Applied Ergonomics 127 (2025) 104516

Contents lists available at ScienceDirect

Applied Ergonomics

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

Avoiding automation surprise: Identifying requirements to support pilot intervention in automated Uncrewed Aerial Vehicle (UAV) flight

Ben Grindley [a,b,*], Katie Phillips [b], Katie J. Parnell [a], Tom Cherrett [a], James Scanlan [a], Katherine L. Plant [a]

[a] Transportation Research Group, Boldrewood Innovation Campus, University of Southampton, UK
[b] Defence Science and Technology Laboratory (Dstl), Salisbury, UK

ARTICLE INFO

Keywords:
Uncrewed air vehicle
Drone
Decision-making
Interview
Automation
Ergonomics

ABSTRACT

The breadth and depth of Uncrewed Aerial Vehicle (UAV) operations are expanding at a considerable rate. With expansion comes challenges for the design of automation to support decision making. This research takes the perceptual cycle model (PCM) and the derived trust version of the Schema World Action Research Method (T-SWARM), to identify the issues and challenges of pilot intervention in UAVs operating during highly automated states. Nine UAV pilots with current experience operating medium to large UAVs were interviewed, using T-SWARM, about incidents in which they initiated an intervention in system operation (i.e. to avoid weather or collision) and an event where the system initiated the intervention (i.e. due to system failure). The coded responses highlighted the challenges with what information is displayed, how it is displayed and how it influences decision-making in the UAV context. In addition, the responses also identified aspects that influence trust in the system, including personal disposition, affect interventions with the automation. Against each of the key factors identified recommendations are made to increase safety and operational efficiency of UAV operations. This research adds to the growing body of literature that supports the application of T-SWARM for eliciting knowledge in the aviation domain and specifically within the UAV domain.

1. Introduction

The amount and type of Uncrewed Aerial Vehicle (UAV) operations have become increasingly accessible not just to military units but businesses and hobbyists (Buissing, 2018; Mohsan et al., 2022). The technologies facilitating the use and ownership and operation of UAVs is both decreasing in cost and increasing in capability (Gupta et al., 2021). Whilst a popular area for development, several human factors challenges are exposed when operating UAVs compared to crewed flight (Grindley et al., 2024a,b; Hobbs, 2010; Hobbs and Lyall, 2016; Hobbs and Shively, 2013; Kaliardos and Lyall, 2015). Notably, the operator is not in the system, which limits tactile, vestibular, audible and high-fidelity visual cues resulting in additional effort in managing and maintaining situational awareness (SA). UAV sorties or flight profiles, particularly for medium and large commercial and military systems, can often be monotonous and fatiguing with automated systems relied upon in most phases of flight control (Hobbs and Lyall, 2016). Operators must be able to use the information and trust the data going to, and coming from the UAV, and trust that the automation is going to behave as they expect (Parnell et al., 2023). The role of automation reliability plays a large part in the trust that people have in UAVs, their subsequent reliance on the automation and how this influences decision-making, workload and SA when human intervention is required (Lee and See, 2004; Ruff et al., 2004).

By law in many countries UAVs are still required to be ‘commanded’ by a human operator, an individual who is legally responsible for the air platform and may be required to intervene in the case of unplanned events (EASA, 2015; R. R. Murphy, 2014; Murray and Chu, 2015; Sah et al., 2021; Stöcker et al., 2017). Whilst some aspects of UAV operations require close human control, there are many that necessitate higher levels of automation to accomplish complex goals and tasks (Davies et al., 2018). This is particularly the case during periods when the control station(s) cannot communicate with the air platform with sufficient bandwidth (i.e. Beyond Visual Line of Sight; BVLOS), or to

* Corresponding author. Transportation Research Group, Faculty of Engineering & Physical Science (FEPS), Boldrewood Innovation Campus, University of Southampton, SP16 7QF, UK.
E-mail address: b.grindley@soton.ac.uk (B. Grindley).

https://doi.org/10.1016/j.apergo.2025.104516
Received 25 October 2024; Received in revised form 22 January 2025; Accepted 29 March 2025
Available online 3 April 2025

compensate during system failure, or periods of high task load (H. Chen et al., 2009). This increased level of automation, whilst designed to support the operator and reduce workload presents unique challenges. It is these challenges that must be overcome to improve the utility of UAVs as capabilities increase. Breakdown in the interaction between human operators and automated systems, can result in what has been described as automation surprise (Sarter et al., 1997; Wiener, 1989) this phenomena can become apparent during interventions between the human and the system/agent and vice versa. This can leave the operator questioning what the system is doing, why it did it, what it did, and what it plans to do next. Sarter et al. (1997) and Bainbridge (1983) describe how workload was not necessarily reduced when using automation, rather that it places new attentional and knowledge demands on the operator. This can be in relation to how the system design can invoke breakdowns in mode awareness, increase complacency and trust and requires new approaches to coordination and training, all of which are influential in transitioning between automated control states and human operator control states. This research sets out to understand the key factors that influence decision making during intervention on automated flight, from the perspective of the operators themselves.

1.1. Perceptual cycle model and SWARM

The Perceptual Cycle Model (PCM; Neisser, 1976) proposes a model of how "schema", a template in which we formulate mental representations of the world (Bartlett, 1995; Chalmers, 2003), can be used to guide behaviour. It suggests that prior training, experience and knowledge is entrenched in a cyclical relationship with an individual's actions and their environment which is continually updated and adapted. Whilst similar to the concept of sensemaking (Attfield and Baber, 2017; Klein et al., 2007), the PCM proposes that how individuals sample the environment combined with their existing preconceptions can influence actions taken; also how subsequent information within the environment influences how an individual perceives and gains understanding within different circumstances. Importantly for this research the PCM places emphasis on understanding the processes involved in decision making within the wider system (Banks et al., 2021; Plant and Stanton, 2015), rather than focussing entirely upon decision output. Appreciating the role of the wider system in the decision making process is critical for a greater understanding on how individuals can be supported within the environment they are working in (Plant and Stanton, 2012; Stanton et al., 2009). The PCM is an effective framework for gaining an understanding of all factors impacting UAV operator decision making, from the training that pilots receive, the information that they are presented as part of the control system as well as the actions that they take. The PCM has previously been used in retrospective analysis of incidents or accidents to examine decision making processes (Banks et al., 2018; Plant and Stanton, 2012; Stanton and Walker, 2011). However, applying the PCM can be challenging (Plant and Stanton, 2013). In order to apply it in a structured manner to draw out useful information the Schema World Action Research Method was developed (SWARM; Plant and Stanton, 2016). SWARM was designed specifically to understand aeronautical critical decision making, in order to elicit information about how an individual's schema, actions and world (their environment; SAW) influence decision making. SWARM assists the collection of data from subject matter experts (SME) in relation to the three categories of the PCM; SAW. Each of these has sub-themes which are clearly articulated in Plant and Stanton (2016). In order to elicit the information from SMEs an interview protocol with prompts against each sub-theme were produced; appendix 1. Following its initial development it was validated against critical incident aviation where the method demonstrated high theoretical validity and high test–retest reliability in commercial aviation (Plant and Stanton, 2016) it has subsequently been further developed and applied to understand UAV pilot decision making, including swarm technologies, in order to develop requirements (Parnell et al., 2023) and in the design of avionics systems (Banks et al., 2021; Parnell et al., 2022). In addition to the existing SAW prompts, Parnell et al. (2023) adapted the SWARM prompts with the addition of trust (T) questions from relevant trust scales, and demonstrated that T-SWARM is able to elicit factors within an individual's schema, behaviours and from the world that influence or are influenced by trust. Given the inherent remote and automated nature of UAV operations there is a requirement for pilots to trust the systems they are operating (Mouloua et al., 2019). It must be considered at the system level and as an emergent property arising through the interaction between system elements, personal traits or prior experience (Lee and See, 2004). The addition of these questions allows researchers to explore how trust is formed and maintained in relation to the influences of the wider system as captured by T-SWARM.

Using the PCM and T-SWARM, this research aims to explore the underlying mechanisms of UAV pilot decision making during automation intervention. Whilst this has been conducted across other domains (see Grindley et al., 2024a), there is little literature specifically considering UAV pilot decision making during automation intervention. Whilst previous published research has been conducted on UAV pilots using similar interview methodologies (Alon et al., 2021; Christ et al., 2016; Jenkins, 2012; Ljungblad et al., 2021; Steen et al., 2024) to the authors knowledge this the first time a knowledge elicitation method has been used to draw out the factors that influence decision making around interventions from specific events and accounts that operators recall. This will facilitate identification of improvements to operational safety, efficacy and efficiency through exploring the gamut of Human Factors (i.e. training, design, recruitment), a challenge given the relatively few and hard to reach operators. Through semi structured interviews with UAV operators and examining their own experiences, a user centred approach was taken to understand the factors that support and implicate the challenge of transfer of control between higher levels of automated flight and pilot-controlled flight due to system and operator intervention. It will identify challenges being faced by current operators and present design requirements to support future UAV development.

2. Method

2.1. Participants

Participants were sought who had experience with UAV's for both military and civilian (industry or academic) purposes which would fall into the Civilian Aviation Authorities' *Specific* or *Certified* categories and therefore:

- have a characteristic dimension (wingspan or length) of 3m or more; and/or
- designed for transporting people; and/or
- designed for the purpose of transporting dangerous goods and requires a high level of robustness to mitigate the risks for third parties in case of an accident.

