseline of flying with one's eyes closed, which was to receive a rating of 0. Note that negative numbers indicate that the tasks detract from SA, zero equals no effect of a task on SA, and positive values indicate that performing the task contributes to SA. Participants were asked to disregard the effects of workload on their SA scores and provide any remarks that they felt were appropriate. Because it is expected that SA demands change as a function of mission phase, these ratings were made for three different mission phases: cruise flight, formation airdrop, and a maximum effort landing.

4 Results

Since not every survey participant was familiar with every mission phase (i.e. F-15E pilots do not perform airdrop), or are in suitably equipped aircraft, each task did not necessarily receive 21 responses. Most notably, tasks involving the head's-up display (HUD) during the airdrop and maximum effort landing mission phases was limited to C-130 J and MC-130 J participants (n = 4). However, the standard deviation for the HUD tasks ($\sigma = 0.52, 0.96, 0.00$) was normal compared to standard deviations of other tasks with 14-21 respondents. Indeed standard deviations were relatively large due to the integer scoring and small sample size (σaverage = 1.33). Noticeably, tasks that induced positive SA tended to have lower standard deviations than negative SA tasks as shown in Table 4.

By far, the strongest sentiment in participant feedback was, "it depends." What is clearly evident from Tables 3 and 4 is that the situation itself is a strong independent variable. Objective area events, such as airdrop and airland, produced slightly stronger opinions than in cruise flight. Tasks that related to visual cues, such as looking out of the window or HUD have stronger SA scores when the aircraft was lower to the ground and when those visual cues provide richer and more detailed information. The more cognitively demanding tasks of keyboard entry, complex talking (i.e. giving a briefing), writing, and manual computations produced stronger SA penalties during objective area events. Human-machine interaction tasks that provided strong positive SA tended to receive even higher ratings during objective area events; non-machine interaction tasks that incurred strong SA penalties received lower ratings during objective area events (i.e. events near targets, drop zones, and landing zones). Conversely, human-machine interaction SA penalties did not get appreciably worse during objective area events (except keyboard entry), while non-machine interaction tasks with strong SA scores did not improve noticeably during objective area events (except for looking out of the window).

When viewing human-machine interaction tasks (Table 3), a strong preference emerges from four tasks: viewing the HUD, reading a multifunction display (MFD) or other similar moving map or digital display, reading instruments and gauges, and radar/sensor interpretation. The HUD scored moderately well during cruise flight, but the SA value increased remarkably during airdrop and airland, receiving the only scores above 2.0. Clearly the relevance of HUD information, whether by design or its relevance, increases dramatically during these two objective area events where aircraft are lower to the ground. The MFD SA averages were consistently high across all phases of flight. The quantity, quality, and relevance of this data does not seem to depend on phase of flight as much as it depends on crew position (flight engineers and some navigators consistently gave it lower SA scores, while pilots gave considerably higher SA scores to the MFD).

Table 3. SA scores, human-machine interaction tasks

Machine interaction tasks	Averages			Standard deviation		
	Cruise	Airdrop	Airland	Cruise	Airdrop	Airland
Viewing Head's Up Display (HUD)	1.33	2.25	3.00	0.52	0.96	0.00
Reading MFD/moving map	1.95	1.93	1.93	1.02	1.14	1.49
Reading instrument or gauge	1.35	1.08	1.62	0.88	1.19	1.33
Radar/sensor interpretation	1.48	1.29	1.14	0.93	1.44	1.46
Simple maneuvering	0.00	0.00	0.15	0.97	1.08	0.99
Reading raw computer data	−0.05	0.07	0.00	1.47	1.27	1.66
Simple button/switch actuation	−0.10	−0.14	−0.29	0.89	0.86	0.91
Complex maneuvering	−1.00	−0.86	−0.79	1.22	1.83	1.81
Cumbersome button/switch actuation	−0.86	−0.93	−0.93	1.39	1.54	1.77
Keyboard/data entry	−0.80	−1.00	−1.00	1.24	1.58	1.63

Surprisingly, instrument/gauge reading and radar/sensor interpretation did not exhibit clearly defined trends when viewed as averages. Reading an instrument infers reading a single component and gaining few elements of information. Compared to digital displays such as a HUD or MFD (which superimpose multiple data sources into the pilot's field of view or onto a moving map display), the gauges–by design–present far less data. However, it is surprising that in cruise flight, the SA value of raw instrument reading is nearly identical to viewing the HUD. Furthermore, raw instrument data scored lowest during airdrop and highest during airland. These trends can partially be explained when considering the non-C-130 aircrew did not answer the airdrop and airland portions. Pilots had stronger opinions of the value of instrument reading than navigators (and even some flight engineers), and pilots made up the bulk of non-C-130 participants that were represented only in the cruise flight category. This same trend was also partially evident in the radar/sensor interpretation task, but several participants indicated that it was of less value during the objective area than cruise flight. It is unclear if this is caused by radar/sensor information being degraded at low altitude, the information having less relevance during those phases, or a preference for other tasks, such as looking outside the window or HUD.

