‘Airmanship’ on the Radar: Military Aircrew Instructors’ Perceptions of Non-Technical Assessment Methods, Training Strategies and Standards

Jonathan Allsop
RAF Central Flying School
Sleaford, Lincolnshire, UK
Jonathan.Allsop102@mod.gov.uk

Rob Hurcomb
RAF 22 Gp Directorate of Flying Training
Abbey Wood, Bristol, UK
Robert.Hurcomb496@mod.gov.uk

ABSTRACT

Airmanship refers to the non-technical skills required to successfully operate an aircraft. While it is critical for aircrew performance, there is a scarcity of research that has examined airmanship in military aviation. Therefore, a mixed-methods study was conducted to examine qualified aircrew instructors’ perceptions of: (1) trainees’ airmanship standards, and (2) airmanship assessment and training methods, with a particular focus on situational awareness (SA). In the quantitative phase, a cross-sectional survey was completed by 234 instructors who were recruited by convenience sampling from within flying training squadrons (n = 74), and aircraft-specific training units (n = 160). The survey collected demographic information, and assessed perceptions of trainees’ overall airmanship, as well as individual airmanship skills. In the qualitative phase, 36 instructors selected from across the flying streams (i.e., fast-jet, rotary, multi-engine and remotely piloted air systems) participated in one-to-one semi-structured interviews.

Survey results indicated that while 42% of aircraft-specific training unit instructors perceived trainee airmanship as meeting the required standard to commence their unit, 50% believed it was either slightly below, or below, the required standard. When examining perceptions of individual airmanship skills, SA and decision making were rated lower than other airmanship skills, indicating that these are key areas requiring further development. Results from the semi-structured interviews indicated several key themes that could contribute to these results. Specifically, while aircrew instructors shared a consistent operational definition of SA, they reported a lack of formal training in conceptual frameworks that could support the instruction and debriefing of trainee SA. Instead, a heavy reliance on informal knowledge transfer was apparent, with inconsistencies also identified in how airmanship skills are assessed.

Overall, the findings highlight the need for improved conceptual training, standardisation, and assessment tools. Addressing these gaps could enhance airmanship training, and ultimately improve aircrew performance and safety.

ABOUT THE AUTHORS

Jonathan Allsop, PhD is a chartered psychologist who works within the Royal Air Force’s (RAF) Central Flying School (CFS) as a Research & Development Psychologist. Jon has a range of research interests, but with a background predominantly examining the effects of acute psychological stressors on performance, and skill acquisition. In his role at CFS, Jon has conducted research into various novel technologies and techniques that have the potential to enhance aircrew instruction. This includes multiple studies investigating the effectiveness of XR technologies for aircrew training. Jon also supports several human performance initiatives that aim to enhance trainee and instructor performance.

Rob Hurcomb, BA joined the RAF from university in 1996, graduating to front line operations as a helicopter navigator in 2001. After multiple tours of Northern Ireland, Iraq, Bosnia, and Afghanistan he re-trained as a helicopter pilot in 2007. He was sent to the Search and Rescue (SAR) Force where he completed over 150 UK-based rescues and a year of command in the Falkland Islands. In 2016 he became a Qualified Flying Instructor delivering Elementary Flying Training to pilot trainees, RPAS trainees and university students. In 2020 he spent 6 months in Afghanistan as a liaison officer to a small US Army unit and then joined CFS to teach instructor training and manage Human Performance projects. Another command tour in the Falkland Islands followed before he joined the Directorate of Flying Training where he currently manages front line assignment of graduating military pilots.

‘Airmanship’ on the Radar: Military Aircrew Instructors’ Perceptions of Non-Technical Assessment Methods, Training Strategies and Standards

Jonathan Allsop
RAF Central Flying School
Sleaford, Lincolnshire, UK
Jonathan.Allsop102@mod.gov.uk

Rob Hurcomb
RAF 22 Gp Directorate of Flying Training
Abbey Wood, Bristol, UK
Robert.Hurcomb496@mod.gov.uk

INTRODUCTION

Modern military air operations have evolved significantly, confronting aircrew with versatile threats and increasingly complex missions, all within unpredictable, dynamic environments (Fornette et al., 2012; Marcu & Ivan, 2024; Ministry of Defence, 2018; Ministry of Defence, 2025). Along with these increasing operational demands, aircraft have become more complex and multirole, with advancements in technology simultaneously simplifying handling characteristics (Mitchell et al., 2003), while also providing an abundance of sometimes ambiguous information (Ebbage & Spencer, 2003; Vedula et al., 2023). Aircrew competency therefore increasingly relies less on the mastery of knowledge and skills sufficient to handle expected situations, processed mainly based on fast associations and readily applicable procedures (Fornette et al., 2012), but instead on the ability to adapt and make decentralised decisions (Bronk, 2021). Taken together, there is an increasing emphasis on the critical role of airmanship, alongside technical skills, in ensuring adequate aircrew performance and safety (Tsifetakis & Kontogiannis, 2017).

Airmanship has been characterised in various ways (e.g., Ebbage & Spencer, 2003; Kern, 2010; Mane, 1981; Negård, 2014), but definitions have consistently emphasised its multidimensional nature, encompassing both social and cognitive skills, encapsulated as “the non-technical skills required to operate the aircraft safely … from planning to debrief” (Royal Air Force Central Flying School, 2025). In aviation, interest in non-technical skills particularly grew during the 1980’s (Helmreich, Merritt, & Wilhelm, 1999), with the aim to improve safety through the structured development and assessment of a core set of non-technical skills. While the specific set of non-technical skills can vary slightly between aviation contexts (e.g., Hamlet, Irwin & McGregor, 2020), following definition and development work with subject matter experts (SMEs; see Heycock & Brown, 2000), United Kingdom (UK) military flying training adopted and predominantly continues to employ the following set of skills: situational awareness, decisiveness, communication, resource management, and mental capacity. Views regarding overall proficiency levels of these skills are frequently contentious, with changes to aircrew training often debated in terms of their perceived impact on non-technical skills standards (e.g., Emerson et al., 2022; Penney, 2025). Despite this continual link, there is commonly very little large-scale systematic capture, or quantitative analysis, of instructor perceptions of airmanship standards. This lack of systematic capture and analysis limits the ability of flying training organisations to identify consistent variations between training streams (the training track that aircrew are assigned to), or trends across time. In other related domains, the structured capture of non-technical skills proficiency levels has resulted in actionable insights (e.g., França et al., 2022; Irwin et al., 2016). For example, in an aircraft maintenance context, an analysis of maintainers' perceptions of their non-technical skills revealed specific areas of relative weakness, as well as differences in standards across geographical locations (Irwin et al., 2016). An aim of the present study was therefore to address this gap by capturing UK military aircrew instructors’ perceptions of airmanship standards. Given the central role of situational awareness in influencing aircrew performance, the present study also specifically examined SA as a critical component of overall airmanship.