To take part, participants were required to be over 18 years old and hold a UAV operator/flyer qualification (i.e. CAA A1/2/3, FAA Part 107, RAF Remote Pilot Qualification), there was no limit on flying hours but must have been able to recall an event where they or the system initiated an intervention. Nine participants took part, 8 male and 1 female, all had experience flying fixed wing BVLOS UAVs. Of the nine participants, five were serving in the Royal Air Force (RAF) and four operating UAVs for commercial organisations. In total they had logged an average of 527 flying hours operating medium or large UAVs (range 100–1000 h) in addition some participants had experience operating as pilots in crewed fixed wing ($n = 6$) or rotary wing ($n = 1$) aircraft (average flying hours = 302, range 55–1500).

2.2. Interview questions

Interview questions were based on a set previously used in the UAV domain for the development of T-SWARM (Parnell et al., 2023). Each theme (SAW) includes several interview prompts that allow interviews to be conducted with operators to extract information aligned to the PCM (Niesser, 1976), the interview protocol can be found in appendix 1. The original SWARM includes a repository of 95 prompts, with the intention that researchers down select them based on the goals of their research project. To keep the interview open to all contributory factors, all probe questions were retained and irrelevant questions (such as where the operator was located) were not asked.

The interviews were conducted, recorded and transcribed on Microsoft Teams (MSTeams). The transcripts were downloaded and corrected by the primary researcher who listened and watched the audio recording back and amended the output. The video files were subsequently deleted, and the transcripts were anonymised. Nvivo 14 was used to qualitatively analyse the data.

2.3. Procedure

Participants were invited to take part via e-mail, from companies supporting the UK Department for Transport, Solent Transport's Future Transport Zone project and relevant UK military units who operate medium and large UAVs. Interviews were arranged for a time that suited the participant and an invite to Microsoft Teams shared; participants were asked to undertake the interview in a quiet and private place where they were comfortable talking about their experiences. The primary researcher opened the interview with a brief overview of the research and participants were then offered the opportunity to ask questions and provide consent. Participants were then asked demographic questions before being asked to briefly describe a scenario or event they had experienced (which did not result in a negative outcome, i.e. a crash) in which either they initiated an intervention during automated flight or where the system had initiated intervention whilst they were piloting/operating the UAV. On completion of the initial scenario or event description the interviewer asked the probe questions as per the T-SWARM methodology. Participants were then given a short break before they were then asked to describe the second event or situation they had experienced and asked questions on the other scenario. The total time for each of the participant interviews, which consisted of discussion and questions of two events, was around 1 h 30 min (40 min each with 10 min for introduction and a short debrief). At the end of the interview participants were offered a debrief and thanked for their time. The research received ethical approval from the University of Southampton ethics board (Ethics ID: ERGO 89015) and a favourable opinion from the Ministry of Defence Research Ethics Committee (2270/MODREC/23).

2.4. Data analysis

The primary researcher coded the interview transcripts against the SWARM themes and codes outlined in the SWARM handbook by Plant and Stanton (2016), the codes are in appendix 2. A second researcher then coded approximately 10 % of the transcribed references to check for inter-rater reliability against the same codes, following the guidance of (O'Connor and Joffe, 2020). The second researcher had over 10 years' experience in applied human factors and is a Member of the Institute of Ergonomics and Human Factors. They were provided with excerpts of the transcripts and the SWARM code book and were asked to use it to code the excerpts. Both the primary researcher and the secondary coder initially coded independently before meeting together to discuss their codes. Hruschka et al. (2004) notes that with a large number of codes, such as with SWARM, there is a high likelihood of lower levels of agreement. The inter-rater reliabilities of the codes calculated on a reference-by-reference basis were assessed using Cohen's Kappa (Cohen, 1960). A moderate agreement can be considered when the κ coefficient is between 0.40 and 0.60 and substantial agreement when k above 0.60 (Landis and Koch, 1977). The overall κ coefficient for the two coders in this study was 0.51 thus providing support for sufficient inter-coder agreement.

In addition to the SWARM codes, additional trust codes relating to positive and negative trust were created, in a similar manner to (Parnell et al., 2023). This allowed a discussion of the factors that influenced trust in the UAVs during these events in addition to the aspects of SA and decision making afforded by the SWARM codes. Where transcribed references did not fit in, the SWARM or trust additional codes were created and described using an inductive grounded approach as per Parnell et al. (2023).

3. Results

Each of the participants discussed two scenarios or events with the majority of the recorded aspect of the interviews (when the participants discussed their events) lasting about 65 min, with the shortest duration being 48 min and longest 77 min. One where they had initiated an intervention in highly automated flight and one where the system had initiated or indicated a requirement for intervention from the crew. Table 1 outlines the types of events described by the participants by intervention type.

In all the scenarios, the pilots were operating as part of a crew with either payload operators, safety/external pilots and on occasion a pilot in command, the UAVs were either conducting flight testing (3) or transiting to or from operational Intelligence, Surveillance and Reconnaissance (ISR; 6) missions.

3.1. SWARM coding

Using Nvivo14 all the interviews were coded against the SWARM codes (Plant and Stanton, 2016) and two additional codes on negative trust and positive trust, whilst there is more to trust than just positive and negative the research was interested in the factors from SAW that influenced trust so kept the trust codes at a high level. The frequency of reference to each SWARM code is shown in Fig. 1, hierarchically arranged by frequency within the SAW themes. The highest number of references were associated with the World theme (n = 1032, 27.3 % of transcript coverage) followed by Action (n = 679, 22.9 % of transcript coverage) and Schema (n = 277, 12.1 % of transcript coverage), each participant transcript was in the region of 5000–10000 words. When comparing the system-initiated intervention events and the operator-initiated events the total number of references was considerably less (albeit not significant) for the system-initiated events (n = 845, $M = 93.89$ $SD = 32.61$) than the operator-initiated events (n = 1136, $M = 126.22$, $SD = 47.26$), $t(8) = 1.78$, $p = .056$. This difference may be down to fewer inputs being required for detection, diagnosis and resolution when the system initiates intervention. However, proportionately there was very little difference in the frequency of references associated with the SWARM themes or the codes. The biggest proportionate difference in codes was in the *natural environment monitoring* during the operator-initiated interventions. This occurrence is most likely as a result of the large number of described events being about avoiding

Table 1
Frequency of type of event described by intervention type.

Intervention type	Type of event described	Frequency
Operator-initiated intervention	*Avoiding poor emergent weather*	*3 (P2, P4, P6)*
	Dealing with an emergent failure with the air vehicle that the system had not yet reported	*3 (P1, P9, P3)*
	Correcting a user input error	*2 (P5, P7)*
	Avoiding other airspace users	*1 (P8)*
System-initiated intervention	*Air data sensor failure*	*4 (P2, P4, P5, P6)*
	Generator failures	*1 (P1)*
	Engine failure	*1 (P9)*
	Battery sensor failure	*1 (P3)*
	Inertial measurement unit failures	*1 (P7)*
	Communications link failure	*1 (P8)*

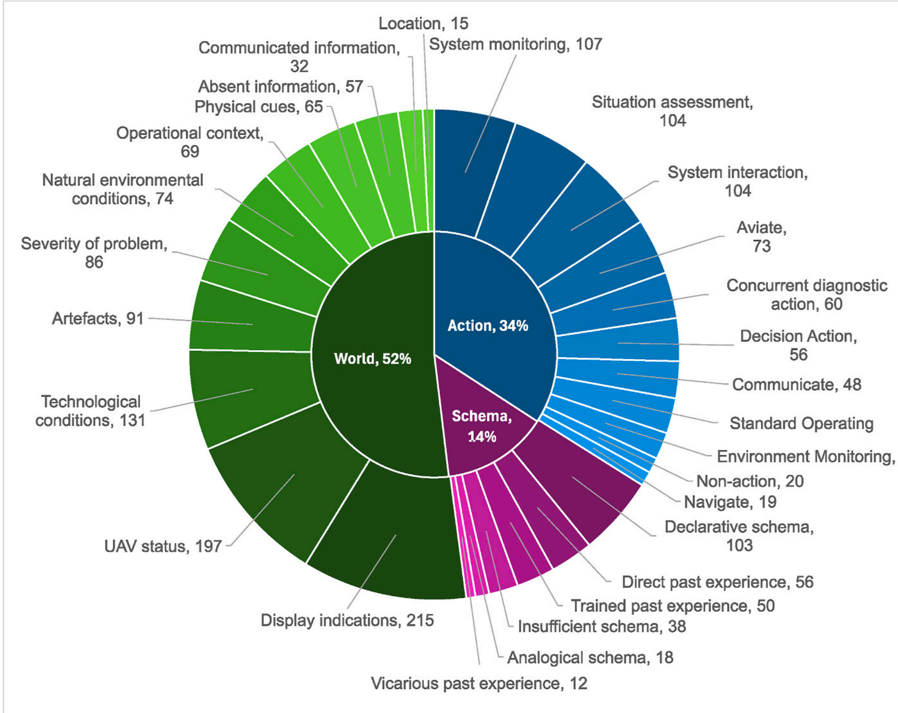

Fig. 1. Frequency of SWARM code occurrence.

weather. During the system-initiated interventions *concurrent diagnosis,* where the participants regularly discussed trying to identify the cause of the issue the UAV has presented in order to choose a course of action, had the biggest proportionate difference.