Similar to the HUD values shown in Table 3, looking out of the window increased in value during objective area events as shown in Table 4. In cruise flight, the window had similar SA values to the HUD, despite the HUD providing extra data. While pilots naturally have a strong SA value derived from looking outside, those situated farther from the cockpit windows did not feel as strongly. Naturally, the one respondent who stated, "I can barely see out the window from where I sit," rated the value of looking out of the window as marginal. Using this task as an example, one pilot emphasized that not only are phases of flight an independent variable, but specific situational elements add a host of other variables. For example, when looking out of the window, "night or low illumination will lower SA. If flying in formation through the weather, SA goes down, but looking through the window in daylight with good visibility in the mountains will greatly increase SA."

Table 4. SA scores, non-machine interaction tasks

Non-machine interaction	Averages			Standard deviation		
	Cruise	Airdrop	Airland	Cruise	Airdrop	Airland
Looking out of window	1.33	1.79	1.93	1.46	1.42	1.27
Listening, simple	0.90	0.50	0.86	1.00	1.29	1.29
Reading charts, approach plates	1.14	0.57	0.36	1.28	1.74	1.69
Listening, complex	0.81	0.14	0.14	1.33	1.75	1.83
Talking, simple	0.38	0.14	0.29	1.16	1.03	1.27
Background listening	−0.14	−0.29	−0.21	1.42	1.38	1.53
Talking, complex	−0.14	−0.64	−0.57	1.56	1.78	1.79
Writing	−0.19	−0.57	−0.93	1.12	1.28	1.21
Manual computations	−0.90	−1.15	−1.08	1.62	1.77	1.85

Naturally, reading a chart, map, or approach diagram will provide the operator with SA. Participants also showed a preference for simple listening tasks (i.e. "too low" or "5 knots fast") over more complex listening. Complex listening provided more SA during cruise flight than during objective area operations, which reinforces operational practices to conduct crew briefings during cruise portions of flight where possible. However, it is indeterminate if this trend has to do with typically lower attentional demands during cruise that allows for more complex listening, or because cruise flight grants more relevance, quantity, or quality of that data. One crew member indicated the value added depends on the information that was being conveyed: "crew feedback is good for SA, background noise is not." Naturally, listening to static, or non-relevant radio chatter might produce little or no SA, but to the electronic warfare officer, background listening through the radar warning receiver was a significant exception to that rule.

In reporting SA scores, some participants provided negative SA scores, which we hypothesize represents an opportunity cost that most likely trades off tactical and strategic SA. Performing task A (tactical task) consumes attentional resources that may have been spent on a hypothetical task B (strategic task) that would have yielded more SA. Thus, despite survey instructions, it is difficult to cognitively separate SA lost to higher workload (or to opportunity cost) from truly negative SA. In reality, a negative SA score would imply that by accomplishing a task, erroneous SA would be gained leading to an incorrect mental model.

Participant opinion appears to reflect this. One noted that the "largest loss of SA in terminal events is due to time constraints: time must be split between [being an] active participant in [the] event and other, coordination tasks." While many participants generalized anecdotal evidence that some tasks add SA and some tasks taketh away, others lament that certain cockpit tasks, such as the complex manual computations required for determining aircraft performance data "can potentially take me out of the game almost completely." Keyboard entry is considerably more complicated than simple switch actuation, and it also requires more attention devoted to ensuring that the keyboard entry is performed correctly. Furthermore, improper syntax, hitting a wrong button, and other errors create more work and perhaps distract the operator from performing other SA-generating tasks.

Perhaps what is most telling about operator opinion regarding the SA/workload tradeoff is summed up by one pilot's opinion that, "rating is a balancing act between SA gained by the task and the SA lost by time spent completing the act." Thus, the SA value of a heads-down multifunction display (MFD) was offset by the need to take the pilot's eyes out of the window or HUD to consume that information. Reading raw computer data (data displayed in number or tabular form and not integrated into an MFD, HUD, or other visual display) did not yield a positive or negative SA value, potentially because of this tradeoff.

5 Future Research

While this paper explores the concept of task-dependent SA, further research is required. This survey heavily samples C-130 and other tactical aircrew. A much larger survey should be conducted that incorporates more aircraft (with and without HUD) across the spectrum of flight. More mission sets can be added including tactical and non-tactical situations at both high and low altitudes. Within each mission phase, more resolution is needed. Many participants queried "what part of airdrop/airland are you talking about?" This infers that even within the particular mission phase, SA values could change. If, for example, viewing the HUD increases value from cruise to airland missions, how much does it increase from initial approach to the actual landing?

Furthermore, how do the SA values of specific tasks relate to the operator's dynamic priorities for SA?

Aircrews have many opinions about SA, but these may not neatly fit into current academic research into SA. For example, operators appear to conflate opportunity costs of performing a task that does not generate SA at the expense of another SA-generating task with tasks that actually reduce certainty through providing misleading information. Future studies should consider exactly what "negative SA" means and account for the difference between opportunity costs and incorrect mental models.