Situational awareness (SA) has consistently been recognised as one of the most fundamental components of airmanship, with the generation and maintenance of SA being crucial for aircrew performance and safety (e.g., Endsley, 1995). Indeed, an analysis of commercial aviation accidents occurring between 2000 and 2016, identified SA as the most common human factors cause, followed by non-adherence to procedures (Kharoufah et al., 2018), with similar findings in military aviation contexts (Hartel, Smith & Prince, 1991). A widely accepted definition of SA is “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future” (Endsley, 1988, p. 97). Endsley’s accompanying

comprehensive cognitive model of SA, initially developed and validated within military aviation contexts (Endsley, 1987), has become a cornerstone of SA research, and training, across a wide range of domains (Wickens, 2008). Summarily, the model explicitly identifies the interactions between a number of cognitive processes (e.g., attention, memory), and various factors (e.g., individuals' goals and training, cockpit design, etc.) that combine to produce SA. In particular, the model distinguishes between three levels of SA, with the initial level involving the perception and awareness of attributes, dynamics, and status, of task-relevant elements in the environment. Level two SA goes beyond just awareness, to include an understanding of the significance of those elements in relation to an individuals' goals. A novice pilot may be able to perceive individual elements equally as well as an expert but not be able produce an integrated comprehension of their importance, and impact, to what they're trying to achieve. Finally, level three SA involves the capability to predict the future actions of elements in the environment, accomplished by understanding the status and dynamics of those elements (Endsley, 1999). The diagnostic differentiation among these three levels is crucial because each highlight distinct perceptual and cognitive processes, resulting in different approaches to both training (Walker et al., 2009; Wickens, 2008) and assessment (Endsley & Garland, 2000). While previous research in general aviation contexts has examined instructional methods aimed at fostering SA (e.g., Prince & Salas, 1998), there is a lack of research specifically investigating perceptions, practices, and challenges of training and assessing SA among military aviation instructors.

Research Aims

The aims of this study were twofold:

1. The quantitative phase aimed to analyse instructors' perceptions of both overall airmanship standards, and the standards across airmanship skills. These analyses were conducted for both operational conversion units (OCU; aircraft-specific training units, akin to Formal Training Units in the USAF), and flying training schools (FTS), and across a wide range of instructors to allow comparisons between training streams and instructor roles.
2. The qualitative phase specifically focused on SA, aiming to explore instructors' conceptualisations, identify effective practices for fostering SA, and investigate practical challenges associated with its assessment.

QUANTITATIVE PHASE: INSTRUCTORS' PERCEPTIONS OF TRAINEES' AIRMANSHIP STANDARDS

Methods

234 aircrew instructors from a wide range of aircraft types provided responses with almost all ab-initio facing OCUs and FTSs represented; their demographics are summarised in Table 1. A survey collected demographic data and asked the responding instructor to rate the airmanship standards of trainees commencing their unit. Separate questions captured both an overall assessment of trainees' airmanship standards, and the standards of individual airmanship skills (i.e., SA, decisiveness, communication, resource management, and mental capacity), with participants responding on a five-point Likert scale anchored in relation to perceived required standards, for example '1 - Below the Standard Required'. The survey results were collected anonymously via Microsoft Forms, and instructors first indicated whether they had observed enough trainees to comment; if not, an N/A option was enabled for subsequent questions.

Table 1. Mean and Standard Deviation (SD) of Instructor (*N* = 234) Demographics Across Streams

Sample	Category	Characteristics	Fast Jet	Multi Engine	Rotary	RPAS
OCU	Role	Pilot *n*	11	12	56	3
		Rear crew *n*	-	21	50	7
	Hours	Flying Hours	2279.5 (738.0)	3822.1 (1739.0)	3296.0 (1691.3)	2949.6 (1751.0)
		Instructional Hours	694.5 (330.7)	936.4 (846.3)	1003.1 (1081.6)	380.8 (380.8)
FTS	Role	Pilot *n*	25	9	24	-
		Rear crew *n*	-	-	16	-
	Hours	Flying Hours	3340.0 (2883.4)	4311.1 (1996.5)	4347.8 (2428.5)	-
		Instructional Hours	850 (1140.5)	1133.3 (995.9)	1098.9 (1077.9)	-

Overall airmanship ratings were analysed using two-way between-subjects analysis of variance (ANOVA), testing the effects of training stream and instructor type. Individual airmanship skill ratings were examined with a mixed-factorial ANOVA, treating training stream and instructor type as between-subjects factors and airmanship skills as a within-subjects factor. Separate analyses were conducted for OCU and FTS instructors. For conciseness, apart from ANOVA effect sizes, probability values are the only test statistics reported. Data analyses were performed in SPSS, Python, and Prism.

RESULTS

Instructor Perceptions at Operational Conversion Units

OCU Perceptions of Overall Airmanship Standards

At OCUs, the main effects for training stream (i.e., fast jet, multi engine, rotary or RPAS), instructor type (i.e., pilot or rear crew), and the interaction effect between training stream and instructor type were all not statistically significant (p's > .12). In summary, instructors' perceptions of trainees' overall airmanship standards were consistent across training streams and instructor roles. To aid understanding of the distribution of responses, a percentage breakdown for response categories is shown in Figure 1.

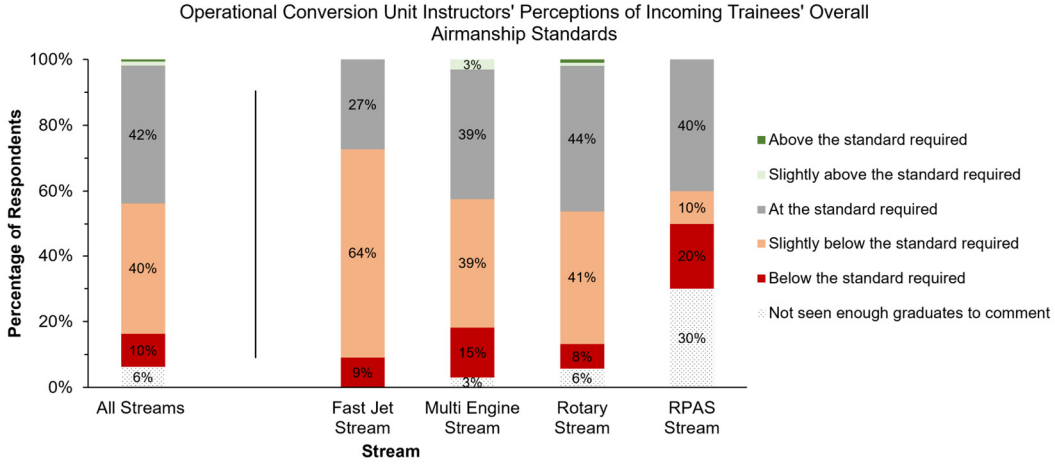

Figure 1. OCU instructors' perceptions of incoming trainees' overall airmanship standards – proportions of responses

OCU Perceptions of Individual Airmanship Skills

The analyses revealed a significant main effect for airmanship skills (p = .007, η_p^2 = .03), meaning that instructor perceptions of trainees' airmanship differed significantly across the five airmanship skills. Post-hoc tests revealed that SA and decisiveness were rated as lower than all other airmanship constructs (p < .05). The main effects for training stream and instructor role, were both non-significant (p's > .22). Interestingly however, the interaction between airmanship skills and training stream was statistically significant (p = .05, η_p^2 = .04). This means that the relative rating of each airmanship skill differed depending on which training stream the instructors were from. In other words, instructors from one training stream evaluated certain airmanship skills (e.g., SA or decisiveness) differently compared to instructors from another stream – this can be seen in the different pattern of results in Figure 2. All other interactions were non-significant (p's > .4). To explore the interaction between airmanship skills and training stream, separate repeated measures ANOVAs were conducted for each stream.