3.2. Identifying key factors associated with intervention

Due to the large number of SWARM codes referenced in the interviews, the most frequently discussed factors in relation the

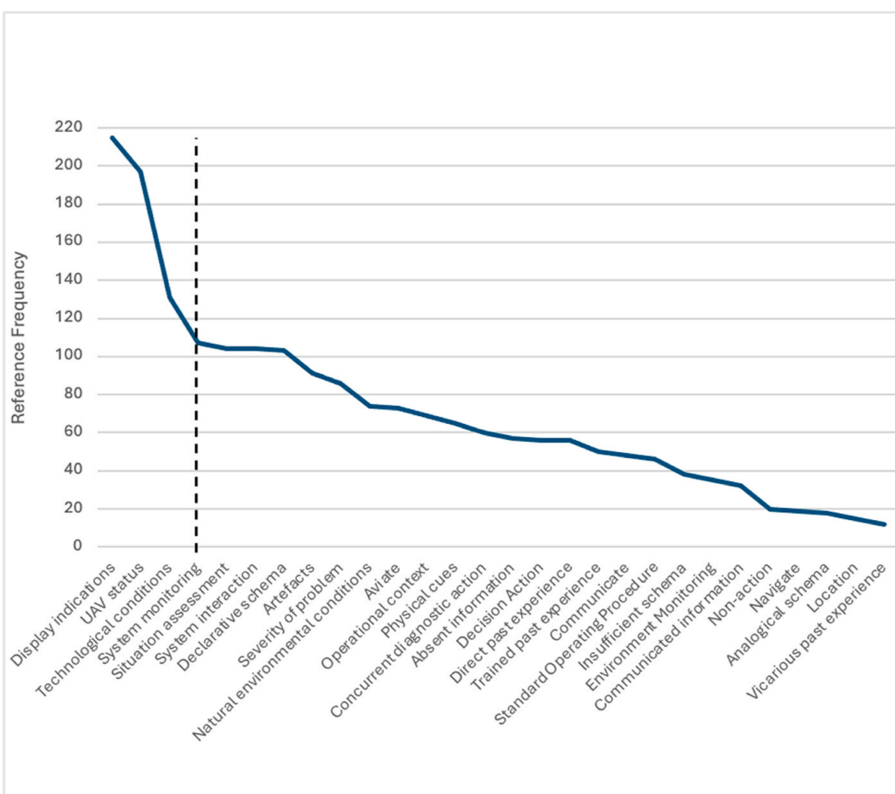

Fig. 2. Scree plot outlining the most frequently referenced SWARM factors associated with automation intervention.

intervention scenarios needed to be derived. A method used by Parnell et al. (2023) to determine the number of factors to retain and present in the discussion was followed. This process uses a graphical scree plot, following the principles outlined in Cattell (1966) in which the frequency of factors are ranked and presented as a line graph from most frequent to least. The factors or components to the left of the point at which the frequencies start to level off is identified as "the elbow" and are retained for discussion (Parnell et al., 2023; Parnell et al., 2016; Rafferty et al., 2010). This can be seen in Fig. 2 and resulted in 4 factors being retained: *Display Indications, UAV Status, Technological Conditions, and System monitoring,*

The most commonly referenced code across all of the interviews, and outlined in Fig. 1, was *Display Indications.* This should come as no surprise to those familiar with UAV operations as display indications are the primary method for the pilot to build and maintain SA. References to *Display Indications* were closely followed by *UAV Status* and *Technological Conditions* with operators regularly commenting on the what the UAV was doing at the time, and common 'gripes' or explanations for why they were able or unable to detect issues or form an assessment of the situation. An *action* item then followed in *Situational Monitoring.* Table 2 describes challenges identified by the participants, against each one a recommendation or design requirement is provided. These recommendations can broadly be grouped into the following areas that will be discussed further in the discussion: visual information display, alternative modalities for information presentation, training to support system knowledge and trusting, understanding individual influences on trust in

Table 2
Design requirements and recommendations as a factor of the most frequently referenced SWARM codes.

SAW Taxonomy	SAW Theme	Freq	Emergent challenges within code	Recommendation/design requirement	Recommendation topic
Display Indications	*World*	*215*	*• Displayed visual information is often the only source of current information about the aircraft as operators often don't have haptic or audible cues to support understanding of whether the aircraft condition has changed.*	*R1 - Information provided to the operators on displays must be unambiguous, accurate and timely to allow the operator to understand the current situation and predict future states.* *R2 - Designers should consider alternative modalities for concurrent information presentation (i.e. aural/haptic)*	*• Visual information display* *• Alternative modalities of information presentation*
			• Issues are often detected and determined through information candidly provided to the pilot through warnings and cautions. If thresholds are not met detection can be hampered.	*R3 - Displays should be designed to enable to operator to detect latent issues and trends in the system status that support detection of issues prior to parameters being met for system failures.*	*• Visual information display*
			• Displayed information changing or not changing supporting issue detection.	*R4 - Where appropriate, visual information displays should support operators to detect changes rather than just status to support trend analysis. Numbers and text are challenging to detect changes and monitor trends (Harris, 2004).*	*• Visual information display*
UAV Status	*World*	*197*	*• Operators reported a primary means for detecting a change in UAV status is when a conflict arises between the expected situation and actual (i.e. aircraft descending when they expect it to remain level).*	*See R4* *See R2*	
			• Some participants, particularly those from crewed flight backgrounds, highlighted not trusting the UAV to support itself and continue as planned when issues arise and therefore required intervention.	*R5 – training must consider the levels of automation design and resilience during failures, to build familiarisation around aircraft responses during gradients of sensor, communications, servo and flight control failures.* *R6 – further, research should consider sociocultural impacts on trust in automation.*	*• Training for system knowledge and building trust* *• Individuals' differences of trust in automation*
			• Operators reported that not having physical cues to understand status can impact course of action selection, the contrary is safety pilots closer to the platform reported being able to hear and see changes to the platform and informed course of action selection.	*See R2,*	
Technological conditions	*World*	*131*	*• Data staling or incorrect sensor information can be inconspicuous to the operator but can hamper or even derail threat detection.*	*R7 – system design should implement methods to detect and report to the operators old, stale or incorrect data using methods such as redundancy checks and providing system and communications status.*	*• Visual information display*
			• Some systems required displayed information to be compared against a mental model of expected parameters (i.e. that they are still within limits or how close they are to limits).	*R8 – systems should not rely on operator memory to determine issues with the system but present safe parameters and support operators' detection.*	*• Visual information display*
			• Participants reported that information being displayed that was required to maintain safe flight was presented in a manner that inhibited understanding due to location.	*R9 – information critical to the pilot to maintain safe flight should be located based around an expected eye position and prioritised based on frequency of use, severity of issue if not detected, speed of access required, following human engineering guidance such as ISO 921.*	*• Physical interface design*
System monitoring	*Action*	*107*	*• A challenge was commonly discussed regarding pilots' need to monitor displayed information at all times to detect particular issues (i.e. incorrect altitudes being set), if not looking at that information, not remembering the prior status of the information or distracted by other events this can be missed.*	*See R2, R3 and R9*	
			• Operators identified a challenge associated with a need to become familiar with a set picture on the displays so that they can use pattern matching to identify changes.	*See R9* *R10 - Training should not just focus on non-typical events and gradients of situations to support identification of events through cues and artefacts rather than specific events.*	*• Training for system knowledge and building trust*
			• Interviews highlighted that continually monitoring systems is fatiguing, and that if monitoring drops off it will quickly increase once triggered by an event.	*R11 – Physical system design, including the design of standard operating procedures, must consider physical and mental fatigue associated with persistent system monitoring and mitigate both underload and overload. Solutions might include tools to maintain vigilance, automation design to manage self-monitoring and operator trust.*	*• Physical interface design*

automation and physical interface design.

3.3. SWARM factors associated with trust during intervention

To identify SWARM factors associated with trust two codes were used, a ‘positive trust’ code was used where references associated with SAW themes were linked with increasing or maintaining trust and a ‘negative trust’ code which was used when the participant referenced or associated with reducing or maintaining poor levels of trust in the UAV. Table 3 presents the SWARM factors that had greater than 3 references associated with positive trust, Table 4 with negative trust.

During the coding process two references were made about the events or scenarios that did not fit with the existing SWARM prompts, they were primarily around the *operators' physical* and mental state. The first was related to *Comfort* and refers how an individual described feeling about the situation, commonly this is in conjunction with existing SAW codes (i.e. comfort from having previously experienced a similar situation) for example *“I'm, you know, more likely to go back to fully manual control”* (P3) or “*I feel quite comfortable there*” (P2). The other additional area was titled Fatigue and relates to comments on the participants having feelings or perceptions of personal fatigue or tiredness; an example reference might be “Especially when you're on a low circadian rhythm cycle and you're maybe not fully paying attention to every bit of Information” (P7). SWARM doesn't currently account for the psychophysiological condition of the decision maker and how this might influence course of action selection.

Table 3
SWARM factors associated with positive trust.