6 Conclusion

Capturing the effect of cockpit tasks on SA can allow for the refinement of predictive SA algorithms that incorporate both workload-dependent and task-dependent SA [1]. Clearly understanding how each task builds a pilot's mental model of his environment has clear advantages for avionics engineers when designing components, as well as for cockpit designers when incorporating those components into the aircraft's primary machine interface. This research shows that modern avionics, such as the HUD, MFD, and other sensors can provide strong positive SA, but depending on the mission phase and other factors, they may not be significantly more advantageous than analog counterparts. Furthermore, designers must not only consider the displays, but the interface required to control those displays, as well as the attention demanded by other cockpit tasks and their opportunity costs.

7 Disclaimer

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the U.S. Government.

References

1. Meyer, D.R.: Effects of Automation on Aircrew Workload and Situation Awareness in Tactical Airlift Missions. Air Force Institute of Technology, Wright-Patterson AFB (2015)
2. Secarea, J.V.: Beyond knobs and dials: towards and intentional model of man-machine interaction. In: Procedings of the IEEE Aerospace and Electronics Conference, pp. 763–783. IEEE (1990)
3. Endsley, M.R.: Situation awareness in aviation systems. In: Garland, D.J., Wise, J.A., Hopkin, V.D. (eds.) Handbook of Aviation Human Factors, pp. 257–276. Lawrence Erlbaum Associates, Mahwah (1999)
4. Endsley, M.R.: Toward a theory of situation awareness in dynamic systems. Hum. Factors **37**, 32–64 (1995)
5. Press, M.: Situation awareness: Let's get serious about the clue-bird (1986)

6. 57th Figther Wing. Intraflight command, control, and communications symposium final report. United States Air Force, Nellis AFB (1986)
7. Endsley, M.R.: A taxonomy of situation awareness errors. In: 21st Conference, Western European Association of Aviation Psychology. Western European Association of Aviation Psychology, Dublin (1994)
8. Hartel, C.E., Smith, K., Prince, C.: Defining aircrew coordination: searching mishaps for meaning. In: Sixth International Symposium on Aviation Psychology, Columbus (1991)
9. Klein, G.A., Calderwood, R., Clinton-Cirocco, A.: Rapid decision making on the fire ground. In: Proceedings of the Human Factors Society - 30th Annual Meeting (1986)
10. Endsley, M.R.: Situation awareness global assessment technique. In: Proceedings of the IEEE 1988 National Aerospace and Electronics Conference: NAECON 1988, vol. 3, pp. 789–795. IEEE, Dayton (1988a)
11. Endsley, M.R.: Measurement of situation awareness in dynamic systems. Hum. Factors: J. Hum. Factors Ergon. Soc. **37**(1), 65–84 (1995)
12. Venturino, M., Hamilton, W.L., Dverchak, S.R.: Performance-based measures of merit for tactical situation awareness. In: Situation Awareness in Aerospace Operations, pp. 4/1–4/5 (1989)
13. Pew, R.W.: Defining and measuring situation awareness in the commercial aircraft cockpit. In: Proceedings of the Conference on Challenges in Aviation Human Factors: The National Plan, pp. 30–31. American Institute of Aeronautics and Astronautics, Washington, D.C. (1991)
14. Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., Jenkins, D.P.: Human Factors Methods: A Practical Guide for Engineering and Design. Ashgate, Surrey (2005)
15. Vidulich, M.A. Hughes, E.R.: Testing a subjective metric of situation awareness. In: Proceedings of the Human Factors Society 35th Annual Meeting, pp. 1307–1311 (1991)
16. Kahneman, D.: Attention and Effort. Prentice-Hall, Englewood Cliffs (1973)
17. Billings, C.E.: Human-Centered Aircraft Automation: A Concept and Guidelines. NASA Ames Research Center, Moffett Field (1991)
18. Wiener, E.L., Curry, R.E.: Flight-Deck Automation: Promises and Problems. NASA, Washington, D.C. (1980)
19. Hart, S.G., Sheridan, T.B.: Pilot Workload, Performance, and Aircraft Control Automation. Ames Research Center. National Aeronautics and Space Administration, Moffett Field (1984)
20. Hart, S.G.: Helicopter human factors. In: Wiener, E.L., Nagel, D.C. (eds.) Human Factors in Aviation. Academic Press (1988)
21. Lovesey, J.: Information overload in the cockpit. In: IEEE Colloquium on Information Overload, pp. 5/1–5/5. IEEE, London (1995)
22. Haynes, W.: Validating the Need for a Heads up Display in the C-130 Avionics Modernization Program. Combat Aerial Delivery School, Little Rock AFB (1998)
23. Shinaberry, J.A.: (D. R. Meyer, Interviewer) (15 November 2013)
24. Kennedy, W.: (D. R. Meyer, Interviewer) (22 April 2015)
25. Endsley, M.R.: Automation and situation awareness. In: Parasuraman, R., Mouloua, M. (eds.) Automation and Human Performance: Theory and Applications, pp. 163–181. Lawrence Erlbaum, Mahwah (1996)