Figure 2A shows the fast jet data, the analysis was statistically significant (p = .01, η_p^2 = .29) with post-hoc tests showing that decisiveness was rated lower than SA, communication and resource management, and mental capacity rated lower than communication and resource management. Figure 2B shows the multi engine data, the analysis was marginally significant (p = .06, η_p^2 = .07), with SA and decision making rated lower than resource management and mental capacity (p < .05). Communication was rated lower than mental capacity, and the difference between decisiveness and mental capacity approached significance (p =.07). Figure 2C shows the rotary data, this was also statistically significant (p = .01, η_p^2 = .14), with essentially SA, decisiveness, and communication all rated lower than

resource management and mental capacity (although two of these results were significant at p = .052). The analysis for RPAS was not statistically significant (p = .52).

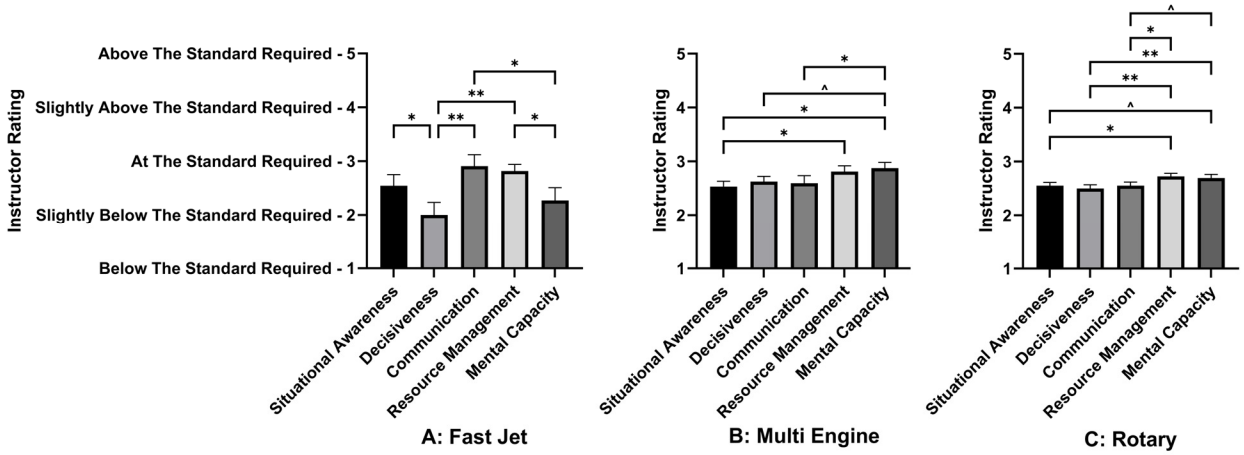

Figure 2. OCU instructors' average ratings of trainees' individual airmanship skills. Error bars represent standard error of the mean. Horizontal brackets indicate statistically significant comparisons between airmanship skills (^ p < .07, * p < .05, ** p < .01).

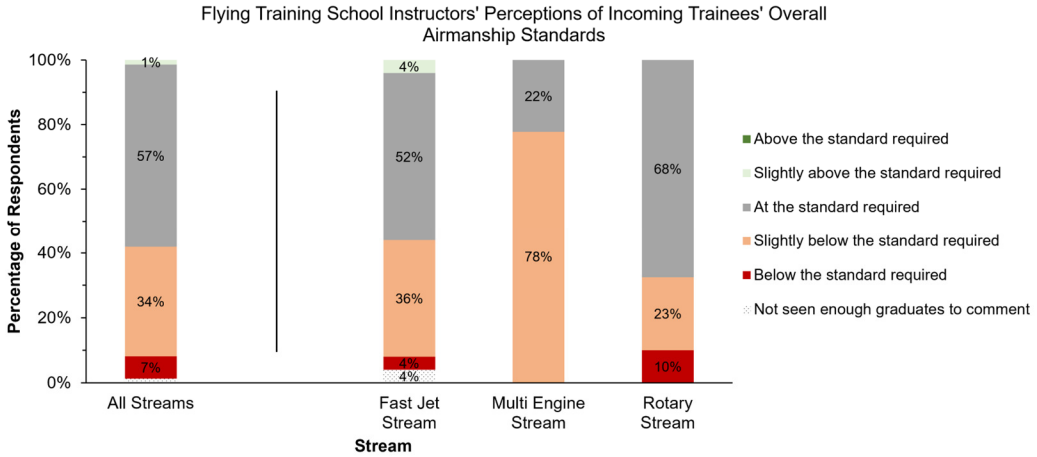

Figure 3. FTS instructors' perceptions of incoming trainees' overall airmanship standards – proportions of responses

Instructor Perceptions at Flying Training Schools

Flying Training School Perceptions of Overall Airmanship Standards

At FTSs the main effects for training stream (i.e., fast jet, multi engine, rotary) and instructor type (i.e., pilot or rear crew) were non-significant (p's > . 27), and the interaction was not calculated as non-pilot (e.g., rear crew) data was only collected from squadrons within the rotary stream. In summary, instructors' perceptions of trainees' overall airmanship standards were consistent across training streams, and instructor types. To aid understanding of the distribution of responses, percentage of respondents is shown in Figure 3.

Flying Training School Perceptions of Individual Airmanship Constructs

The statistical analyses revealed a significant main effect for airmanship construct (p = .000, η_p^2 = .09), with post-hoc tests revealing that SA and decisiveness were rated as lower than all other airmanship constructs (p < .05). The main effect for training stream approached significance (p = .06, η_p^2 = .08), with post-hoc tests revealing that multi-engine instructors' perceptions of trainees' airmanship entering their squadron were statistically lower than their rotary counterparts (p = .04) and approached significance for their fast jet counterparts (p = .06).