Factor	Number of references	Description	Example references
Direct past experience	14	*It appears from the data that when the participants had experienced the system working or experienced it working following a failure this improved their trust in the automation.*	*“it's essentially just knowledge and experience and just looking back on the scenario and thinking about what happened and why it happened that then repaired the trust in the system.” (P4)*
UAV status	11	*Participants reported that their trust in the system remained positive if they were able to determine the status of the UAV especially when it was still flying, whilst they may have had to intervene if it was still flying they could be more trusting.*	*“So with fixed wing [platforms] like if it loses an engine, you can still glide back. Whereas with a multirotor it loses a rotor. You lose that aircraft.” (P8)*
Declarative schema	6	*References to facts about the system, knowledge of the system were often aligned to positive trust.*	*“I've trusted the system, you know, we already knew our fail-safe heights were pre-planned in. I knew the airbag and the parachute was going to deploy. You know we've lost an engine”(P3)*
Display indications	6	*Related to the UAV status, participants reported having a positive indication that provided system information supported trust in the system even though it may have had to intervene. The information meant they were able to determine what was happening and how to manage the situation and therefore retain trust in different parts of the system.*	*“Yeah, but kind of by seeing that my indications were mostly normal, I could trust it.” (P2)*

Table 4
SWARM factors associated with negative trust.

Factor	Number of references	Description	Example references
Insufficient schema	17	*Participants commonly cited not having enough knowledge of how the system or automation operates driving feelings of distrust or trust reduction.*	*“speed or height holds can catch you out because it's doing the opposite to what you asked for and then the urge is to take control until you realize what it's doing.” (P6)*
UAV status	16	*As the opposite to UAV status as a factor of positive trust, references were commonly found relating to not knowing the exact status or mode of the system negatively influenced trust or led operators to be less trusting.*	*“I would if you'd asked me at the time, do you trust what the aircrafts' doing? I'd say no, because I don't know what it's doing” (P4)*
Technological conditions	14	*How information is displayed/generated was found to negatively influence trust as individuals were concerned about not being able to determine system status in an unambiguous or timely manner.*	*“it depletes my trust that it doesn't verbosely like it doesn't succinctly show me and loudly tell me that it's [an event] occurred.” (P1)*
Display indications	10	*Similar to technological conditions as a factor influencing negative trust, the fact that the majority if not all of the information on the UAVs state is conveyed through displays caused some distrust due to the complexity and opportunity for failure through the system. This appeared particularly prevalent in operators with crewed experience.*	*“the information I'm getting is coming through a series of computers, conversions, operating systems, and then it's relayed to me on a on a display as opposed to a simple route and you know in a manned aircraft it would be from the back of the aircraft at the front.” (P6)*
Absent information	9	*Not having information that they thought they should have was reported as a factor negatively influencing trust, particularly after a significant event. (i.e. a warning should have gone off).*	*“the only part I didn't trust was the initial bit before we understood [what the issue was] because we didn't receive a warning, which was weird.” (P9)*
System interaction	7	*System interaction manifested in the data alongside negative trust as an action participants took when they didn't trust the system to continue safety, often they reported ‘taking control’ when they felt uncertain or trusted what the UAV/automation was going to do next.*	*“if it's getting bad information, I don't necessarily trust it to do the right thing at the right time, so I'm going to I'm protect it by giving it, in this case, more airspeed” (P5)*

4. Discussion

Interviews with 9 UAV operators, on two separate events, were conducted to determine factors that influence or impact interventions in automated flight of medium to large UAVs. The goal was to identify factors, as defined by the SWARM, in order to propose requirements and/or recommendations to support the evolving operational and design community. In addition, the use of T-SWARM as applied to UAVs builds on Parnell et al. (2023) further demonstrating its applicability.

4.1. Practical contribution

Key factors that influenced or impacted automation intervention in UAVs were identified through applying the SWARM. The plot in Fig. 2 identifies the most important SAW factors by frequency of reference. This methodology highlights the importance of both the pilot's and crew's schema, the actions they take to understand and build SA and the information they are getting from the world primarily through sensors and via communication channels from the airborne platform. The interviews also highlighted the crews' inherent and situationally informed trust (both positive and negative) in the systems but also the requirement to trust the platforms they are operating and how this influences the actions they take to try and protect them. The most common codes from the transcripts and the challenges within them are summarised in Table 2 alongside suggested design recommendations. The design recommendations provided against each of the most common codes in Table 2 have been collapsed into a number of themes that are discussed in more detail in the following paragraphs. Collating the recommendations into themes allows for practical, and generalised recommendations to be made for further research or operational design improvements.

4.1.1. Visual information display

The display of information, particularly in aircraft cockpits has been the charge of many successful ergonomist, human factors engineer, human computer interface designer and user experience expert since the Wright brothers first flew with just a stopwatch, a tachometer and an anemometer (Chorley, 1976). With the ever-expanding amount of information available to the pilot and increasing complexity of operations in both crewed and uncrewed flight there has been a need to maximise the efficiency of displayed information so that the operator can interpret, process and use the information available to them to maintain safe flight, reduce automation surprise (Woods and Sarter, 2000) and maximise operational effectiveness (see Lovesey (1977) for an early review and Landry (2021) and Carroll and Dahlstrom (2021) for more recent commentary). With the multitude of options available to designers, focus has been to reduce workload and manage SA, providing crews with the right information at the right time in the right place. This often results in very similar displays of information due to the influence of international standards, fashion, user expectation and familiarity from crews. Many have adopted a 'dark/quiet' philosophy with a focus on dials, pointers and counters (Harris, 2017; Wickens, 2003). To optimise display of information in the UAV context we must go back to basics and follow human/user centred design approaches combined with the latest theory and expertise (Lee and Seppelt, 2012; Parnell et al., 2021; Parnell et al., 2023; Stanton et al., 2010). The output from the interviews highlighted that often-displayed information was falling short in a number of areas for UAV ground control stations when automation required intervention. Primarily, challenges were identified during automated operation regarding a lack of historical information to support trend analysis and projection (akin to automation surprise, Woods and Sarter (2000) and out of the loop phenomenon, Endsley and Kiris (1995)) leading to operators not being able to quickly identify the pace of degradation or detect issues before they occur. An area that has been widely researched with regard to visual presentation of information is the importance of transparency (Bhaskara et al., 2021; T. Chen et al., 2015; Chien et al., 2019; Sadler et al., 2016; W. Zhang et al., 2021) with results commonly indicating that presenting information in a transparent manner supports trust, decision-making time and certainty and reduces the requirement for verification but can also induce complacency. However, outside of transparency at the time of writing, very little recent published research could be found (with the exception of Li et al., 2020) that specifically considered how workload, SA or trust during automation monitoring is influenced by different information presentation approaches of the same information. Whilst this may be done by design organisations in the development of systems, little is published in the open literature. Future research should consider how different presentation of the same information influences workload, SA and trust during automated monitoring therefore managing trend assessment and reducing out of the loop phenomenon and automation surprise during intervention events.

4.1.2. Alternative modalities of information presentation

Wickens (2021), amongst others, have discussed two concepts of attention allocation: first that we have a filter that selects and admits channels of information from the environment to be processed (Broadbent, 1982, 2013); the second that attention is a resource that is allocated, selected and divided to enable information processing, limited by the demand of tasks or multiple tasks needing to be performed concurrently defining the limits of multitasking and implications of distraction (Lavie, 1995). If the latter is true, as a theory that has received much support in recent decades (see Lavie, 2005; Lavie et al., 2014; G. Murphy et al., 2016), then there is precedent that information from multiple modalities is attended to simultaneously, albeit potentially sub/pre-consciously. Whilst there may be a distracting element, benefits of concurrent information presentation in different modalities to support threat detection may be valuable (Klemen et al., 2009; Wickens, 2002). Interviewees regularly commented on either an absence of sound cues from the systems or the benefit gained from being able to hear the system (e.g. engine rpm, gearbox sounds or wind over the wings) when acting as an external pilot or from experience in crewed flight. This aligns to Tvaryanas (2004) who showed through eye tracking that engine scan was reduced in the absence of non-visual cues in UAVs. The addition of audio cues have been found to increase the pilot's ability to maintain a steady airspeed, altitude, and bank angle when no visual cues were present, but not up to the level of performance achieved when visual cues were available (Brungart and Simpson, 2008; Gröhn et al., 2004; Lyons et al., 1990) with similar systems being developed to support helicopter pilot trend analysis (Edworthy et al., 1995). Whilst it is necessary that auditory instrumentation does not interfere with an operator's performance and must not pose any safety issues the domain would benefit from practical investigation into the empirical benefit of acoustic system-based information to support trend analysis in UAV systems. Noting, that several operators interviewed in this study referenced the benefit gained in situations where they could hear the UAV and how that improved their decision making (with regard to reducing uncertainty or increasing error detection time) or where they couldn't hear it and negatively influenced decision making and intervention (for example increasing time required to problem solve). At the time of writing no published research was identified on the benefit (or perhaps implications) of acoustic system-based information aimed at supporting operators to maintain performance, error detection and problem solving in UAV systems that do not benefit from proximity to the air vehicle.