Similar to the OCU results, the interaction between individual airmanship skills and training stream was significant (p = .003, η_p^2 = .08), Separate RM-ANOVAs were therefore conducted for each stream. Figure 4A shows the fast jet data, the analysis was statistically significant (p = .002, η_p^2 = .17), with SA being rated as poorer than all other airmanship skills. In addition, decision making was rated lower than mental capacity. Figure 4B shows the rotary data, analysis was again significant (p < .001, η_p^2 = .22), with decision making being rated significantly lower than all other airmanship skills. The multi engine analysis was not significant so is not displayed (p = .28, η_p^2 = .14), and all other main effects (e.g., instructor type) and interactions (e.g., airmanship skill x instructor type) were also non-significant (p > .1).

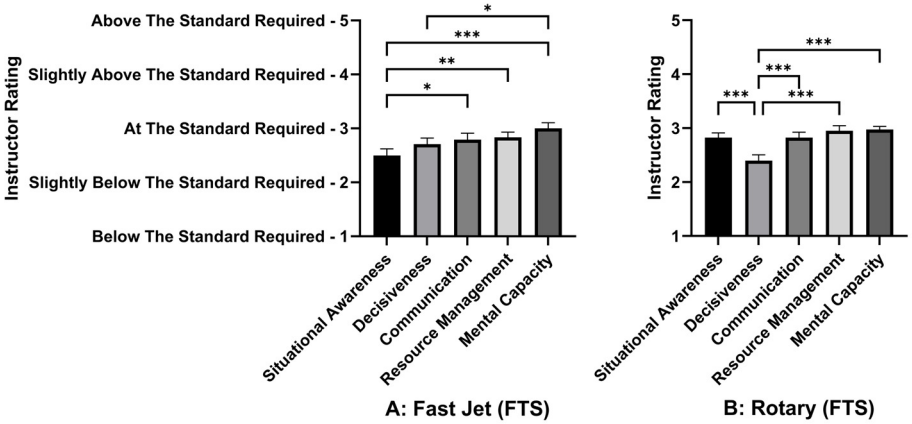

Figure 4. FTS instructors' average ratings of trainees' individual airmanship skills. Error bars represent standard error of the mean. Horizontal brackets indicate statistically significant comparisons between airmanship skills (* p < .05, ** p < .01, *** p < .001).

Predictors of Overall Airmanship Ratings

A multiple linear regression was conducted on data pooled from both FTS and OCU instructors, to examine the extent to which flying experience, instructional experience, and each of the individual airmanship skills predicted instructors' overall ratings of airmanship. The overall model was statistically significant (p < .001) and explained 44.6% of the variance in overall airmanship ratings (R^2 = .446). Situational Awareness (β = .362, p < .001) emerged as the strongest predictor, followed by Communication (β = .248, p < .001) and Decisiveness (β = .188, p = .001). However, Resource Management, and Mental Capacity were not significant predictors. Flying Hours and Instructional Hours also did not predict overall ratings of airmanship (p > .05), indicating that impressions were similar across instructors.

QUALITATIVE PHASE: INSTRUCTORS' PERCEPTIONS OF SITUATIONAL AWARENESS TRAINING AND ASSESSMENT METHODS

Methods

36 aircrew were recruited from within OCUs (n = 27) and FTSs (n = 9), all but two were aircrew instructors. Purposive sampling was used to ensure a wide range of instructors were represented across a wide breadth of types and streams. To illustrate, FTSs were represented by 9 pilots (2 elementary, 3 fast jet, 2 multi engine and 2 rotary), with OCUs being represented by 18 pilots (4 fast jet, 6 multi engine, 4 rotary, and 4 RPAS) and 9 rear crew (5 multi engine, 2 rotary, and 2 RPAS). Participants had an average of 2757.4 flying hours (SD = 1685.9) and 624.0 instructional hours (SD = 679.5).

Semi-structured interviews ($M_{minutes}$ = 64.4, SD = 15.4) were conducted either online or in-person, with the question schedule and further probes being developed in advance around the research questions. Analyses were guided by reflexive thematic analysis (Braun & Clarke, 2006, 2020); both authors reviewed the interview transcripts or recordings, or both, on multiple occasions as part of an iterative analytical process. Theme development incorporated both inductive and deductive approaches, with the deductive component rooted in Endsley's model of SA. Themes

are provided with direct instructor quotes to authentically illustrate the findings, a practice essential for ensuring accuracy and credibility (King, 2004).

RESULTS

Conceptual Foundations of Situational Awareness – Instructor Perspectives

Shared but General Operational Definition

Instructors consistently converged on similar themes when asked to provide a definition of SA. A key theme related to having knowledge of aircraft positioning, both currently and into the future, for example "knowing where the aircraft is in space and time, knowing where it's going to be, and knowing where those around you are, and where they're expected to be in the near future". Another theme related to wider safety factors and tactical factors. "An understanding of the area and factors around the aircraft … everything to allow you to understand where the aircraft is and how you can employ it properly". For multi crew platforms the theme of shared SA arose for some participants, "know where your aircraft sits in space …. and then just having your crew at the same at the same point as you". These high-level elements closely align with elements of Endsley's (1995) model.

Lack of Conceptual Handrails

A key theme that emerged was a lack of awareness, or use of, conceptual handrails to support SA-related instructional activities. For instance, almost all instructors stated that they did not have an explicit awareness of Endsley's model as a conceptual handrail to support SA assessment or development, "as I didn't even know what it stood for, I'd suggest no [I do not have an awareness of Endsley's model]". A few instructors suggested they may be applying the model implicitly even when they are not consciously aware of it "I've never heard them under those three levels, but I think I do that internally".

One notable exception was one instructor's explicit use of the acronym NUTA as a conceptual handrail to support their SA-related instruction, "the thing we do speak to trainees about is NUTA, so for SA – have you *Noticed* it? If you haven't noticed it then you can't *Understand* it, and if you haven't understood it then you can't *Think Ahead*". The instructor stated that this acronym was particularly important for fault diagnosis and debriefing trainees "they really understand it [NUTA] when you come back to it, 'Oh, do you know why you did that badly? …. It's because you didn't notice this, so you couldn't understand the implications, and therefore you couldn't think ahead'". This handrail is directly aligned with Endsley's three levels of SA, and could be a useful way of enhancing SA instructional activities.

Sparse Conceptual Training and Reinforcement

Instructors consistently described a lack of conceptual training on SA, and negligible systematic checks on whether that conceptual knowledge is carried through into instructional practice. An experienced instructor with significant front line experience on two aircraft streams (i.e., fast jet and multi engine) typified this theme, "You know, in my whole time, I've never seen a PowerPoint on SA, certainly never had one delivered to me and it's not one I've ever delivered", with many others sharing similar sentiments "I can't think of any formal training I've had to develop it [understanding of SA as a concept]". This sentiment included those instructors who had very recently completed flying training and instructional courses. Instructors also consistently mentioned a lack of training on practical strategies, or underpinning philosophies, to help develop or assess trainee SA. Taken together, these accounts reveal a systemic training gap: instructors are expected to foster and assess SA without solid conceptual foundations, or ongoing support to do so.