4.1.3. Training for system knowledge and building trust

Both in the *declarative schema* and *display of indications* codes themes emerged around how previous experience and the presentation of information appeared to negatively influence trust in the system. Enhancing a person's understanding of automation process and performance has been shown to improve individual's trust in the automation (Lee et al., 2004) and that a person's trust should be appropriately calibrated to match that of the automation's capabilities (Kazi et al., 2007; Lee and See, 2004; Verberne et al., 2015). Indeed trust repair, tempering and dampening have been shown to provide an interactive means of calibrating trust across sequential situations (De Visser et al., 2020). However, when the performance of the person and the automation are highly correlated, trust calibration does not matter. More generally, in some situations it might be appropriate for people to rely on automation, even when it performs worse on a task than they would, because relying on the automation enables them to shift attention to a more important activity. This was highlighted in our analysis by operators commonly refencing that they don't have a choice but to continue operating the system when issues occur in an automated state as they are not able to manually fly the UAV and often it is better at conducting particular tasks that the pilot would be with limited input (e.g. "If we lose link with the aircraft that that obviously then throws a spanner in the works, but overall the reliability of the aircraft and the system is high enough that we can trust in that" (P4)). An often proposed solution might be to provide explanations or visualisations that show what influenced the system's decision, although many non-experts find these unhelpful (Y. Zhang et al., 2020). In a recent review, Chiou and Lee (2023) suggested that future design of automated systems, particularly when considering high levels of autonomy should be focused on *trusting*, rather than *trust*; that is not to maximise trust, or to even calibrate trust, but to support a process of trusting. Chiou and Lee propose that designers focus on: Identifying where trusting and trust calibration are critical (Situations), providing improved system-level cooperation through the way information is conveyed (Semiotics), recognising aspects of automation responsivity that may influence trusting (Sequence) and promoting cooperative joint action and mutual trusting (Strategy). During the interviews several examples of system design that influence trusting were identified that could be considered in future system design. For example, the operators outlined *situations* where they understood the limits or capability of the system and therefore had the appropriate trust calibration "*You may leave the engine RPM [revolutions per minute] higher so that it's more responsive to enable it to add on power as required, and I suppose you yeah, you're taking something automation away from the aircraft by saying I don't trust you to manage your AOA properly today, so I'm gonna [sic] give you more speed*" (P5) or "*it's slow, it's stable, it's all those things that that long endurance drones are supposed to do*" (P1). Operators highlighted, knowledge around *situations* that supports trusting is developed by experience and importantly should be shared through regular lessons identified and briefing sessions e.g. "*many times during a calendar year we talk through what would happen if emergencies happened*" (P1). *Semiotics* were identified as weaknesses in the design of some of the UAVs operated for example *"Can I trust everything? Is my aircraft still in the air? What's lying to me? What's not lying to me?"* (P7). But they also highlighted the importance of health monitoring of displayed data "you can tell how healthy the data that you're getting is" (P8) but also the benefits of how information is presented for example "*So we have plots. So you can monitor certain characteristic of an airframe. So it just shows you like a graph in time and then the actual like the values and that it's goes values on the left and then the time frame at the bottom*" (P9) display designs like this, including transparent automation may allow the operator to detect drops in information quality and gross changes that might infer issues with the data source. The operators also referenced *sequencing* and how that influenced them trusting the system *"I've trusted the system, you know, we are already knew our fail-safe heights pre planned in. I knew the airbag and the parachute was going to deploy"* (P9). These *sequencing* type features were often reflected in clearly written standard/emergency operating procedures and a good knowledge of the automation design including fall-back or reversionary modes. Finally, aspects around *strategy* were discussed, primarily around building a culture or processes around trusting the automation to do what it is good at or for task that can only be done by the automation for example keeping the platform within the parameters set by the pilot e.g. *"… maintaining the flight path that I asked it to maintain, so in that moment I [can] sort of go eyes off what's going on outside [to focus on system management]" (P5) and "I was quite content to allow the protection feature to do what it's designed to do rather than completely take control back and panic*"(P2). In summary, systems discussed by interviewees were already found to have features that support trusting; however, it is likely trust, or the process of trusting could still be improved in future system design. Implementing a systematic user centred design process that includes considerations around user trust, building on positive design features such as those identified above and looks to reduce negative design features is fundamental to improved operational safety and efficiency.

4.1.4. Individuals' differences of trust in automation

As mentioned previously trust or trusting in the automation is important as it supports a human supervisor (operator) in their use of an automation subordinate. During a number of interviews participants reflected on their trust in the UAV and how if differed from prior experience, particularly that of crewed flight. As many have pointed out, trust in systems has many aspects, a person's disposition, propensity or individual differences being one of them (Hoff and Bashir, 2015; Lee and See, 2004; Marsh and Dibben, 2003). Specifically, the role of individual differences in the perception, adoption, and decision-making with technology has often discussed and researched (for example (Hoff and Bashir, 2015; Kraus et al., 2021; Miller et al., 2021; Parasuraman et al., 2014; Scholz et al., 2024) and posited that propensity to trust automation may be based on a combination of general personality variables and the individual's experience with the technology (Kraus et al., 2020; Scholz et al., 2024). Whilst the findings from the interviews is anecdotal there is merit investigating further the relationship between prior experience in crewed flight and trust in UAV flight. Additionally, whether specific personality traits that influenced an individual's desire to take up or be successful in crewed flight (Chang et al., 2018; Chidester et al., 1991; Glicksohn and Naor-Ziv, 2016) subsequently influences trust or trusting in the UAV. Understanding this concept could have implications in selection, training, design and assurance for UAV operations.

4.1.5. Physical ergonomics and interface design

Those familiar with human factors engineering will be well acquainted with the challenges of workspace design with a need to trade off limited space with large amounts of competing priority information and control requirements. It is well documented that poor ground control station design that doesn't align to human strengths or avoids the limitations of the operator leads to incidents and accidents (Grindley et al., 2024; Grindley et al., 2024a,b; Mohammed et al., 2022; Oncu and Yildiz, 2014; Tvaryanas et al., 2006; Williams, 2006). This challenge is and will become ever more present as systems strive to be more automated, autonomous, portable, universal and are being developed quicker. Many of the participants commented on poor positioning of information, design of interfaces or controls. In order to alleviate some of the risks associated with these design decisions human engineering standards have been developed such as Military Standard 1472H, International Organization for Standardization's ISO 921 and ISO 11064, BS EN ISO 6385. These standards attempt to ensure that principles such as Wickens and Carswell's (1995) Proximity Compatibility Principle, whereby information is located with control inputs are considered in the design of the system to ensure a safe, comfortable and efficient system. Additionally Human Factors process or Integration (HFI) guidance such as the United Kingdom's Joint Service Publication 912 (Ministry of Defence, 2014) and the US Human Systems Integration (Lacson et al., 2017)(HSI) have been developed to ensure that human performance factors are adequately considered during the system engineering process. Many development programmes do not sufficiently include HFI/HSI as an fundamental aspect of planning and execution and are therefore at risk of diminished user performance and total system performance (Lacson et al., 2017). No literature was found on the prevalence of these approaches being applied to UAV design; however, it is expected that if it is being conducted at all its scope and influence is likely limited and informal due to the size of most the companies and the pace at which UAVs are being developed. It is highly recommended that design organisations follow Human Factors integration processes and standards to reduce the risk of errors and issues occurring. UAV regulators should consider directing design organisations to apply appropriate methodologies, as poor ergonomic design as highlighted in a number of the interviews negatively impacts decision making the corollary being the potential for reduced safety and performance of the system.

The 5 areas of recommendation above align to previous research findings by Parnell et al. (2023) who used similar methods within the UAV domain, specifically around aspects such as equipment design and training for trust in the automation. However, this research, by focussing on specific occurrences and events was able to delve into specifics around HMI design and draw out anecdotal evidence that supports an argument for further research to improve the safety of UAV operations particularly around the use of automation. The recommendations on visual display interfaces, multiple modality information presentation and predisposition for trusting automation are offered to the research community to identify best practice and support design organisations to improve automation intervention. Recommendations on physical interface design and designing for trust should be used by those in UAV design now, be it by hobbyists, commercial or military facing organisations to reduce the negative aspects of automation intervention and automation surprise.

4.2. Methodological contribution

When using data from interviews, as with many qualitative methods, the output results must be treated with some care as they may be subject to bias and will only reflect the understanding and experience of those that participated and which they are willing to divulge (Lamont and Swidler, 2014). Whilst there was some variation in gender, age and experience a large amount of the population interviewed were either serving in the Royal Air Force or had prior experience in the military. The military has a wealth of experience operating, often exquisite and capable UAVs, therefore the experience may not be representative of all UAV operations and organisations. However, understanding what is good and what could be improved from experienced operators and organisations can be beneficial for smaller emerging organisations (Brown et al., 2024; Gore et al., 2023; Hoffman et al., 1998; Milton, 2010). Given the relatively small operator population, access to the population, the pace of developments and variance in designs methods like SWARM are critical for eliciting valuable information from small samples sizes that can be generalised to the wider domain. Using this methodology, the researchers were able to extract information dense accounts on specific automation interventions, which hasn't been done before and applied to SWARM. Previously SWARM had only been applied once to UAV operations (Parnell et al., 2023), this paper presents the first application of the SWARM taxonomy to UAV operator interactions during specific type of events, particularly intervention in automated flight. Specifically, the method was able to successfully elicit detailed information about how the environment, experience and actions influence SA and decision makers. Additionally, questions on trust were included in SWARM and once again provided clear emergent relationships between technological systems, the operators previous experience and the actions they take, from which recommendations are provided in the previous paragraphs. The findings support the continued use of the Trust- Schema World Action Research Model, T-SWARM (Parnell et al., 2023).