Informal Peer Learning

In the relative absence of formal conceptual training or reinforcement, instructors consistently identified the prevalence and perceived utility of informal peer learning to develop understanding, as one instructor put it "Probably crew room chat … that's how I develop my understanding .... If I don't know, I'll ask somebody and then I'll get them to sort of put an example onto that.", and another echoed the importance of informal sharing of conceptual understanding between peers, "I think just crew room, instructor, office banter and discussions … just chatting amongst people, listening to various ideas". Although peer learning is viewed as crucial, its dominance as the primary method for sharing SA-related instructional practice leaves knowledge sharing ad-hoc and open to variability.

Perceptions of Situational Awareness Development and Training Strategies

Exposure

Exposure to realistic scenarios in a frequent, varied, and increasingly complex manner emerged as a key theme, with three sub-themes: live flying, scaffolding/progressive complexity, and effective simulator usage. For instance, one instructor succinctly suggested "more flying, more reps" as the method to improve SA, with others providing similar sentiments related to the quantity of exposure "exposing them to more of the real airplane, I'm flying them more. More experience and time in the air". Instructors also frequently stated that the specific nature of the exposure was important for promoting SA, with a particular focus on scaffolding and ensuring progressive complexity, "take it back down to bare bones nuts and bolts … then just start layering on the tasks … now let's try and build the complexity as we go through.", and manipulating the non-technical elements specifically, "You want the flying to be simple, and then I would want some more things to think about externally … more timings to be hit, more users, more frequencies to contact, more airspace. We can create airspace on our maps that's fake, but they'd still have to abide by it".

Finally, instructors described how simulators offer a controlled environment to expose trainees to scenarios that foster SA, "we do it really well in the simulator …. we can put lots of aircraft in the stack, we can put enemy aircraft … it allows us to expose students to tasks that they won't get live, potentially, allowing them to build SA". This was echoed by other instructors, but a tension was sometimes noted with the resourcing of simulators "Our synthetic training, the sim is brilliant, but we've under resourced it 100%. We've got a QFI [qualified flying instructor] that's running air traffic, that's running the operation of the simulator, they're running the pilot monitoring role, which you're obviously supposed to be developing Crew Resource Management, and they're also the instructor …. if they were utilised properly, you could really explore areas of SA that people aren't comfortable in". Taken together, these themes suggest that SA development should be grounded in a deliberately scaffolded syllabus that blends ample live flying and fully resourced simulated training, with complexity being added methodically rather than opportunistically.

Cognitive Groundwork

Participants consistently framed SA development as being heavily supported by ground-based activities, with four main themes emerging: the importance of effective preparation and briefing, checking mental models, chair flying, and effective debriefing. As one instructor explained, "I'd have a look at their briefing as to how prepped they are before they get airborne as to what they've got in their mind, what they're mental model is", supported by "In the brief you can usually identify pinch points from experience as to where there are going be threats…. as we're arriving at it [when airborne] from a SA point of view, did they take that on board, and did they factor in or … have they taken no notice of it and ploughed on regardless?". The impact of effective chair flying was also emphasised by instructors, "the first time round their heads were on fire, as expected, and their SA was not great … so we got them back in the classroom [and got them chair flying] … there's been a great improvement, the chair flying has really paid dividends". Finally, instructors described effective debriefing as the keystone "I think the debrief is the keystone, the most key one … walking through that and seeing where the plan went well or wrong is the strongest thing you can do, and then asking why". These themes emphasise the importance of a range of effective cognitive groundwork strategies, techniques and assessment.

Attention Management Techniques

Three main sub-themes emerged under a theme of attention management: work-cycles, anticipatory prompts, and ensuring attentional resources are freed via consolidation of technical skills. Participants repeatedly referred to the importance of work-cycles in fostering SA "I think anytime that you can give somebody an easy work cycle to remember it will only assist with their situational awareness", with one instructor suggesting that this particularly helps to distribute attention, a crucial component of SA development and maintenance "For the camp of people that get very tightly focused, then it would be … giving them that very rigid sort of work cycle to work on. And just constantly do that over and over again". Coupling work-cycles with the use of anticipatory prompts was commonly voiced as an effective strategy to manage attention, for example "what's next?" questioning and "If-Then" planning, "What is their [the trainee's] work-cycle like? What is the next event? What could happen next? … If this happens, these will be my actions, if the weather comes down, we'll run this check.". Finally, building capacity by ensuring that individual tasks do not take up too much time, attention or cognitive resources was suggested to be a key enabler for effective attention management "I'm going to expose them to doing those tasks. Lots, lots and lots and lots … they can almost do it without thinking about it…. it doesn't take much of their capacity, which then allows them to use that capacity on other areas, new areas or arising information".

Use of Crew Members

In multi-crew aircraft, instructors stressed the importance of proactively using crewmembers in order to develop SA, "delegation and not abdication is quite an important one in that we'll have a lot of inexperienced pilots that will fly with experienced sensors, so they can use the experience in the other seat to kind of lighten their workload, which then will grow their SA", with structured updates on certain information being used to help maintain SA "sensor for me today, you're going to be my primary for these three windows and you are to brief the crew every time there's a change".

Appropriate Instructor Intervention

Instructors repeatedly emphasised that well-timed, forward-looking questions are central for developing SA, as one explained "in the quieter times … I'll go, 'OK guys, what could happen next?'" with instructors particularly noting the type of question as being important "I'll try and ask open questions, try and get someone to explain what happened first in their own words… why did you do X or what did you want to do at this point?". Instructors also offered caution suggesting that ill-timed questioning can backfire, "I think open questioning at the wrong time can be really ineffective when you [a trainee/pilot] don't know what is missing, and you're being asked an open question … it's to do with that pull and push of information at the correct time" with the goal being to intervene in a manner that allows regaining of SA "you want to alleviate the pressure so they can get back in front of the aircraft.". Together, these accounts portray instructor intervention as a balancing act: skilful prompts that expose and extend the trainee's mental model, whereas ill-timed probes erode the cognitive resources required for SA.

Perceptions of Situational Awareness Assessment Methods

Subjectivity

When talking about their experiences of monitoring and assessing SA, instructors consistently characterised the process as being, to varying degrees, unavoidably subjective as typified by the following quote "the subjectivity has to come out, the quality of the instructor's subjective assessment will be different from one to the other, even with the most objective grading system, you're still going to have a healthy dose of subjectivity", with other participants supporting this "it's just not possible to characterise objectively airmanship or situational awareness. It's not possible". Such views were linked to the variability of live flying and the nature of flying training, "we have to [assess subjectively], we don't fly in the same conditions every day, the trainee hasn't had the same build-up to each sortie, they haven't had the same continuity whether temporal or instructional…".