As Parnell et al. (2023) discussed, compared to crewed aviation, in general the UAV domain does not have the same level of regulation, focus on checklists or requirements for levels of training and certification, and therefore do differ in their design and influences. That is not to say that UAV operations are not safe, but that the requirement to achieve the same level of safety as crewed or commercial flight is not present (Grindley et al., 2024a,b) nor required, due to many not carrying passengers or flying over populated areas. Unlike Parnell et al. (2023) whilst there was plenty of variation within the events described the researchers noticed that the influences that participants discussed started to be repeated in the later interviews. Whilst it is unlikely that saturation had occurred due to the large variety of potential events that could be discussed, a broad spectrum of influences and experience was captured. The repetition may also have been a result of over half of the interviews being conducted with Royal Air Force MQ-9A Reaper crews and all being conducted on medium to large UAV operators rather than small UAVs where the greatest variation in operation is seen.

During the coding process it was noted by both the primary and secondary researcher that participants referenced 'condition of self' or 'psychophysiological' aspects that they believed were influencing their decision making or SA. Specifically, participants mentioned fatigue, workload and comfort, for example, "*especially when you're on a low circadian rhythm cycle and you're maybe not fully paying attention*" (P7), "*I was on the limit. I was on. I was at capacity.*" (P7), "*I was comfortable enough*" (P9). The impact of workload (Hockey, 1997; Ji et al., 2022; Orasanu, 2017; Recarte and Nunes, 2003; Vidulich and Tsang, 2012; Wang et al., 2023) and fatigue (Bendak and Rashid, 2020; Harrison and Horne, 2000; Jia et al., 2022) on decision making and SA is well documented. Specifically, literature outlines how fatigue and workload influences top down processes in the form of the use of biases and heuristics (Engle-Friedman et al., 2018; Raby and Wickens, 1994) and bottom-up in the form of factors such as scan reduction (Di Nocera, Camilli and Terenzi, 2007; Rainieri et al., 2021; Tole et al., 1983) which are discussed within the PCM. The model however does not currently account for the psychophysiological condition of the decision maker and how this might influence course of action selection, subsequently the SWARM does not account for or aim to explicitly elicit psychophysiological influences. This research was able to elicit some of the personal factors that influenced decision making due to the detailed accounts about specific events, where previous research has been more aligned to case studies and the human-system interaction there within rather than individual recollection of events. Future uses of (T-)SWARM and the PCM should consider how information could be collected on the relevant psychophysiological state of the individuals and how that influenced their decision making and actions. For example, whilst more in depth research is needed to identify the most appropriate areas, prompt questions could be included around interviewees perception of: their fatigue at the time; how well they had been sleeping; the operational tempo at the time; task load in the run up to the event; and importantly how they perceived this to affect the SAW factors identified by other prompts. Similarly, in regard to workload prompt questions could be asked on their level of perceived workload at the time, whether they were having to drop or miss out tasks, what they were or whether they believed they had spare capacity that could impact their ability to make decisions. Understanding psychophysiological impacts on decision-making is critically important in design of future systems and interaction with automation. Whilst we can attempt to control aspects such as fatigue and workload through policy and process the nature of operations and propensity for periods of high workload (H. Chen et al., 2009) means systems must be able to accommodate, through design, for these variations in operator condition and remain as safe as reasonably practicable.

5. Conclusions

As operations with UAVs expand in complexity and character the role of the human operator as commanders, supervisors and monitors of automation will continue to be a challenge. This paper aimed to understand the factors that influence and impact UAV operators' intervention in automated flight, utilising the T-SWARM, with the addition of questions on trust. The interviews highlighted the importance of what information is displayed, how it is displayed and how it is used to detect and diagnose issues with the UAV. Participants commented on how a lack of particular cues, specifically physical, influenced their ability to assess situations. The questions on trust drew out how the design of the system and perceived reliability drove the level of trust in the system and supported the process of trusting albeit in many cases they didn't have a choice but to trust the system. Similarly, that compared to previous roles piloting crewed aircraft trust was less in the UAVs they were currently operating. From these identified challenges recommendations could be made to improve operational safety through further research. Caution remains, however, on the application of traditional aviation methodologies to the UAV operator due to its relative infancy and limited standardisation or regulation. This research adds to the growing evidence of the utility of T-SWARM for eliciting requirements and knowledge within the UAV domain.

CRediT authorship contribution statement

Ben Grindley: Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Katie Phillips:** Writing – review & editing, Formal analysis. **Katie J. Parnell:** Supervision. **Tom Cherrett:** Writing – review & editing, Supervision, Funding acquisition. **James Scanlan:** Writing – review & editing, Supervision. **Katherine L. Plant:** Writing – review & editing, Supervision.

Copyright

Funding

This research received a grant from the UK's Department of Transport (Ref. RITM0369035).

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.apergo.2025.104516.

References

Alon, O., Rabinovich, S., Fyodorov, C., Cauchard, J.R., 2021. Drones in firefighting: a user-centered design perspective. Paper Presented at the Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction.

Attfield, S., Baber, C., 2017. Elaborating the frames of data-frame theory. Paper Presented at the 13th International Conference on Naturalistic Decision Making.

Bainbridge, L., 1983. Ironies of automation. In: Analysis, Design and Evaluation of Man–Machine Systems. Elsevier, pp. 129–135.

Banks, V.A., Allison, C.K., Plant, K.L., Parnell, K.J., Stanton, N.A., 2021. Using the perceptual cycle model and schema world action research method to generate design requirements for new avionic systems. Hum. Factors Ergon. ManufactService Ind. 31 (1), 66–75.

Banks, V.A., Plant, K.L., Stanton, N.A., 2018. Driver error or designer error: using the Perceptual Cycle Model to explore the circumstances surrounding the fatal Tesla crash on 7th May 2016. Saf. Sci. 108, 278–285.

Bartlett, F.C., 1995. Remembering: A Study in Experimental and Social Psychology. Cambridge university press.

Bendak, S., Rashid, H.S., 2020. Fatigue in aviation: a systematic review of the literature. Int. J. Ind. Ergon. 76, 102928.

Bhaskara, A., Duong, L., Brooks, J., Li, R., McInerney, R., Skinner, M., Loft, S., 2021. Effect of automation transparency in the management of multiple unmanned vehicles. Appl. Ergon. 90, 103243.

Broadbent, D.E., 1982. Task combination and selective intake of information. Acta Psychol. 50 (3), 253–290.

Broadbent, D.E., 2013. Perception and Communication. Elsevier.

Brown, O., Power, N., Shortland, N., Gore, J., 2024. Expertise in action: insights from naturalistic decision making (NDM). In: Philosophy, Expertise, and the Myth of Neutrality. Routledge, pp. 128–146.

Brungart, D.S., Simpson, B.D., 2008. Design, validation, and in-flight evaluation of an auditory attitude indicator based on pilot-selected music. In: Paper Presented at the Proceedings of the International Conference on Auditory Display (ICAD2008), Paris, France.

Buissing, N., 2018. Aviation law and drones: unmanned aircraft and the future of aviation. Air Space Law 43 (6), 655–657.

Carroll, M., Dahlstrom, N., 2021. In: Human Computer Interaction on the Modern Flight Deck, vol. 37. Taylor & Francis, pp. 585–587.

Cattell, R.B., 1966. The scree test for the number of factors. Multivariate Behav. Res. 1 (2), 245–276.

Chalmers, P.A., 2003. The role of cognitive theory in human–computer interface. Comput. Hum. Behav. 19 (5), 593–607.

Chang, M.-C., Lee, T.-H., Lung, F.-W., 2018. Personality characteristics of fighter pilots and ground personnel. Mil. Psychol. 30 (1), 70–78.

Chen, H., Wang, X.-m., Li, Y., 2009. A survey of autonomous control for UAV. Paper Presented at the 2009 International Conference on Artificial Intelligence and Computational Intelligence.

Chen, T., Campbell, D., Gonzalez, L.F., Coppin, G., 2015. Increasing Autonomy Transparency through Capability Communication in Multiple Heterogeneous UAV Management. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Paper presented at the 2015.

Chidester, T.R., Helmreich, R.L., Gregorich, S.E., Geis, C.E., 1991. Pilot personality and crew coordination: implications for training and selection. Int. J. Aviat. Psychol. 1 (1), 25–44.

Chien, S.-Y., Lewis, M., Sycara, K., Kumru, A., Liu, J.-S., 2019. Influence of culture, transparency, trust, and degree of automation on automation use. IEEE Trans. Human-Machine Syst. 50 (3), 205–214.

Chiou, E.K., Lee, J.D., 2023. Trusting automation: designing for responsivity and resilience. Hum. Factors 65 (1), 137–165.

Chorley, R., 1976. Seventy years of flight instruments and displays. Aeronaut. J. 80 (788), 323–342.

Christ, P.F., Lachner, F., Hösl, A., Menze, B., Diepold, K., Butz, A., 2016. Human-drone-interaction: a case study to investigate the relation between autonomy and user experience. Paper Presented at the Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part II 14.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20 (1), 37–46.

Davies, L., Bolam, R.C., Vagapov, Y., Anuchin, A., 2018. Review of unmanned aircraft system technologies to enable beyond visual line of sight (BVLOS) operations. Paper Presented at the 2018 X International Conference on Electrical Power Drive Systems (ICEPDS).

De Visser, E.J., Peeters, M.M., Jung, M.F., Kohn, S., Shaw, T.H., Pak, R., Neerincx, M.A., 2020. Towards a theory of longitudinal trust calibration in human–robot teams. Int. J. Social Robotics 12 (2), 459–478.

Di Nocera, F., Camilli, M., Terenzi, M., 2007. A random glance at the flight deck: pilots' scanning strategies and the real-time assessment of mental workload. J. Cogn. Eng. Decis. Mak. 1 (3), 271–285.

EASA, 2015. Riga declaration on remotely piloted aircraft (drones). Framing the Future of Aviation. Retrieved from. https://eu2015.lv/images/news/2016_03_06_RPAS_Riga_Declaration.pdf.

Edworthy, J., Hellier, E., Hards, R., 1995. The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals. Ergonomics 38 (11), 2341–2361.

Endsley, M.R., Kiris, E.O., 1995. The out-of-the-loop performance problem and level of control in automation. Hum. Factors 37 (2), 381–394.

Engle-Friedman, M., Mathew, G.M., Martinova, A., Armstrong, F., Konstantinov, V., 2018. The role of sleep deprivation and fatigue in the perception of task difficulty and use of heuristics. Sleep Sci. 11 (2), 74–84.

Glicksohn, J., Naor-Ziv, R., 2016. Personality profiling of pilots: traits and cognitive style. Int. J. Personality Psychol. 2, 7–14.

Gore, J., Shortland, N., Power, N., Brown, O., 2023. In: Naturalistic Decision Making (NDM): Epistemic Expertise in Action, vol. 14. Frontiers Media SA, 1303098.

Grindley, B., Parnell, K.J., Cherett, T., Scanlan, J., Plant, K.L., 2024a. Understanding the human factors challenge of handover between levels of automation for uncrewed air systems: a systematic literature review. Transport. Plann. Technol. 1–26.

Grindley, B., Phillips, K., Parnell, K.J., Cherrett, T., Scanlan, J., Plant, K.L., 2024b. Over a decade of UAV incidents: a human factors analysis of causal factors. Appl. Ergon. 121, 104355.

Gröhn, M., Lokki, T., Takala, T., 2004. An orientation experiment using auditory artificial horizon. Paper Presented at the ICAD.

Gupta, A., Afrin, T., Scully, E., Yodo, N., 2021. Advances of UAVs toward future transportation: the State-of-the-Art, challenges, and Opportunities. Future Transport. 1 (2), 326–350.

Harris, D., 2004. Head-down flight deck display design. Human Factors Civil Flight Deck Design 69–102.

Harris, D., 2017. Human Factors for Civil Flight Deck Design. Routledge.

Harrison, Y., Horne, J.A., 2000. The impact of sleep deprivation on decision making: a review. J. Exp. Psychol. Appl. 6 (3), 236.

Hobbs, A., 2010. Unmanned aircraft systems. In: Human Factors in Aviation. Elsevier, pp. 505–531.

Hobbs, A., Lyall, B., 2016. Human factors guidelines for unmanned aircraft systems. Ergon. Des 24 (3), 23–28.

Hobbs, A., Shively, R.J., 2013. Human factors guidelines for UAS in the national airspace system. In: Proceedings of the Association for Unmanned Vehicle Systems International (AUVSI), Washington, DC, pp. 12–15.

Hockey, G.R.J., 1997. Compensatory control in the regulation of human performance under stress and high workload: a cognitive-energetical framework. Biol. Psychol. 45 (1–3), 73–93.

Hoff, K.A., Bashir, M., 2015. Trust in automation: integrating empirical evidence on factors that influence trust. Hum. Factors 57 (3), 407–434.

Hoffman, R.R., Crandall, B., Shadbolt, N., 1998. Use of the critical decision method to elicit expert knowledge: a case study in the methodology of cognitive task analysis. Hum. Factors 40 (2), 254–276.

Hruschka, D.J., Schwartz, D., St John, D.C., Picone-Decaro, E., Jenkins, R.A., Carey, J.W., 2004. Reliability in coding open-ended data: lessons learned from HIV behavioral research. Field Methods 16 (3), 307–331.

Jenkins, D.P., 2012. Using cognitive work analysis to describe the role of UAVs in military operations. Theor. Issues Ergon. Sci. 13 (3), 335–357.

Ji, X., Huang, H., Li, Z., Guo, Z., Rau, P.-L.P., 2022. Comparing interventions to reduce boredom in a low mental workload environment. Int. J. Occup. Saf. Ergon. 28 (3), 1973–1979.

Jia, H., Lin, C.J., Wang, E.M.-y., 2022. Effects of mental fatigue on risk preference and feedback processing in risk decision-making. Sci. Rep. 12 (1), 10695.

Kaliardos, B., Lyall, B., 2015. Human factors of unmanned aircraft system integration in the national airspace system. In: Valavanis, K.P., Vachtsevanos, G.J. (Eds.), Handbook of Unmanned Aerial Vehicles. Springer Netherlands, Dordrecht, pp. 2135–2158.

Kazi, T.A., Stanton, N.A., Walker, G.H., Young, M.S., 2007. Designer driving: drivers' conceptual models and level of trust in adaptive cruise control. Int. J. Veh. Des. 45 (3), 339–360.

Klein, G., Phillips, J.K., Rall, E.L., Peluso, D.A., 2007. A data–frame theory of sensemaking. In: Expertise Out of Context. Psychology Press, pp. 118–160.

Klemen, J., Büchel, C., Rose, M., 2009. Perceptual load interacts with stimulus processing across sensory modalities. Eur. J. Neurosci. 29 (12), 2426–2434.

Kraus, J., Scholz, D., Baumann, M., 2021. What's driving me? Exploration and validation of a hierarchical personality model for trust in automated driving. Hum. Factors 63 (6), 1076–1105.

Kraus, J., Scholz, D., Stiegemeier, D., Baumann, M., 2020. The more you know: trust dynamics and calibration in highly automated driving and the effects of take-overs, system malfunction, and system transparency. Hum. Factors 62 (5), 718–736.

Lacson, F.C., Risser, M.R., Gwynne, J.W., Kosnik, W.D., 2017. The human systems integration framework: enhanced HSI support for system acquisition. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61. SAGE Publications, Sage CA: Los Angeles, CA, pp. 1720–1724.

Lamont, M., Swidler, A., 2014. Methodological pluralism and the possibilities and limits of interviewing. Qual. Sociol. 37, 153–171.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 159–174.

Landry, S.J., 2021. Human factors and ergonomics in aviation. Handbook Human Factors Ergon. 1460–1479.

Lavie, N., 1995. Perceptual load as a necessary condition for selective attention. J. Exp. Psychol. Hum. Percept. Perform. 21 (3), 451.

Lavie, N., 2005. Distracted and confused?: selective attention under load. Trends Cognit. Sci. 9 (2), 75–82.

Lavie, N., Beck, D.M., Konstantinou, N., 2014. Blinded by the load: attention, awareness and the role of perceptual load. Phil. Trans. Biol. Sci. 369 (1641), 20130205.

Lee, J.D., Hoffman, J.D., Hayes, E., 2004. Collision warning design to mitigate driver distraction. In: Proceedings of the SIGCHI Conference on Human factors in Computing Systems, pp. 65–72.

Lee, J.D., See, K.A., 2004. Trust in automation: designing for appropriate reliance. Hum. Factors 46 (1), 50–80.

Lee, J.D., Seppelt, B.D., 2012. Human factors and ergonomics in automation design. Handbook Human Factors Ergon. 1615–1642.

Li, W.C., Zakarija, M., Yu, C.S., McCarty, P., 2020. Interface design on cabin pressurization system affecting pilot's situation awareness: the comparison between digital displays and pointed displays. Hum. Factors Ergon. ManufactService Ind. 30 (2), 103–113.

Ljungblad, S., Man, Y., Baytaş, M.A., Gamboa, M., Obaid, M., Fjeld, M., 2021. What matters in professional drone pilots' practice? An interview study to understand the complexity of their work and inform human-drone interaction research. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Paper presented at the.

Lovesey, E., 1977. The instrument explosion—a study of aircraft cockpit instruments. Appl. Ergon. 8 (1), 23–30.

Lyons, T.J., Gillingham, K.K., Teas, D.C., Ercoline, W.R., Oakley, C., 1990. The effects of acoustic orientation cues on instrument flight performance in a flight simulator. Aviat Space Environ. Med. 61 (8), 699–706.

Marsh, S., Dibben, M.R., 2003. The role of trust in information science and technology. Annu. Rev. Inf. Sci. Technol. 37, 465–498.

Miller, L., Kraus, J., Babel, F., Baumann, M., 2021. More than a feeling—interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety. Front. Psychol. 12, 592711.

Milton, N., 2010. The Lessons Learned Handbook: Practical Approaches to Learning from Experience. Elsevier.

Ministry of Defence, M., 2014. Human factors integration in defence systems (JSP 912). Retrieved from. https://www.gov.uk/government/publications/human-factors-integration-in-defence-systems-jsp-912.

Mohammed, A., Ibrahim, B.G., Momoh, M.O., Ter, K.P., Adetifa, A.O., Oluwole, D.A., 2022. Challenges of ground control system in ensuring safe flights for unmanned aerial vehicles. Mekatronika: J. Intell. Manuf. Mechatronics 4 (1), 8–19.