Subjectivity often took the form of tacit, experience-based "gut feelings" or "instincts" being used to make judgements about trainees, with one participant saying "It's a gut feeling. It's just that instinct, that something is missing", and another saying "100% gut feeling is a factor here". However, others talked about the need for a balance between gut feelings and striving to be as evidence-based as possible, looking for objective behaviours or cues to support their assessments, "I think there's always going to be a little bit of gut feeling with experience. But for SA I'd want to be able to say to the trainee. This is why I feel that your SA was poor, or very good", supported by "I do try and avoid gut feelings where I can, I think hopefully all instructors will say that they only use a gut feeling where they are lacking other evidence".

Grade Sheets

Instructors described various grade sheets for assessing SA, with several themes emerging concerning their design, including: the number of items, the grading frame-of-reference, scale resolution, and the importance of the supporting narrative. Instructors suggested that having a large number of assessment items can lead to straight-lining or non-differentiation, "there's a lot of boxes that require a lot of numbers…when there are too many you kind of gloss over". In support of this, other instructors suggested that having less items, that are explained more thoroughly, is beneficial "find a way of keeping it simple but have the aircrew instructors understand those things in detail. It's like, yes, I've got 10 things [in total, not just airmanship skills] to assess, but I've been provided with 10 in depth descriptions on what that means.".

Further scale design tensions were apparent, with the choice of evaluation anchor, or frame of reference, sometimes leading to difficulties. An advocation for criterion-referenced grading, over norm-referenced grading, was apparent as exemplified by: "while sometimes there's a bit of a lack of correlation [in criterion-referenced grading] between the verbiage as to what a number might be, I think it's absolutely better than me assessing a trainee and saying, well, I think at this stage of the course [you're at about the average standard] versus another instructor who might have a

different point of view". Tensions were also apparent in relation to scale resolution and granularity, "on the OCU it was out of eight, and now it's one to five. I preferred out of eight … not all threes [centre point grades] are created equal", echoing sentiments that grading SA was easy at the extremes, with midscale judgements being more difficult "it's very hard to put a really objective line in the sand as to what is good. It's quite easy to tell when they're at the top end of the score". Taken together, these issues were often linked with instructors prioritising the narrative aspects of the assessment "the content of the report that I really care about, because that will give me the best indicator …. you can give someone a score, but what does that actually mean?".

Assessment Supporting Guidance

Instructors' views differed on the supporting assessment guidance that accompanies the grading of SA, with one participant saying, "I quite like the guidance … what makes it a very good grade, or what makes it marginal - well it needed instructor intervention. So, I actually find the description quite useful". However, instructors also provided negative views on the supporting guidance, especially in relation to the lack of clarity and precision "There is some guidance, but I would say, it's not subjective it's just very vague guidance", and "the definitions need to be a lot clearer to say somebody did this to an average standard, to a high average standard, to an above average standard. You know, that doesn't really mean anything. What's average for starters?". Other instructors supported this sentiment, but also offered a view that more detailed, aircraft specific, definitions of SA could be helpful, "The descriptions for situational awareness in our taxonomy, or all of the subtopics, they're maybe two lines, could even be one line ….. so I think the descriptions could be much better. You could, put it in the taxonomy, an A4 page on what we [our squadron/aircraft type] think is demonstrating good SA".

Behavioural Indicators

Only top-level, cross-platform, sub-themes are provided for this theme as more granular, stream-specific, themes are outside of the scope and aims of this paper, and form part of a wider ongoing project. Instructors consistently discussed how behaviours that are indicative of poor SA are easier to observe than those indicative of good SA, "it's the absence of a negative, makes a positive … that we were looking for, rather than you know actual positives". Relatedly, there was a consistent link between poor SA being observable through poor, or unexpected, decision making "it's kind of the absence of the normal … they've just gone spearing off, making a different decision. So it makes it much easier to spot".

Instructors also consistently discussed attentional issues being linked to degraded SA, with frequent examples of inattentional deafness "sensor operator keeps asking to reposition, but the pilot is simply not hearing them - their SA is so degraded", and lack of visual scanning, "any situation where you get drawn into one very narrow area at the expense of everything else that's going on around you, that shows that you've lost SA". In contrast to the consistent discussion of negative behavioural indicators, one consistent positive behavioural indicator was crew communication in multicrew aircraft, "There's so much in crew comms … almost everything important should get verbally acknowledged or described … to make sure the whole crew awareness is kept up".

DISCUSSION

The present research aimed to quantitatively analyse instructors' perceptions of overall airmanship standards and its component skills, compare results across streams and roles, and identify which individual skills most strongly relate to perceptions of overall airmanship. The aim of the qualitative phase was to develop a clear understanding of instructor perceptions of SA as a concept, capture effective ways of fostering SA, and examine the practices and challenges of assessing SA. Collectively, the findings offer unique insights into standards, perceptions, and practices related to airmanship and non-technical skills in military aviation, but most importantly have practical implications for military flying training organisations aiming to prepare future aircrew for operational activities.

The quantitative analysis revealed that perceptions of overall airmanship standards of incoming trainees were not statistically different between streams and instructor roles, both at OCUs and within FTSs. Descriptive statistics showed that approximately half of instructors (50% at OCUs, 41% at FTSs) perceived the airmanship of their incoming trainees to be either slightly below, or below, the required standard, highlighting significant opportunities to improve current training practices. The slightly higher proportion of OCU instructors who indicated that incoming trainees' airmanship falls below required standards might reflect differing operational demands and complexities at the OCU

level. This highlights a potential gap between the training outcomes achieved at FTSs and the expectations of modern front-line platforms.

While perceptions of overall airmanship were consistent across streams, distinct differences were revealed when examining individual airmanship skills. SA and decisiveness repeatedly emerged as areas of relative weakness. Analyses revealed that, overall, these two skills were rated lower than communication, resource management and mental capacity, independent of any stream-specific differences. The regression analyses confirmed SA as the strongest predictor of instructors' overall perceptions of airmanship, followed by communication and decisiveness. This underscores the central role of SA (Hartel et al., 1991; Kharoufah et al., 2018) and shows that overall impressions of trainees' airmanship predominantly depends on their ability to generate and maintain SA. Taken together, any broad interventions at both the instructor-level, and training system-level, should initially focus on these two skills, particularly SA.

Interestingly, perceptions of individual airmanship skills varied significantly between streams, indicating stream-specific nuances. For instance, rotary and multi engine OCU instructors rated all three of SA, decisiveness and communication lower than other airmanship skills. In contrast, fast jet OCU instructors highlighted specific deficiencies in decisiveness and mental capacity, while fast jet FTS instructors emphasised weaknesses primarily in SA, followed by decisiveness. Together, these results suggest an intriguing developmental pattern within the fast jet stream, where deficiencies initially perceived as SA issues at the FTS, may manifest more clearly as decisiveness problems at the more complex OCU stage. This potentially implies that underlying sub-skills, techniques, and strategies associated with SA become particularly crucial – and more noticeably lacking - during advanced training phases. These stream-specific insights, suggest that tailored interventions addressing stream-specific developmental trajectories and training demands may enhance the effectiveness of any intervention efforts that aim to improve overall airmanship.

The qualitative findings revealed that instructors shared a consistent but broad understanding of SA as a concept, however this understanding was based on informal peer learning, with minimal formal training. Instructors consistently highlighted a lack of training on strategies or philosophies for developing and assessing trainee SA. With few exceptions, instructors also lacked explicit conceptual handrails to guide instructional activities, or training design, and did not recognise validated SA theory or models. This is despite these being presented on mandated instructional and human factors courses. Such a finding is perhaps unsurprising, given that previous meta-analyses examining the effectiveness of non-technical skills training courses have shown that such training usually results in statistically significant improvements in learner attitudes (i.e., positive initial reactions), but not improvements in post-course knowledge, or changes in actual on-the-job behaviours (O'Connor, Campbell, Newon, Melton, Salas, & Wilson, 2008). However, a notable caveat is that non-technical skills training becomes significantly more effective when it emphasises behavioural aspects through a combined approach of conceptual classroom lectures, opportunities to practice desirable behaviours, and behaviourally based feedback (e.g., Chang et al., 2017; Salas et al., 2006; Walshe et al., 2019). Taken together, addressing this training gap requires either making the initial training more effective, systematically consolidating and verifying that conceptual knowledge is applied in actual instructional practice, or preferably, a combination of both.

The results captured a number of strategies and philosophies that instructors believe are effective for promoting SA. Exposure to realistic, appropriately scaffolded, and progressively challenging scenarios, was viewed as foundational for SA development, with an emphasis on live flying. However, there was also acknowledgement that simulated training that is well-resourced can offer enhanced training opportunities. This supports previous research, for example Sarter & Woods (1991, p. 53) state: "to give this phenomenon [SA] a chance to occur, it is necessary to stage complex dynamic situations that require resources comparable to high-fidelity full-mission simulations". Several options exist to facilitate this type of complex simulation, such as employing multi-network simulators that integrate Air Traffic Controllers and multiple aircrew (see Mishler et al., 2022), or utilising a software-driven Simulated Air Traffic Control Environment (SATCE; see Allsop & Keeling, 2023). However, a philosophical shift – or in some cases either a return to previously held principles or an increase in emphasis - is also perhaps equally as important. Just as military flying training has deliberately scaffolded the development of technical skills for decades, a similarly intentional approach is required for non-technical skills. This should include scenario-based simulations, or airborne training, explicitly designed around non-technical skills, rather than implicitly treating these skills as secondary outcomes of ample airborne training hours that are focused primarily on technical skills.

A number of attention related strategies were also viewed as critical for supporting SA, with a particular focus on work-cycles that include anticipatory prompts. Bolstering the existing structured training of workcycles with visual attention monitoring tools such as eye-tracking, could assist SA development (e.g., Cooper et al., 2013) by ensuring that attention distribution is continually, and automatically, reinforced and assessed. Consolidation of technical skills emerged as another strategy to ensure that attention can be distributed effectively for SA development. While continuity of exposure to realistic scenarios (live or simulated) is a primary basis for the consolidation of technical skills, ensuring that trainees' cognitive groundwork is as robust as possible is also essential, as echoed by the instructors. Previous research has shown that students often rely on passive preparation and learning strategies (Anthony, 1996; Miyatsu, Nguyen & McDaniel, 2018), that are in almost direct opposition to well-established research findings on effective, effortful learning and preparation strategies. Therefore, ensuring all trainees are aware of established principles from learning science (e.g., spaced repetition, testing effect etc.) and applying a wide variety of evidence-based techniques to consolidate knowledge and skills (e.g., peer role-playing, tabletop exercises, regular self-assessments, peer-teaching, etc.) could help to ensure that attentional resources are not overly consumed by technical skills.

The findings related to assessing SA reveal a reliance on the absence of negative trainee behaviours, with instructor assessments often not helped by vague supporting guidance, broad grading gradations, and inherently subjective frames of reference. Instructors consistently stated that SA assessment will always be subjective within flying training, to a greater or lesser extent, but obvious avenues for improvement are still possible across several areas. Firstly, the development of behavioural anchored rating scales, at appropriate levels of specificity (i.e., each aircraft type, and also mission-set) may help to provide more assessment guidance to support instructors (see Hamlet et al., 2023), especially those new to instructional roles. Secondly, and relatedly, more consistent utilisation of criterion-based assessment frames of reference across training courses, may help to standardise and calibrate instructor assessment to more objective, and well-documented standards. Thirdly, holding airmanship standardisation sessions using video vignettes (e.g., Yule et a., 2008) from actual trainees may help to, at the very least, raise awareness of differences in opinions, and at best, help calibrate scoring. Finally, trialling more objective SA measurement techniques such as freeze-probe (Endsley, 1987) or real-time probe (Durso & Dattel, 2004) techniques may assist in both assessment, but also help to raise awareness of question and intervention content that can be used to assess across the three levels of SA.

In conclusion, this study highlights important gaps in the standards and instructional practices regarding airmanship in a military aviation context. Instructor perceptions indicate that SA and decisiveness are areas in greater need of targeted improvement. Qualitative insights, centred around SA, further underscore the need for improved training and utilisation of conceptual frameworks, structured training scenarios, and more detailed and objective assessment tools to reliably develop and assess trainees' SA. Systematically addressing these gaps could significantly enhance aircrew training, ultimately improving aircrew performance and safety.

REFERENCES

Allsop, J. & Keeling, R. (2023). On Approach to Reality: The Impact of a Simulated Air Traffic Control Environment (SATCE) on Workload and Situational Awareness in Military Aviators. *Proceedings of the 2023 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC),* Orlando, Fl.

Anthony, G. (1996). When Mathematics Students Fail to Use Appropriate Learning Strategies. *Mathematics Education Research Journal, 8(1)*, 23-37.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3(2),* 77–101.

Braun, V., & Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology, 18(3)*, 328–352.

Bronk, J. (2021). De-Centralized Command and Control in Air Operations: Implications for Air Battle Management and Mission Command. In *Multi-Domain Operations, Artificial Intelligence and Information Dominance*; The Air Power Journal

Chang, A.L., Dym, A.A., Venegas-Borsellino, C., Bangar, M., Kazzi, M., Lisenenkov, D., et al. (2017). Comparison between simulation-based training and lecture-based education in teaching situation awareness: A randomized controlled study. *Ann Am Thorac Soc, 14(4),*529–35

Cooper, S., Black, B., Bogossian, F., Browning, M., Munro, G., O'Meara, P., et al. (2013). *Can Eye-tracking Technology Improve Situational Awareness and Student Feedback during Simulation?* Office for Learning and Teaching, Innovation, Science, Research and Tertiary Education, Sydney

Durso, F. T., & Dattel, A. R. (2004). SPAM: The real-time assessment of SA. In S. Banbury & S. Tremblay (Eds*.), A cognitive approach to situation awareness: Theory and application* (pp. 137–154). Hampshire, UK: Ashgate.

Ebbage, L. & Spencer, P.D. (2004). Airmanship training for modern aircrew. Fort Belvoir, VA: Defense Technical Information Center.

Emerson, S. N., Halverson, K. C., Rizzardo, C., Knight, R., Brown, J., Reinert, A., Hoelscher, M. G., Schmidt, T., Tripp, L. & Mills, D. (2022). Pilot training transformation: Early results and lessons learned. *Proceedings of the 2022 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC),* Orlando, Fl.,

Endsley, M.R. (1987). SAGAT: A methodology for the measurement of situation awareness (NOR DOC 87-83). Hawthorne, CA: Northrop Corporation.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society 32nd annual meeting*. Santa Monica, CA: Human Factors Society, 97–101.

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

Endsley, M.R. (1999). Situation awareness in aviation systems. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Human factors in aviation systems* (pp. 257-276). Hillsdale, NJ: Lawrence Erlbaum.

Endsley, M.R., & Garland, D. J. (2000). Pilot situation awareness training in general aviation. *In Proceedings of the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 357–360). Santa Monica, CA: Human Factors and Ergonomics Society.

Fornette M.P., Bardel, M.H., Lefrançois, C., Fradin, J., Massioui, F.E., & Amalberti, R. (2012). Cognitive-Adaptation Training for Improving Performance and Stress Management of Air Force Pilots. *Int J Aviation Psychol*, *22(3)*, 203–223.

França, J., Oliveira, A., Silva, L., & Karlsson, P. (2023). Analyzing non-technical skills in the sharp end of facilities/utilities operations in onshore and offshore O&G process plants. *Environment Systems and Decisions, 43*, 251–264

Hamlet, O.E.D., Irwin, A. & McGregor, M. (2020) Is it all about the mission? Comparing non-technical skills across offshore transport and search and rescue helicopter pilots, *Int. J. Aer. Psych*. 30 (2020) 215–235,

Hamlet, O.E.D., Irwin, A., Flin, R., Sedlar, N. (2023). The development of the helicopter non-technical skills (HeliNOTS) behavioural marker systems. *Ergonomics. 2023, 3*, 1–10.

Hartel, C. E., Smith, K., & Prince, C. (1991). Defining aircrew coordination: Searching mishaps for meaning. Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH

Helmreich, R. L., Merritt, A. C., & Wilhelm, J. A. (1999). The evolution of crew resource management training in commercial aviation. *The International Journal of Aviation Psychology, 9*(1), 19–32

Heycock, S. A., & Brown, M. F. (2000). The Subjective Objective Assessment of Airmanship. *Proceedings of the 2000 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC),* Orlando, Fl.,

Irwin, A., Taylor, S., Laugerud, E., & Roberts, D. (2016). Investigating non-technical skills in Scottish and English aircraft maintenance teams using a mixed methodology of interviews and a questionnaire. *The International Journal of Aviation Psychology, 26(3–4)*, 105–119.

Kern, T. (2010). Foundations of Professional Airmanship and Flight Discipline. Convergent Performance. Colorado Springs, Colorado.

Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences, 99*, 1-13.

King, N. (2004). Using templates in the thematic analysis of text. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organizational research* (pp. 257–270). London, UK: Sage.

Mane, A. (1981). Airmanship: An introduction. First symposium on aviation psychology, 1981. *Proceedings of the symposium on aviation psychology*, 161-166. Columbus, Ohio.

Marcu, I.I., & Ivan, B.M. (2024). Agile combat employment: A paradigm shift in modern warfare. In *International Scientific Conference Strategies XXI*, 20, 308-314)

Ministry of Defence (2018). Combat air strategy: an ambitious vision for the future. https://www.gov.uk/government/publications/combat-air-strategy-an-ambitious-vision-for-the-future

Ministry of Defence (2025). Strategic Defence Review Making Britain Safer: secure at home, strong abroad. https://www.gov.uk/government/publications/the-strategic-defence-review-2025-making-britain-safer-secure-at-home-strong-abroad

Mishler, A., Severe-Valsaint, G., Natali, M., Seech, T., McCoy-Fisher, C., Cooper, T. & Astwood, R. (2019). Project Avenger Training Effectiveness Evaluation. *Technical Report: Naval Air Warfare Center Training Systems Division - NAWCTSD Public Release 22-ORL006*.

Mitchell, D. G., Doman, D. B., Key, D. L., Klyde, D. H., Leggett, D. B., Moorhouse, D. I., Mason, D. H., Raney, D. L., and Schmidt, D. K., (2004). Evolution, Revolution, and Challenges of Handling Qualities. *Journal of Guidance, Control, and Dynamics, 27 (1)*, 12–28

Nergård, V. (2014). Airmanship - a qualitative approach. *Aviation, 18(3),* 147-156.

Penney, H. (2025). *Want combat airpower? Then fix the air force pilot crisis.* The Mitchell Institute for Aerospace Studies Air & Space Forces Association Arlington, VA

Prince, C., & Salas, E. (1998). Situation assessment for routine flight and decision making. *International Journal of Cognitive Ergonomics, 1(4)*, 315- 324.

Royal Air Force Central Flying School (2025). Handbook of Military Aircrew Instruction: Vol 1 – Aircrew Instructor Theory Foundation. *Internal RAF CFS manual: unpublished.*

Salas, E., Wilson, K. A., Burke, C. S., Wightman, D. C., & Howse, W. R. (2006). A checklist for crew resource management training. *Ergonomics in Design, 14*, 6–15.

Sarter, N., & Woods, D. (1991). Situation Awareness: A critical but ill-defined phenomenon. *International Journal of Aviation Psychology*, *1(1),* 45-57.

Tsifetakis, E., & Kontogiannis, T. (2019). Evaluating non-technical skills and mission essential competencies of pilots in military aviation environments. *Ergonomics, 62*, 204-218

Vedula, P., Tingstad, A., Menth, L., Mehta, K. R., Roberts, J., Guffey, R. A., Crawford, N.W., Bemish, B. A., Payne, R., & Schuh, E. (2023). Outsmarting Agile Adversaries in the Electromagnetic Spectrum. *RAND Corporation*, Santa Monica, CA, USA:, 2023

Walker, G.H., Stanton, N.A., Kazi, T.A., Salmon, P.M., & Jenkins, D.P. (2009). Does Advanced Driver Training Improve Situation Awareness? *Applied Ergonomics, 40*, 678-687

Walshe, N. C., Crowley, C. M., O'Brien, S., Browne, J. P., & Hegarty, J.M. (2019). Educational interventions to enhance situation awareness: A systematic review and meta-analysis. *Simulation in Healthcare, 14(6)*, 398.

Wickens, C. D. (2008). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors, 50,* 397–403.

Yule, S., Flin, R., Maran, N., Rowley, D., Youngson, G., & Paterson-Brown, S. (2008). Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. World J Surg, *32*, 548-56