Mohsan, S.A.H., Khan, M.A., Noor, F., Ullah, I., Alsharif, M.H., 2022. Towards the unmanned aerial vehicles (UAVs): a comprehensive review. Drones 6 (6), 147.

Mouloua, M., Ferraro, J.C., Kaplan, A.D., Mangos, P., Hancock, P.A., 2019. Human factors issues regarding automation trust in UAS operation, selection, and training. In: Human Performance in Automated and Autonomous Systems. CRC Press, pp. 169–190.

Murphy, G., Groeger, J.A., Greene, C.M., 2016. Twenty years of load theory—where are we now, and where should we go next? Psychonomic Bull. Rev. 23, 1316–1340.

Murphy, R.R., 2014. Disaster Robotics. MIT press.

Murray, C.C., Chu, A.G., 2015. The flying sidekick traveling salesman problem: optimization of drone-assisted parcel delivery. Transport. Res. C Emerg. Technol. 54, 86–109.

Niesser, U., 1976. Cognition and Reality: Principles and Implications of Cognitive Psychology. T W Freeman.

O'Connor, C., Joffe, H., 2020. Intercoder reliability in qualitative research: debates and practical guidelines. Int. J. Qual. Methods 19, 1609406919899220.

Oncu, M., Yildiz, S., 2014. An Analysis of Human Causal Factors in Unmanned Aerial Vehicle (Uav) Accidents. Naval Postgraduate School, Monterey, California.

Orasanu, J.M., 2017. Decision-making in the cockpit. In: Decision Making in Aviation. Routledge, pp. 103–138.

Parasuraman, R., de Visser, E., Wiese, E., Madhavan, P., 2014. Human trust in other humans, automation, robots, and cognitive agents: neural correlates and design implications. Paper Presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Parnell, K.J., Banks, V.A., Allison, C.K., Plant, K.L., Beecroft, P., Stanton, N.A., 2021. Designing flight deck applications: combining insight from end-users and ergonomists. Cognit. Technol. Work 23, 353–365.

Parnell, K.J., Banks, V.A., Wynne, R.A., Stanton, N.A., Plant, K.L., 2023a. Human Factors on the Flight Deck: A Practical Guide for Design, Modelling and Evaluation. CRC Press.

Parnell, K.J., Fischer, J.E., Clark, J.R., Bodenmann, A., Galvez Trigo, M.J., Brito, M.P., Ramchurn, S.D., 2023b. Trustworthy UAV relationships: applying the schema action world taxonomy to UAVs and UAV swarm operations. Int. J. Hum. Comput. Interact. 39 (20), 4042–4058.

Parnell, K.J., Stanton, N.A., Plant, K.L., 2016. Exploring the mechanisms of distraction from in-vehicle technology: the development of the PARRC model. Saf. Sci. 87, 25–37.

Parnell, K.J., Wynne, R.A., Plant, K.L., Banks, V.A., Griffin, T.G.C., Stanton, N.A., 2022. Pilot decision-making during a dual engine failure on take-off: insights from three different decision-making models. Hum. Factors Ergon. ManufactService Ind. 32 (3), 268–285.

Plant, K.L., Stanton, N.A., 2012. Why did the pilots shut down the wrong engine? Explaining errors in context using Schema Theory and the Perceptual Cycle Model. Saf. Sci. 50 (2), 300–315.

Plant, K.L., Stanton, N.A., 2013. The explanatory power of Schema Theory: theoretical foundations and future applications in Ergonomics. Ergonomics 56 (1), 1–15.

Plant, K.L., Stanton, N.A., 2015. The process of processing: exploring the validity of Neisser's perceptual cycle model with accounts from critical decision-making in the cockpit. Ergonomics 58 (6), 909–923.

Plant, K.L., Stanton, N.A., 2016. The development of the Schema World Action Research Method (SWARM) for the elicitation of perceptual cycle data. Theor. Issues Ergon. Sci. 17 (4), 376–401.

Raby, M., Wickens, C.D., 1994. Strategic workload management and decision biases in aviation. Int. J. Aviat. Psychol. 4 (3), 211–240.

Rafferty, L.A., Stanton, N.A., Walker, G.H., 2010. The famous five factors in teamwork: a case study of fratricide. Ergonomics 53 (10), 1187–1204.

Rainieri, G., Fraboni, F., Russo, G., Tul, M., Pingitore, A., Tessari, A., Pietrantoni, L., 2021. Visual scanning techniques and mental workload of helicopter pilots during simulated flight. Aerospace Med. Human Performance 92 (1), 11–19.

Recarte, M.A., Nunes, L.M., 2003. Mental workload while driving: effects on visual search, discrimination, and decision making. J. Exp. Psychol. Appl. 9 (2), 119.

Ruff, H.A., Calhoun, G.L., Draper, M.H., Fontejon, J.V., Guilfoos, B.J., 2004. Exploring Automation Issues in Supervisory Control of Multiple UAVs. Retrieved from.

Sadler, G., Battiste, H., Ho, N., Hoffmann, L., Johnson, W., Shively, R., Smith, D., 2016. Effects of transparency on pilot trust and agreement in the autonomous constrained flight planner. Paper Presented at the 2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC).

Sah, B., Gupta, R., Bani-Hani, D., 2021. Analysis of barriers to implement drone logistics. Int. J. Logist. Res. Appl. 24 (6), 531–550.

Sarter, N.B., Woods, D.D., Billings, C.E., 1997. Automation surprises. Handbook Human Factors Ergon. 2, 1926–1943.

Scholz, D.D., Kraus, J., Miller, L., 2024. Measuring the propensity to trust in automated technology: examining similarities to dispositional trust in other humans and validation of the PTT-A scale. Int. J. Hum. Comput. Interact. 1–24.

Stanton, N.A., Salmon, P.M., Walker, G.H., Jenkins, D., 2009. Genotype and phenotype schemata and their role in distributed situation awareness in collaborative systems. Theor. Issues Ergon. Sci. 10 (1), 43–68.

Stanton, N.A., Salmon, P.M., Walker, G.H., Jenkins, D.P., 2010. Is situation awareness all in the mind? Theor. Issues Ergon. Sci. 11 (1–2), 29–40.

Stanton, N.A., Walker, G.H., 2011. Exploring the psychological factors involved in the Ladbroke Grove rail accident. Accid. Anal. Prev. 43 (3), 1117–1127.

Steen, R., Håheim-Saers, N., Aukland, G., 2024. Military unmanned aerial vehicle operations through the lens of a high-reliability system: challenges and opportunities. Risk Hazards Crisis Publ. Pol. 15 (3), 347–373.
Stöcker, C., Bennett, R., Nex, F., Gerke, M., Zevenbergen, J., 2017. Review of the current state of UAV regulations. Remote Sens. 9 (5), 459.
Tole, J., Stephens, A., Vivaudou, M., Ephrath, A., Young, L., 1983. Visual Scanning Behavior and Pilot Workload. Retrieved from.
Tvaryanas, A.P., 2004. Visual scan patterns during simulated control of an uninhabited aerial vehicle (UAV). Aviat Space Environ. Med. 75 (6), 531–538.
Tvaryanas, A.P., Thompson, W.T., Constable, S.H., 2006. Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years. Aviat Space Environ. Med. 77 (7), 724–732.
Verberne, F.M., Ham, J., Midden, C.J., 2015. Trusting a virtual driver that looks, acts, and thinks like you. Hum. Factors 57 (5), 895–909.
Vidulich, M.A., Tsang, P.S., 2012. Mental workload and situation awareness. Handbook Human Factors Ergon. 243–273.
Wang, L., Gao, S., Tan, W., Zhang, J., 2023. Pilots' mental workload variation when taking a risk in a flight scenario: a study based on flight simulator experiments. Int. J. Occup. Saf. Ergon. 29 (1), 366–375.
Wickens, C.D., 2002. Multiple resources and performance prediction. Theor. Issues Ergon. Sci. 3 (2), 159–177.
Wickens, C.D., 2003. Aviation displays. Principles and Practices of Aviation Psychology, pp. 147–199.
Wickens, C.D., 2021. Attention: theory, principles, models and applications. Int. J. Hum. Comput. Interact. 37 (5), 403–417.
Wickens, C.D., Carswell, C.M., 1995. The proximity compatibility principle: its psychological foundation and relevance to display design. Hum. Factors 37 (3), 473–494.
Wiener, E.L., 1989. Human Factors of Advanced Technology (" Glass Cockpit") Transport Aircraft, vol. 177528. NASA Ames Research Center, Mountain View, CA, USA.
Williams, K.W., 2006. Human factors implications of unmanned aircraft accidents: flight-control problems. Human factors Remotely Operated Vehicles 7, 105–116. Emerald Group Publishing Limited.
Woods, D.D., Sarter, N.B., 2000. Learning from automation surprises and going sour accidents. Cognitive Engineering in the Aviation Domain, pp. 327–353.
Zhang, W., Feltner, D., Kaber, D., Shirley, J., 2021. Utility of functional transparency and usability in UAV supervisory control interface design. Int. J. Social Robotics 13 (7), 1761–1776.
Zhang, Y., Liao, Q.V., Bellamy, R.K., 2020. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making. Paper Presented at the Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency.