What we got Here, is a Failure to Coordinate: Implicit and Explicit Coordination in Air Combat

Mansikka, H. P., Virtanen, K. & Harris, D

DOI 10.1177/15553434231179566
ISSN 1555-3434

Publisher: SAGE Publications

What we got Here, is a Failure to Coordinate: Implicit and Explicit Coordination in Air Combat

Heikki Mansikka and **Kai Virtanen**, Aalto University, National Defence University, Helsinki, Finland, **Don Harris**, Coventry University, Coventry, UK

Air combat is the ultimate test for teamwork, as teams of fighter pilot (or flights), must coordinate their actions in a highly complex, hostile, dynamic and time critical environment. Flights can coordinate their actions using communication, that is, explicitly, or by relying on team situation awareness (SA), that is, implicitly. This paper examines how these two forms of coordination are associated with performance when prosecuting or evading an attack in simulated air combat. This was done by investigating the flights' team SA, number of SA-related communication acts and performance in these two types of critical events during air combat. The results exhibit a quadratic dependence between team SA and communication. The rate of change of SA-related communication frequency with respect to change of team SA was negative: communication was needed to build team SA, but once an appropriate level of team SA was established, fewer communications were required. If, however, team SA deteriorated the number of SA communication acts increased. However, during time critical events, the flights did not always have enough time to coordinate their actions verbally. If the flights' team SA in such situations was low, the flights' explicit coordination attempts were not sufficient to avoid poor performance.

Keywords: air combat, explicit coordination, implicit coordination, team performance, team situation awareness

Address correspondence to Heikki Mansikka, Department of Mathematics and Systems Analysis, Aalto University, National Defence University, P.O. Box 18000, Helsinki 00076, Finland.
Email: heikki.mansikka@aalto.fi
Journal of Cognitive Engineering and Decision Making
Vol. 0, No. 0, ■■ ■, pp. 1-15
DOI:10.1177/15553434231179566
Article reuse guidelines: sagepub.com/journals-permissions

Introduction

Team and its Coordination Mechanisms

A team is a group of two or more individuals (Annett & Stanton, 2000), who work with some level of interdependence (Salas et al., 1992) towards a common goal (Mathieu et al., 2008) within sequential and simultaneous cycles of goal directed activity. To complete its assignment, a team interacts with its tasks, machines and systems (Bowers et al., 1997). For a team to be successful, these types of task-related interactions, that is, taskwork, must be supplemented with interactions between team members and between the team and its environment, that is, teamwork (Fisher, 2014; LePine et al., 2008).

While there is a myriad of teamwork models available (see, e.g., Roberts et al., 2022; Rousseau et al., 2006; Salas et al., 2005 for reviews), there are similarities amongst them. Taskwork and teamwork are generally contextualized as an input-process-output (I-P-O) framework (McGrath, 1964, 1984) or some variation or extension of it (Guzzo & Shea, 1992; Ilgen et al., 2005; Marks et al., 2001). According to the I-P-O-framework, conditions such as the characteristics of team members, available resources and contextual factors serve as inputs for teamwork processes or as mediators that convert the inputs into collective outcomes, that is, team performance outputs (Cannon-Bowers et al., 1995; Hackman, 2012; Mathieu et al., 2000, 2008). The mediators represent a collection of processes and emergent states, which may not directly affect the performance outputs as such, but which can serve as proximal outputs and inputs for other mediators (Ilgen et al., 2005). Teamwork is often viewed as having a temporal aspect such that teams utilize different teamwork processes during distinct performance episodes

(Weingart, 1997; Zaheer et al., 1999), that is, within periods of time during which performance accrues and feedback is available (Button et al., 1996).

Marks et al. (2001) make the temporal aspect explicit by splitting the teams' performance episodes into transition and action phases. During the transition phase, a team either plans its activity for a future action phase or episode, or evaluates its performance in the previous action phase or episode. During the action phase, a team engages directly with taskwork. While authors differ in the precise nature of the interactions that teamwork consists of, some generalizations can be made. Broadly speaking, teams engage in various interpersonal interactions, which besides supporting performance in the long run, affect other aspects of team efficiency, such as team cohesion (Fleishman & Zaccaro, 1992), team members' frustration (Cannon-Bowers et al., 1995), motivation (Chen et al., 2002) and bonding (Ilgen et al., 2005). In addition, during action phases teams engage in task-related interactions, which directly assist the team in achieving the desired level of performance. Such interactions include monitoring of systems' and team members' performance (Dickinson & Mclntyre, 1997; Salas et al., 2005) as well as monitoring task progression (Jentsch et al., 1999; Marks et al., 2001), and backup behaviors (Porter et al., 2003) in case a system or a team member is not performing as expected.

The task-related interactions require coordination. Coordination is the process of orchestrating the sequence and timing of team members' inter-dependent actions (Marks et al., 2001). The higher the level of interdependency between the team members and the more complex and time compressed the team's task is, the more critical it is for the team to be able to coordinate its members' individual efforts effectively during the action phase (Salas et al., 2005; Shaw, 1976; Zalesny et al., 1995). The mechanisms for team coordination can be broadly divided to explicit and implicit ones. Explicit coordination relies on the active use of task programming mechanisms and communication, whereas implicit coordination depends upon the team to coordinate its members' actions without consciously trying (Espinosa et al., 2004). Implicit coordination is based on team members' shared knowledge about the team, its task and its environment, thereby enabling team members to anticipate each other's actions and needs without need for overt communication (Entin & Serfaty, 1999; Rico et al., 2008; Stout et al., 2017).

Implicit coordination requires that team members have common knowledge about the task situation and that during the action phase they successfully perform situation assessment (Salmon et al., 2008) by updating their knowledge with observations. The proximal output of the team members' individual situation assessments is an emergent state known as situation awareness (SA).

Situation Awareness was initially developed as a construct relating to the individual. It is often viewed as a hierarchical construct with three levels: the team members' perception of the relevant elements within their environment (SA level 1), their comprehension regarding the meaning of those elements (SA level 2) and their projection of the elements' status in the near future (SA level 3) (Endsley, 1995). She also suggested that team SA can be defined as "*the degree to which every team member possesses the SA required for his or her responsibilities*" (Endsley, 1995; p. 39). These SA components are inter-dependent in meeting the overall goal of the team's task. Nevertheless, Endsley's definition still adopts a largely individual SA perspective when functioning as a part of a team. However, in addition to team members' individual SA, the team possesses shared common cognitive ground often referred to as shared mental models (Cannon-Bowers et al., 1993; Salas et al., 2005), team mental models (Langan-Fox et al., 2004) or team situation awareness (TSA) (Sulistyawati et al., 2009). The more accurate and similar the team members' SA is, the better TSA is and more likely the team is to succeed in implicit coordination.

Both implicit and explicit coordination have challenges. As SA also directs a person's attention allocation, a confirmatory bias can make it difficult for an individual to know whether his/her SA is accurate and if SA should be updated or not (Fracker, 1988). With inaccurate SA it may be difficult to find and identify relevant information from the environment. This complicates the process of regaining SA once it is lost. In case of TSA, the situation is even more challenging. TSA is seldom perfect, and when the accuracy of team members' SA is low, the SA of individual team members can be similarly or dissimilarly false (Mohammed et al., 2010; Stout et al., 2017). Such uncommon cognitive ground can make a team's implicit coordination efforts difficult or impossible. Explicit coordination, on the other hand, entails building and maintaining a common understanding of the situation (Salas et al., 1997; Serfaty et al., 1998) and enables coordination of team members' activities (DeChurch & Mesmer-Magnus, 2010) even when TSA is low. For explicit coordination to be effective, it must be timely (Park & Kim, 2018), rapid and frequent (Caldwell, 2008) as well as task relevant (Kim et al., 2010). However, communication is a vulnerable means of coordination. It is prone to misinterpretations and misunderstandings, and when such problems occur, the efficiency of coordination efforts may be hampered (Svensson & Andersson, 2006). Furthermore, teams do not always have the time or opportunity to communicate freely. If a team has to rely on radio communication during the most intensive phases of its task, time pressure may result in overlapping transmissions (Lahtinen et al., 2010) and in omissions of necessary transmissions (Kleinman & Serfaty, 1989; Mathieu et al., 2000; Orasanu & Salas, 1993).

While TSA has been found to be a key predictor of team performance in complex and dynamic environments (Salmon et al., 2006), the returns in performance diminish as TSA improves (Mansikka et al., 2021a). In addition, the confidence that individuals and teams have in their SA also has an impact on performance (Hamilton et al., 2017). In dynamic tasks, it may be hard to detect all the changes in the environment (Durlach, 2004), especially as the change and threat detection performance deteriorates with an increase in workload (Matthews et al., 2015). Challenges in change detection combined with the tendency of humans to overestimate their abilities to glean information from their environment can make situation assessment difficult during times of stress and high levels of time pressure (John & Smallman, 2008). For a team, critical situations place high demand on situation assessment (Kozlowski et al., 2009).

van den Oever & Schraagen (2021) define critical situations as events which have high levels of complexity, hazard and time pressure. For instance, in air combat, examples of critical events include situations where a friendly aircraft has launched a weapon against an enemy or when an enemy has launched a weapon against a friendly aircraft. During critical events, military teams may have to adapt their coordination strategies (van den Oever & Schraagen, 2021) to maintain their combat effectiveness (Roberts & Dotterway, 1995). Sulistyawati et al. (2009) noted that explicit coordination is necessary for situation assessment. This is in line with studies which have suggested that explicit coordination frequency during critical events is positively correlated with both TSA (Costello et al., 2006) and team performance (Gontar et al., 2017; Sexton & Helmreich, 2000). In contrast, Entin & Serfaty (1999) argue that during critical events teams tend to switch from explicit to implicit coordination to reduce the communication and coordination overhead and to maintain their performance. Mansikka et al. (2022) made a similar finding when they examined TSA during simulated air combat. They found that low TSA had a significant negative impact on performance during critical events and that TSA was higher when the friendly aircraft had launched a weapon against the enemy aircraft, compared to situations when the friendly aircraft themselves were attacked.

While many studies have shown a strong link between communication and TSA (see, e.g., Garbis & Artman, 2004; Hazlehurst et al., 2007; Heath & Luff, 1991; Kiekel et al., 2001), their relationship is reciprocal. On one hand, communication is needed to build sufficient TSA, which, when established, enables coordination without communication (Endsley, 2015; Parush et al., 2011). On the other hand, when TSA in such a situation fails, teams must communicate to re-establish it (Thornton, 1992). Taken together, we believe, unlike Orasanu (1995) and Salas et al. (1995), that the increase in the number of communication acts is a symptom of decreased TSA, especially in critical events with time pressure.

Flight as a Team and Coordination Mechanisms in Air Combat

Fighter aircraft usually operate as a team of four, that is, a flight. A flight consists of a flight leader, a two-ship leader and two wingmen. The flight leader is usually referred to as #1 and his/her wingman as #2. The two-ship leader and his/her wingman are commonly referred to as #3 and #4, respectively. The flight members have clearly established roles and responsibilities. In very generic terms, the wingmen are responsible for searching targets from the volume of airspace assigned to them and to engage targets as directed and as per the flight's tactical contract. The role of the two-ship lead is to direct his/her wingman and to execute tasks either according to the established tactical contract or as directed by the flight leader. Finally, the flight leader has the overall responsibility for the flight's tactical decisions, which ultimately dictate the flight's lethality and survivability. In addition, all flight members are responsible for defending themselves and for providing mutual support for other flight members. For an unclassified discussion of flight members' roles, responsibilities and types of tactical contracts, please see a Korean Air Force Basic Employment Manual, Section 4.8 (Korean Air Force, 2005).

A flight has clear performance episodes with identifiable phases. The flight's mission brief and debrief can be seen as transition phases, whereas the actual mission represents the flight's action phase. During its action phase, a friendly flight's task (referred to as Blue) is essentially to intercept the enemy aircraft (referred to as Red). Within a single performance episode, Blue has identifiable sub-goals, which are typically described as two parallel processes, known as kill-chain and live-chain (Joint Chief of Staff, 2013). The kill-chain describes the progression of a flight's taskwork towards the interception of Red, whereas the live-chain describes how well Blue can deny Red from progressing its own kill-chain. "Red Engaged" is an example of the kill-chain phase, representing that Blue has launched a weapon against Red. In comparison, "Blue Engaged" represents a phase in the live-chain where Red has managed to launch a weapon against Blue. In air combat, the flight's goal is to complete the kill-chain while maintaining its live-chain intact.

To advance its taskwork, Blue orchestrates its available resources using tactics, techniques and procedures (TTPs) (Mansikka et al., 2021b), which are a set of rules and rule values of how the other task-related interactions should be performed (Mansikka et al., 2021c). For a flight, air combat manifests itself as a requirement for constant evaluation of the environment, a need to decide cyclically which TTP to select, and how to execute or adjust it in case contingencies are met. There is also a requirement for constant evaluation concerning how the flight's performance output feeds back to the observed environment. Similar to the perceptual cycle (Neisser, 1976; Plant & Stanton, 2015), the selection, execution and output evaluation of TTPs form a fast-paced I-P-O cycle nested within the action phase (Mansikka et al., 2021b). Within this inner I-P-O cycle, the inputs consist of factors such as the pilots' knowledge about the tactical environment. During the process phase, the flight builds and maintains its TSA to support effective decisions when selecting an appropriate TTP. The sources and types of information the flight utilizes to gain and maintain TSA vary from a dynamic tactical information available via on-board and off-board sensors to more stable information such as coordination contracts decided on the mission brief, intra- and inter-flight contracts and standard-operating procedures. In very general terms, even the flight members' knowledge of each other's attitudes and personality traits contribute to TSA. The flight uses its TSA to match environmental cues, not necessarily with the best, but with a satisfactory decision alternative, that is, TTP. From the perspective of a flight's effectiveness, the optimal situation would be if the flight had perfect TSA. Should this be the case, the flight members would understand the tactical environment in a similar fashion, identify the same TTP as the most feasible and would have a similar view on how to execute it. In other words, the flight could rely solely on implicit coordination (Fisher, 2014). Real-life situations, however, are often sub-optimal: the flight's TSA is less than perfect and as a result, flight members have different views about the most feasible TTP and the way it should be executed. In such cases, the flight must use whatever means to regain and maintain its TSA to enable effective future decision making and actively coordinate the execution of the already decided TTP. When the flight cannot rely on coordination based on its TSA, it has to use explicit coordination, that is, radio communication, to orchestrate the flight members' actions (Fisher, 2014). In air combat, radio frequencies are often busy and radio transmission may overlap, resulting in missed and misunderstood coordination messages. In addition, compared to implicit coordination, building of TSA and coordinating the team members' activities using radio can be time consuming. As a result, even when explicit coordination is possible, the flight may not have enough time to recover sufficiently TSA and reach a desired level of coordination if its implicit coordination has already failed.

In conclusion, we submit that coordination, whether it is explicit or implicit, plays a central role in air combat and can have a significant impact on flights' performance output. This paper concentrates on investigating how communication and the performance output of flights are associated in a simulated air combat. We do this by examining the flights' performance output in two types of kill- and live-chain events, that is, in Blue Engaged events and Red Engaged events. Similar to Sulistyawati et al. (2009), the survival of Blue and the loss of Red are considered to be a success and loss of Blue and survival of Red are considered a failure.

In addition, we consider the flights' SA-related speech acts during and just prior to those critical events. The assumption is that compared to explicit coordination, implicit coordination will result in superior performance when the flights are engaging Red forces and when the flights themselves are being engaged. We maintain, as did Entin and Serfaty (1999), that should the level of TSA permit, the flights prefer implicit coordination when dealing with critical events. We hypothesize that an increase in the number of communication acts, that is, explicit coordination, is an indication of flights' situation assessment efforts in a situation where the level of TSA has deteriorated such that it no longer warrants a more effective type of coordination. Finally, Blue and Red Engaged events are such dynamic events, that we hypothesize that a flight's ability to recover their TSA after implicit coordination had failed, especially when Blue is being engaged, will be inferior.

Method

Participants

Sixteen F/A-18 fighter pilots participated in the study. The mean age of pilots was 30 years (SD = 2.29) and their average experience on F/A-18 aircraft was 412 flight hours (SD = 220). All participants were male.

The pilots operated in flights. The flight leader was referred to as #1 and his/her wingman as #2. The two-ship leader was referred to as #3 with his/her wingman as #4. All pilots were qualified to act in a role they were assigned to. The flight leader was the most qualified pilot in the flight, followed by the two-ship leader. The wingmen were equally qualified and represented the least qualified pilots within a flight. All pilots had passed an aeromedical examination during the past 12 months and were fit to fly at the time of the study. A fighter controller (FC) was assigned to support each flight.

Apparatus

The data for the study were collected during a simulator exercise, which was a part of the pilots' normal flight training. Two types of high-fidelity flight training devices were used in the exercise: one type with a touchscreen display (resolution 1280*1024, frame rate 60 Hz) and a virtual reality (VR) headset providing a 360 field of view (resolution 1920*120, latency less than 6 ms), and the other type with a fully functional cockpit with a 216 degree field of view (resolution 2560*1600, frame rate 60 Hz). The use of the VR headset was at pilots' discretion. The simulators were distributed between two fighter squadrons, separated by several hundred miles. The maximum observed latency of the distributed simulator network between different locations was 21 ms. FC had a command-and-control simulator, which allowed him/her to relay the tactical air picture to the flights and to give them advisories and warnings. All simulators were linked to Modern Air Combat Environment (MACE) simulation software (for details, please visit: https://www.bssim.com/mace/) via a distributed interactive simulation connection. This connection is a standard that provides simulation applications with the ability to exchange information via protocol data units. In addition, there was a tactical L16 datalink connection between all simulators. All enemy aircraft were managed in MACE.

Procedure

The Flying Mission. In the exercise, the pilots formed two F/A-18 flights. The manning of flights varied based on the pilots' training rosters and training objectives. The F/A-18 flights, that is, Blue, undertook a number of beyond-visual-range simulated air combat missions against a computer generated enemy force, that is, Red. The missions were conducted in varying weather and lighting conditions. The blue flights flew defensive counter air (DCA) missions against Red, which conducted offensive counter air (OCA) operations against Blue. The task of the DCA aircraft was to maximize the number of killed Red aircraft within their area of responsibility, while minimizing friendly losses. The task of the OCA air-to-air fighters was to actively engage all DCA aircraft along the attack route of the OCA package. As the OCA package was programmed to attack via the Blue area of responsibility, the missions essentially unfolded as tactical intercepts between Red and Blue air-to-air fighters. Before each mission, the flights were provided with standard mission material about their upcoming mission, including intelligence brief, air tasking order and administrative information such as settings for radios and other aircraft systems. The flights reviewed the mission material and the flight leader briefed the mission to his team. After the briefing, the flight entered the simulators and the simulation was started.

Red was programmed to replicate the capabilities of threat aircraft and to react to Blue's actions according to given behavioral rules. As a result, Red behavior varied dynamically within the boundaries of those rules. Once the simulation had started, it was let to evolve uninterrupted until Blue had completed its mission or the training objectives had been reached. Each mission lasted approximately 40 minutes.

Performance Output. After each mission, the flights' attended a standard debrief. During the debrief, the pilots reviewed the mission using the flight members' cockpit recordings and audio, as well as a computer-animated mission reconstruction of the mission. While the pilots' cockpit recording provided a limited, and sometimes false, picture of simulation reality, the animated mission reconstruction allowed the pilots to see the simulation as it actually happened, that is, the ground truth. During debriefs, the flights identified 29 critical events where the flight had engaged Red (i.e., Red Engaged events) and 29 events where the Red had engaged the flight (i.e., Blue Engaged events). When an aircraft engaged it launched weapon against its target. Being engaged did not automatically mean that the launched weapon would hit its target. As a result, the same aircraft could be engaged several times during one mission.

The flight's performance output in the events was then determined. The possible outputs were either "Success" or "Failure." In a Red Engaged event, Failure was considered as a situation where the Red aircraft evaded the Blue flight's weapon launch. The same event was deemed a Success if the Red aircraft was hit. In a Blue Engaged event, the performance outputs were the opposite: a Blue kill was considered Failure and Blue survival as Success. After the performance output had been determined, the debrief was continued until the next Red Engaged or Blue Engaged event occurred. The debrief was paused again and the output evaluation was repeated. Data collection continued in the same fashion until the first flight member was killed. If none of the flight members was killed during the mission, the data collection was terminated once the first Red was killed.

Team Situation Awareness. Team SA can be assessed using various techniques, such as observer rating, freeze probe, and self- and peer appraisal techniques (see, e.g., Salmon et al., 2006). While appraisal techniques are low-cost and easy to administer, pilots may be unable to accurately report what they and their peers are and are not aware of. The utility of the freeze-probe techniques is highly limited by the fact that stopping a mission in virtual simulation is highly disruptive—and is impossible during real flight. Finally, as the observer rating techniques rely on observable behaviors as indicators of (T) SA, there are doubts to what extent an external observer is able to assess a pilot's subjective reality, which is not necessarily manifested in overt behavior. Post-trial elicitation of knowledge (see, e.g., Rosenman et al., 2018) has proven to be a promising technique to assess (T) SA (Cooke et al., 2017), and such a technique can also be used in a natural air combat training environment (see, e.g., Mansikka et al., 2021a; 2022). As long as the post-trial elicitation technique is administered carefully and the playback of prior activity and the ground truth are available, the technique can reveal a person's situated knowledge, that is, SA, as opposed to reflecting his/her a priori knowledge (Cooke et al., 2017). In this paper, the post-trial elicitation technique was used to assess the pilots' knowledge about the SA attributes in two types of critical events.

Once a critical event was identified, the debrief was paused and a SA attribute which most affected the occurrence of the event was identified. Such a SA attribute is the smallest element in the tactical environment of which a pilot can have SA. For example, "Speed of a non-friendly aircraft" is an attribute. For a complete list of air combat attributes, see Mansikka et al. (2021a). The mission reconstruction as well as the cockpit recordings were rewound for 60 seconds after which the flight members reviewed the last 60 seconds of the engagement again, this time assessing the accuracy of their SA regarding the attribute in question. The pilots evaluated their SA accuracy by comparing their recollection of the situation of interest with its ground truth. Once the last 60 seconds preceding the event had been reviewed, each flight member scored their SA about the attribute on a scale 1 (most inaccurate) to 3 (most accurate). The use of a simple three-point scale enhanced the reliability of SA accuracy assessments during the debriefs (Louangrath, 2018). SA accuracy was rated separately for SA levels 1–3. After the SA scores were obtained, the TSA score of the flight for each SA level in an event was determined by calculating the average of individual pilots' scores. The debrief was then continued until the next critical event occurred. The debrief was paused again and the TSA assessment was repeated. The SA data collection and the performance data collection were terminated at the same time.

Situation Awareness-Related Communication Acts. The pilots utilize all available data about their tactical environment to build and maintain their SA. Visually detectable cues outside the cockpit, information from on-board and off-board sensors displayed on the aircraft's displays as well as system warnings and tactical radio transmissions, all facilitate situation assessment. While all perceived information contribute to SA, this study concentrated solely on SA-related radio communication acts.

The pilots used communication contracts similar to those described in US Air Force TTP for counter air operations (US Government, US Air Force 2001). The radio communications of each flight member and their FC were recorded. Sixty second samples of the radio traffic preceding each Red Engaged and Blue Engaged event were extracted from the audio recording. Separate audio samples were taken from each flight member and FC, resulting in total 290 one-minute samples. The individual samples were listened to, and the SA-related communication acts were identified. These were communications related to building and maintaining the flights' TSA. In addition, the transmitter and recipient of each transmission were identified. Within a flight, a widely accepted leader-subordinate hierarchy was used. In this hierarchy, the flight leader is at the top of the hierarchy followed by the two-ship leader. The wingmen are equal and follow the two-ship leader in the hierarchy. Finally, FC is at the bottom of the hierarchy. As a result, it was possible to determine how much information was sent "upwards" from FC to the flight members, from the wingmen to the two-ship leader or the flight

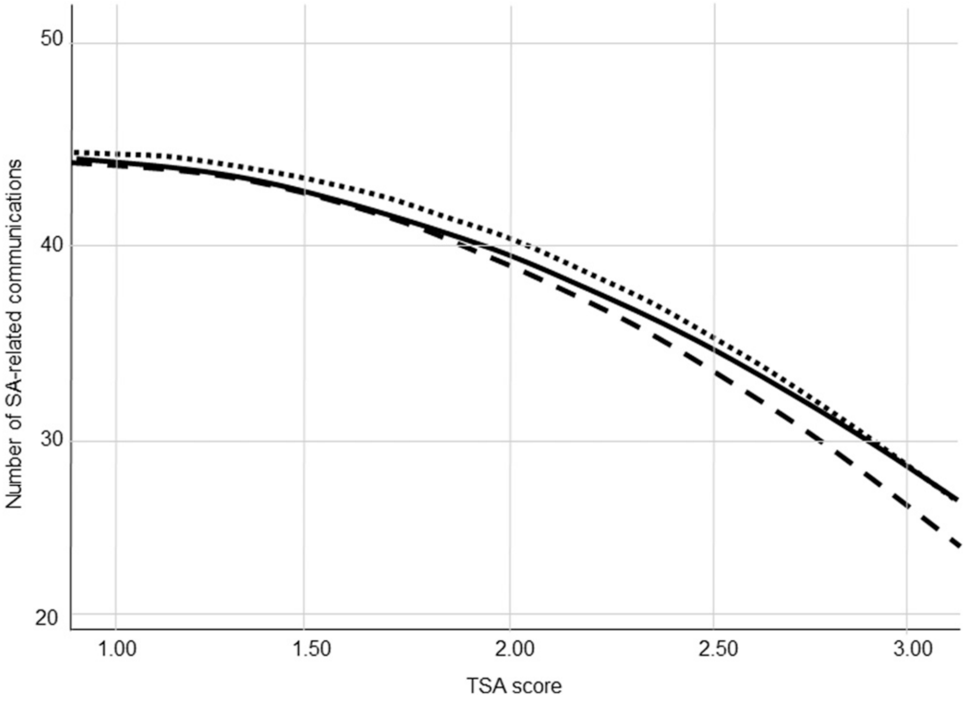

Figure 1. Regression curves for the total number of SA-related communications dependent upon TSA scores on SA levels 1–3. The solid line represents TSA Level 1, the dotted line represents TSA level 2, and the dashed line represents TSA level 3.

leader, or from the two-ship leader to the flight leader. Similarly, the amount of information sent "downwards" was determined by observing the transmissions from the flight leader to other flight members or FC, from the two-ship leader to the wingmen or FC, or from the wingmen to FC.

Results

The unit of analysis was at flight level not for each individual pilot. SA communications were analyzed with respect to TSA. Curvilinear regressions to predict SA-related communications from TSA were calculated for each level of TSA. In these models, TSA accuracy levels 1–3 were used as independent variables and the number of total SA communication acts was used as a dependent variable.

For all levels of TSA, there was a highly significant negative curvilinear relationship between TSA accuracy and total SA communications (TSA level 1: R = 0.394, R^2 = 0.155, R^2_{adj} = 0.124, $F_{2,55}$ = 5.039, $p < .01$; TSA level 2: R = .409 R^2 = .167, R^2_{adj} = .137, $F_{2,55}$ = 5.524, $p < .005$; TSA level 3: R = .438, R^2 = .192, R^2_{adj} = .162, $F_{2,55}$ = 6.516, $p < .005$).

With SA-related communications dependent upon TSA, the best fit models were all quadratic in nature (TSA level 1: SA communications = TSA level 1 *4.591 - TSA level 1^2 * 3.078 + 42.231; TSA level 2: SA communications = TSA level 2 *6.345 - TSA level 2^2 * 3.596 + 41.562; TSA level 3: SA communications = TSA level 3 *5.129 - TSA level 3^2 * 3.492 + 42.286). These regression curves are depicted in Figure 1. All curves are almost identical.

Next, SA communications were analyzed with respect to flights' performance, that is, Failure/Success in critical events, and Red/Blue Engaged events as independent variables. To minimize the probability of a type I error, the SA-related communications (SA UP or SA DOWN) were subject to MANOVA with main effects of Blue or Red Engaged and Success or

Failure performance. This also removed the effects of any inter-correlation between the dependent variables. These data are presented in Table 1.

Both main effects were significant. There was an overall difference in the amount of communications for Blue/Red Engaged (Wilks' Lambda = .797; $F_{(2,53)} = 6.751$; $p < .005$; partial eta^2 = .203) and for Success/Failure performance (Wilks' Lambda = .862; $F_{(2,53)} = 2.298$; $p < .05$; partial eta^2 = .138). The interaction term was also significant (Wilks' Lambda = .859; $F_{(2, 53)} = 4.340$; $p < .05$; partial eta^2 = .141). To aid the interpretation of the multivariate results, the significant main effects were further analyzed using univariate factorial analyses of variance (ANOVA).

With Blue/Red Engaged as the main effect and SA DOWN as the dependent variable, there was a significant difference (SA DOWN: $F_{(1,54)} = 13.250$, $p < .001$, partial eta^2 = .197). Similarly with SA UP as the dependent variable, there was also a significant difference (SA UP: $F_{(1,54)} = 5.797$, $p < .05$, partial eta^2 = .097). There were significantly fewer SA-related communications both DOWN and UP when friendly forces were engaging the enemy (Red Engaged) than when the friendly forces were being attacked (Blue Engaged) (see Table 1).

With Success/Failure performance as the main effect and SA DOWN as the dependent variable, there was no significant difference between success and failure groups (SA DOWN: $F_{(1,54)} = 1.742$, $p > .05$, partial eta^2 = .031). However, with SA UP as the dependent variables, there was a significant difference (SA UP: $F_{(1,54)} = 7.099$, $p < .01$, partial eta^2 = .116). There were significantly fewer SA-related UP communications in successful engagements (see Table 1).

De-composing the contribution of SA-related variables to the significant interaction term showed that there was no significant effect with SA DOWN as the dependent variable (SA DOWN: $F_{(1,54)} = .073$, $p > .05$, partial eta^2 = .001). There was a significant interaction involving SA UP ($F_{(1,54)} = 4.213$, $p < .05$, partial eta^2 = .072). The interaction term is shown in Figure 2. In Red Engaged events, the number of SA UP communications showed large difference between Success and Failure outcomes whereas

Table 1: Means (M), Standard Deviations (SD) and Sample Sizes (N) for SA-Related Communications (COM) UP and DOWN Within the Flight Tabulated by Blue Engaged or Red Engaged and the Performance Output Failure or Success.

			M	SD	N
SA COM DOWN	Blue Engaged	Failure	17.93	7.035	15
		Success	15.93	7.353	14
		Total	16.97	7.134	29
	Red Engaged	Failure	11.50	6.418	14
		Success	8.47	8.096	15
		Total	9.93	7.368	29
	Total	Failure	14.83	7.388	29
		Success	12.07	8.502	29
		Total	13.45	8.016	58
SA COM UP	Blue Engaged	Failure	38.40	13.627	15
		Success	36.36	12.188	14
		Total	37.41	12.763	29
	Red Engaged	Failure	37.21	12.503	14
		Success	21.47	12.403	15
		Total	29.07	14.616	29
	Total	Failure	37.83	12.876	29
		Success	28.66	14.256	29
		Total	33.24	14.236	58

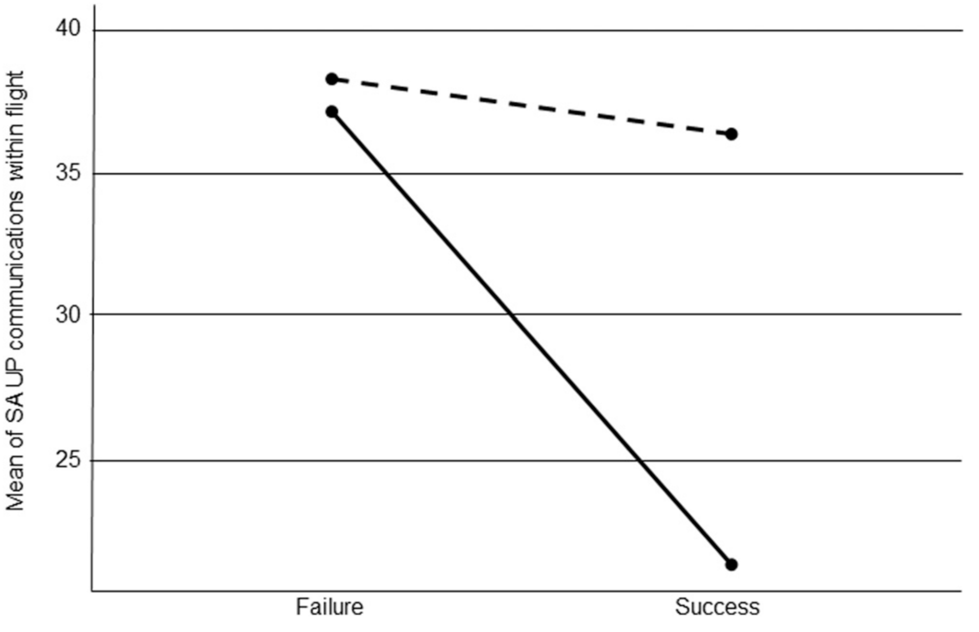

Figure 2. Interaction plot of SA UP communications within a flight with regard to Blue engaged and Red Engaged events resulting in failure or success. Blue Engaged is depicted as a dashed line and Red Engaged is depicted as a solid line.

Blue Engaged events revealed only small difference.

Discussion

This paper investigated how the number of SA-related communications is associated with flights' performance in two types of critical events during simulated air combat: Blue Engaged and Red Engaged events. The number of speech acts was seen as an indirect indication of flights' TSA. Flights' switch to explicit coordination, indicated by the increased number of speech acts, was thought to be an indication of insufficient TSA. In addition, the direction of SA-related communication was studied which provided insights about flights' explicit coordination mechanisms. While explicit coordination has its benefits, the results of this paper leave little doubt about the most effective form of coordination in air combat which is implicit.

TSA accuracy increased in a nonlinear fashion as the number of SA-related communications decreased during the engagement. The regression models presented in Figure 1 revealed that when the level of TSA decreased, the total number of SA-related communication acts increased, suggesting the pilots' verbal attempts to re-establish an acceptable level of TSA. The form of dependence between TSA and communication was quadratic, and the rate of the change of communication frequency with respect to the change of TSA was negative. That is, the increase in the number of communication acts was greater when TSA started to deteriorate from its maximum, compared to a situation where TSA was already low as it started to decline. The results indicated that once TSA had dropped to a low level, additional communication as a means to recover from the situation was no longer effective, resulting in even worse TSA (cf. Orasanu, 1995; Salas et al., 1995; Thornton, 1992). As shown in Figure 1, the described relationship was similar for TSA levels 1, 2 and 3.

The results also revealed a significant relationship between the number of SA-related communication acts and flights' performance in different critical events. Considering the possible performance outputs of these events, Blue Engaged should be considered to be a more critical than Red Engaged. In the Red Engaged events, the worst outcome was Red not being hit, whereas in the Blue Engaged events the encounters could end in a loss of a Blue aircraft. The results clearly showed that the total number of speech acts was higher in the Blue Engaged events compared to the Red Engaged ones (see Table 1). This is logical, as with high TSA the flights should not have ended up being engaged by Red in the first place. The comparison of Success and Failure performance outputs revealed a similar pattern; there were fewer explicit communication in events which resulted in Success compared to those which ended in Failure (see Table 1). From the flights' perspective Blue Engaged was the most critical event and Failure in that event was a disaster. An increase in the number of speech acts reflected the flights' last ditch, but failed, effort to avoid that.

As discussed above, there were more intense SA-related communications in Blue Engaged events compared to Red Engaged events. In fact, some Red Engaged events were completed with no communication at all. The Blue Engaged events, however, can be seen as a result of failed implicit coordination as evidenced by the significant increase in SA-related communications during those events (see Table 1). In contrast to the findings of Sulistyawati et al. (2009), Gontar et al. (2017) and Sexton and Helmreich (2000), the high number of SA-related communication acts in events resulting in Failure, especially in the Blue Engaged events, is a clear indication of the weakness of explicit coordination.

Overall, the same general SA-related communication behavior can be seen in the direction of speech acts. There were significantly fewer SA DOWN and SA UP communication acts in Red Engaged events compared to events when Blue was being engaged (see Table 1). The Blue Engaged events are not under the control of the flight and result from some unexpected changes in the tactical situation where the Blue flight needs to react to the situation. The increased SA UP/DOWN communication was probably motivated by these unexpected changes and served as an attempt to help #1 and #3 in adapting their mental models and decision making to this new situation. This reflects the different roles within a flight, as flight and two-ship leaders, #1 and #3 are predominantly responsible for the tactical decision making within a flight. According to their roles, the other flight members feed the decision makers with SA-related information to support their decision making. Once the tactical decisions are made, #1 and #3 feed the rest of the flight with SA-related information such that each flight member can independently adapt their TTP execution as needed.

In events resulting in Success, there were significantly fewer SA UP communications compared to events resulting in Failure. No significant change was observed in SA DOWN communications between Success and Failure events. In air combat, friendly losses are typically not accepted, and the events evolving towards Failure typically include some unexpected changes in the tactical environment. The increase in the frequency of SA UP communication reflects the flight members' attempt to assist the primary decision makers, that is, #1 and #3, to adapt to these changes. The reevaluation of tactical options in such a tactical situation can be cognitively resources heavy. It is possible that #1 and #3 had to limit their SA-related communications just to keep up with the required pace in tactical decision making.

Salas et al. (2005), Zalesny et al. (1995) and Shaw (1976) all suggested that in complex, time-pressured tasks it was critical to coordinate team members' individual actions, especially when there was high interdependency between them. However, in extremely time-pressured, high workload, highly dynamic situations, attempts at explicit coordination may be counterproductive. Implicit coordination based upon common knowledge about the task situation could be the best option. Providing additional, potentially conflicting SA information may be detrimental to performance rather than enhancing it (Carroll & Sanchez, 2021). Evaluating new SA-related information within an existing mental model of the situation may turn out to be too demanding in such engagements. Suggestion of Salmon et al. (2008) that implicit coordination also requires updating team members' knowledge with observations can be counterproductive in very highly time-pressured situations. Integrating potentially conflicting new information may slow down decision making. Entin and Serfaty (1999) noted that during critical events teams tend to switch from explicit to implicit coordination to reduce the communication and coordination overhead to maintain performance. As discussed above, the flights were faced problems when low TSA mandated an opposite switch, that is, from implicit to explicit coordination. In time compressed and dynamic situations, the flights simply did not have enough time for verbal situation assessment and for the dissemination of SA-related information to all flight members. As a result, when the flights had to switch to explicit coordination, they tended to fail.

In the Introduction section, it was hypothesized that an increase in the number of communication acts is an indication of flights' situation assessment efforts in a situation where the level of TSA has deteriorated such that it no longer warrants a more effective type of coordination. Also, it was hypothesized that flights' ability to recover their TSA after implicit coordination had failed, especially when Blue was being engaged, will be inferior. The results of this paper confirm these hypotheses.

Regarding the limitations of the study presented in this paper, the fighter pilot community may find some of the assumptions and findings of this paper intuitive. At the same time, there seems to have been a gap between what is commonly known by the pilots and what is reported in the scientific literature. As a result, the ability of the unclassified research to contribute to commonly known issues within an air combat domain has been limited. This paper serves to bridge the mentioned gap and is expected to motivate future studies about teamwork in the context of air combat.

As this study focused especially on the coordination mechanisms and performance of a four-ship, the data collection was terminated as soon as the first flight member was killed. In reality, the flight may continue its mission after a friendly loss(es) as a three- or two-ship. The coordination mechanisms of such formations are likely to differ from those of the four-ship and warrant further investigation.

While the focus of this study was limited to SA accuracy at a team level, future studies should evaluate what impact individual SA differences have on the flight's performance and team processes. For example, to what extent a flight is resilient against an inaccurate SA of individual pilots? In addition, future studies are encouraged to examine what is the impact of individual pilots' similarly or dissimilarly inaccurate SA on the flight's performance (see, e.g., Mansikka et al., 2022). The TSA measuring technique described in this study should be helpful in both endeavors. While slightly outside the scope of this paper, it would be theoretically interesting and practically valuable to go beyond the reactions of Blue to critical events, and to investigate the coordination mechanisms and team processes leading to those events.

This this study has several potential practical applications. When air combat simulations are used to evaluate and compare the utility of tactical operating procedures, the competence of teams or the applicability of aircraft systems, it is essential to have robust measures of team performance. This study contributes to the measurement of team performance by revealing how different coordination mechanisms and performance are linked. While the findings of this paper directly support team performance measurement in fighter pilot training, they can also be applied to practically any other domain where the evaluation of team coordination and performance can be supported with the playback of task activities and ground truth.

Regardless of the application domain, the principles of this paper can be used to identify situations where teams' poor performance is associated with a shift from implicit to explicit coordination. Once this observation has been made, it is possible to begin investigating the root causes for this phenomenon. The approach presented in this paper can help in identifying issues in the procedures and equipment used by the teams, deficiencies in teams' training curricula, and even competence shortages of individual team members and complete teams.

Conclusions

In conclusion, this paper can explain why the findings regarding the association between TSA and communication acts have been so conflicting (see, e.g., Endsley, 2015; Garbis & Artman, 2004; Hazlehurst et al., 2007; Heath & Luff, 1991; Kiekel et al., 2001; Orasanu, 1995; Parush et al., 2011; Salas et al., 1995; Thronton, 1992). Communication is needed to build TSA, but once an appropriate level of TSA has been established, less communication is required. If, however, TSA is lost, it is likely that the number of communication acts will again increase. Once TSA collapses, implicit coordination is no longer possible and an alternative coordination mechanism is necessary. At the same time, the increased communication serves as a method to regain TSA. The confusion in the existing literature can be explained by the curvilinear dependence between communication and TSA revealed in this paper where lower TSA resulted in higher communication frequency. In summary, the association is context dependent and has a temporal aspect. As long as these aspects are appreciated, results of future studies will probably be less puzzling.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Heikki Mansikka https://orcid.org/0000-0002-7261-6408

References

Annett, J. A., & Stanton, N. A. (2000). Team work - A problem for ergonomics? *Ergonomics*, *43*(8), 1045–1051. https://doi.org/10.1080/00140130050084860

Bowers, C. A., Braun, C. C., & Morgan, B. B., Jr. (1997). Team workload: Its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance and measurement: Theory, methods, and applications* (pp. 85–108). Lawrence Erlbaum Associates.

Button, S. B., Mathieu, J. E., & Aikin, K. J. (1996). An examination of the relative impact of assigned goals and self-efficacy on personal goals and performance over time. *Journal of Applied Social Psychology*, *26*(12), 1084–1103. https://doi.org/10.1111/j.1559-1816.1996.tb01126.x

Caldwell, B. S. (2008). Knowledge sharing and expertise coordination of event response in organizations. *Applied Ergonomics*, *39*(4), 427–438. https://doi.org/10.1016/J.APERGO.2008.02.010

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making* (pp. 221–246). Lawrence Erlbaum.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). Jossey-Bass.

Carroll, M. B., & Sanchez, P. L. (2021). Decision making with conflicting information: Influencing factors and best practice guidelines. *Theoretical Issues in Ergonomics Science*, *22*(3), 296–316. https://doi.org/10.1080/1463922X.2020.1764660

Chen, G., Webber, S. S., Bliese, P. D., Mathieu, J. E., Payne, S. C., Zaccaro, S. J., Webber, S. S., Mathieu, J. E., & Born, D. H. (2002). Simultaneous examination of the antecedents and consequences of efficacy beliefs at multiple levels of analysis. *Human Performance*. *15*(4), 381–409. https://doi.org/10.1207/S15327043HUP1504_05

Cooke, N. J., Stout, R. J., & Salas, E. (Eds.), (2017). *Situational awareness* (pp. 157–182). Routledge.

Costello, A. M., Strater, L. D., Bolstad, C. A., Cuevas, H. M., & Endsley, M. R. (2006). Communication and situation awareness in ad hoc teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *52*(4), 468–472. https://doi.org/10.1177/154193120805200456

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010). The cognitive underpinnings of effective teamwork: A meta-analysis. *The Journal of Applied Psychology*, *95*(1), 32–53. https://doi.org/10.1037/a0017328

Dickinson, T. L., & Mclntyre, R. M. (1997). A conceptual framework for teamwork measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance and measurement: Theory, methods, and applications* (pp. 19–43). Lawrence Erlbaum Associates.

Durlach, P. J. (2004). Change blindness and its implications for complex monitoring and control systems design and operator training. *Human–Computer Interaction*, *19*(4), 423–451. https://doi.org/10.1207/s15327051hci1904_10

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R. (2015). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making*, *9*(1), 4–32. https://doi.org/10.1177/1555343415572631

Entin, E. E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *41*(2), 312–325. https://doi.org/10.1518/001872099779591196

Espinosa, J. A., Lerch, J., & Kraut, R. (2004). Explicit vs. implicit coordination mechnanisms and task dependencies: One size does not fit all. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance*(pp. 107–129). American Psychological Association. https://doi.org/10.1037/10690-006

Fisher, D. M. (2014). Distinguishing between taskwork and teamwork planning in teams: Relations with coordination and interpersonal processes. *The Journal of Applied Psychology*, *99*(3), 423–436. https://doi.org/10.1037/a0034625

Fleishman, E. A., & Zaccaro, S. J. (1992). Toward a taxonomy of team performance functions. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 31–56). Ablex.

Fracker, M. L. (1988). A theory of situation assessment: Implications for measuring situation awareness. *Proceedings of the Human Factors Society Annual Meeting*, *32*(2), 102–106. https://doi.org/10.1177/154193128803200222

Garbis, C., & Artman, H. (2004). Team situation awareness as communicative practices. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 275–296). Ashgate.

Gontar, P., Fischer, U., & Bengler, K. (2017). Methods to evaluate pilots' cockpit communication: Cross-recurrence analyses vs. speech act–based analyses. *Journal of Cognitive Engineering and Decision Making*, *11*(4), 337–352. https://doi.org/10.1177/1555343417715161

Guzzo, R. A., & Shea, G. P. (1992). Group performance and intergroup relations in organizations. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., pp. 269–313). Consulting Psychologists Press.

Hackman, J. R. (2012). From causes to conditions in group research. *Journal of Organizational Behavior*, *33*(3), 428–444. https://doi.org/10.1002/job.1774

Hamilton, K., Mancuso, V., Mohammed, S., Tesler, R., & McNeese, M. (2017). Skilled and unaware: The interactive effects of team cognition, team metacognition, and task confidence on team performance. *Journal of Cognitive Engineering and Decision Making*, *11*(4), 382–395. https://doi.org/10.1177/1555343417731429

Hazlehurst, B., McMullen, C. K., & Gorman, N. (2007). Distributed cognition in the heart room: how situation awareness arises from coordinated communications during cardiac surgery. *Journal of Biomedical Informatics*, *40*(5), 539–551. https://doi.org/10.1016/j.jbi.2007.02.001

Heath, C., & Luff, P. (1991). Collaborative activity and technological design: Task coordination in London underground control rooms. In L. Bannon, M. Robinson, & K. Schmidt (Eds.), Proceedings of the second European conference on computer-supported cooperative work ECSCW '91 (pp. 65–80). Springer. https://doi.org/10.1007/978-94-011-3506-1_5

Ilgen, D. R., Hollenbeck, J. R., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, *56*(1), 517–543. https://doi.org/10.1146/annurev.psych.56.091103.070250

Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situational awareness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *41*(1), 1–14. https://doi.org/10.1518/001872099779577237

John, M. S., & Smallman, H. S. (2008). Staying up to speed: Four design principles for maintaining and recovering situation awareness. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 118–139. https://doi.org/10.1518/155534308x284408

Joint Chief of Staff (2013). *Joint publication 3-60: Joint targeting*. US Joint Chief of Staff.

Kiekel, P. A., Cooke, N. J., Foltz, P. W., & Shope, S. M. (2001). Automating measurement of team cognition through analysis of communication data. In M. J. Smith, G. Salvendy, D. Harris, & R. J. Koubek (Eds.), *Usability evaluation and interface design* (pp. 1382–1386). Lawrence Erlbaum Associates.

Kim, S., Park, J., Han, S., & Kim, H. (2010). Development of extended speech act coding scheme to observe communication characteristics of human operators of nuclear power plants under abnormal conditions. *Journal of Loss Prevention in the Process Industries*, *23*(4), 539–548. https://doi.org/10.1016/j.jlp.2010.04.005

Kleinman, D. L., & Serfaty, D. (1989). Team performance assessment in distributed decision making. In R. Gilson, J. P. Kincaid, & B. Goldez (Eds.), *Proceedings of the interservice networked simulation for training conference* (pp. 22–27). University of Central Florida.

Korean Air Force. (2005) Tactics, techniques and procedures 3-3 vol 5, basic employment manual. Available online: falcon.blu3wolf.com/Docs/Basic-Employment-Manual-F-16C-RoKAF.pdf

Kozlowski, S. W. J., Watola, D. J., Jensen, J. M., Kim, B. H., & Botero, I. C. (2009). Developing adaptive teams: A theory of dynamic team leadership. In E. Salas, G. F. Goodwin, & C. S. Burke (Eds.), *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches* (pp. 113–155). Routledge.

Lahtinen, T. M., Huttunen, K. H., Kuronen, P. O., Sorri, M. J., & Leino, T. K. (2010). Radio speech communication problems reported in a survey of military pilots. *Aviation, Space, and Environmental Medicine*, *81*(12), 1123–1127. https://doi.org/10.3357/asem.2468.2010

Langan-Fox, J., Anglim, J., & Wilson, J. R. (2004). Mental models, team mental models, and performance: Process, development, and future directions. *Human Factors and Ergonomics in Manufacturing*, *14*(4), 331–352. https://doi.org/10.1002/hfm.20004

LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., & Saul, J. R. (2008). A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria. *Personnel Psychology*, *61*(2), 273–307. https://doi.org/10.1111/j.1744-6570.2008.00114.x

Louangrath, P. (2018). Reliability and validity of survey scales. *International Journal of Research and Methodology in Social Science*, *4*(1), 50–62. https://doi.org/10.5281/zenodo.1322695

Mansikka, H., Virtanen, K., & Harris, D. (2022). *Accuracy and similarity of team situation awareness in simulated air combat*. Manuscript submitted for publication.

Mansikka, H., Virtanen, K., Uggeldahl, V., & Harris, D. (2021a). Team situation awareness accuracy measurement technique for simulated air combat - curvilinear relationship between awareness and performance. *Applied Ergonomics*, *96*(3), 103473. https://doi.org/10.1016/j.apergo.2021.103473

Mansikka, H., Virtanen, K., Harris, D., & Jalava, M. (2021b). Measurement of team performance in air combat – have we been underperforming? *Theoretical Issues in Ergonomics Science*, *22*(3), 338–359. https://doi.org/10.1080/1463922X.2020.1779382

Mansikka, H., Virtanen, K., Harris, D., & Salomäki, J. (2021c). Live–virtual–constructive simulation for testing and evaluation of air combat tactics, techniques, and procedures, Part 1: Assessment framework. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, *18*(4), 285–293. https://doi.org/10.1177/1548512919886375

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *The Academy of Management Review*, *26*(3), 356–376. https://doi.org/10.2307/259182

Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., & Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *The Journal of Applied Psychology*, *85*(2), 273–283. https://doi.org/10.1037/0021-9010.85.2.273

Mathieu, J. E., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management*, *34*(3), 410–476. https://doi.org/10.1177/0149206308316061

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., & Abich, J. (2015). The psychometrics of mental workload: Multiple measures are sensitive but divergent. *Human Factors*, *57*(1), 125–143. https://doi.org/10.1177/0018720814539505

McGrath, J. E. (1964). *Social psychology: A brief introduction*. Holt, Rinehart and Winston.

McGrath, J. E. (1984). *Groups: Interaction and performance*. Prentice-Hall.

Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, *36*(4), 876–910. https://doi.org/10.1177/0149206309356804

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology.* Freeman.

Orasanu, J. (1995). Situation awareness: Its role in flight crew decision making. In R. S. Jensen & L. A. Rakovan (Eds.), Proceedings of the eighth international symposium on aviation psychology (pp. 734–739). Ohio State University.

Orasanu, J., & Salas, E. (1993). Team decision making in complex environments. In G. Klein, J. Orasanu, R. Calderwood, & C. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 327–345). Ablex.

Park, J., & Kim, Y. (2018). A novel speech-act coding scheme to visualize the intention of crew communications to cope with simulated off-normal conditions of nuclear power plants. *Reliability Engineering and System Safety*, *178*(C), 236–246. https://doi.org/10.1016/j.ress.2018.05.013

Parush, A., Kramer, C., Foster-Hunt, T., Momtahan, K., Hunter, A. C., & Sohmer, B. (2011). Communication and team situation awareness in the OR: Implications for augmentative information display. *Journal of Biomedical Informatics*, *44*(3), 477–485. https://doi.org/10.1016/j.jbi.2010.04.002

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. https://doi.org/10.1080/00140139.2014.991765

Porter, C. O. L. H., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P. J., West, B. J., & Moon, H. (2003). Backing up behaviors in teams: the role of personality and legitimacy of need. *The Journal of Applied Psychology*, *88*(3), 391–403. https://doi.org/10.1037/0021-9010.88.3.391

Rico, R., Sánchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge–based approach. *Academy of Management Review*, *33*(1), 163–184. https://doi.org/10.5465/amr.2008.27751276

Roberts, A. P. J., Webster, L. V., Salmon, P. M., Flin, R., Salas, E., Cooke, N. J., Read, G. J. M., & Stanton, N. A. (2022). State of science: Models and methods for understanding and enhancing teams and teamwork in complex sociotechnical systems. *Ergonomics*, *65*(2), 161–187. https://doi.org/10.1080/00140139.2021.2000043

Roberts, N. C., & Dotterway, K. A. (1995). The Vincennes incident: Another player on the stage? *Defense Analysis*, *11*(1), 31–45. https://doi.org/10.1080/07430179508405642

Rosenman, E. D., Dixon, A. J., Webb, J. M., Brolliar, S., Golden, S. J., Jones, K. A., Shah, S., Grand, J. A., Kozlowski, S. W. J., Chao, G. T., & Fernandez, R. (2018). A simulation-based approach to measuring team situational awareness in emergency medicine: a multicenter, observational study. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, *25*(2), 196–204. https://doi.org/10.1111/acem.13257

Rousseau, V., Aubé, C., & Savoie, A. (2006). Teamwork behaviors: A review and integration of frameworks. *Small Group Research*, *37*(5), 540–570. https://doi.org/10.1177/1046496406293125

Salas, E., Cannon-Bowers, J. A., & Johnston, J. H. (1997). How can you turn a team of experts into an expert team?: Emerging training strategies In C. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 359–370). Lawrence Erlbaum Associates, Inc.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their training and performance* (pp. 3–29). Ablex.

Salas, E., Prince, C., Baker, D. P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 123–136. https://doi.org/10.1518/001872095779049525

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a "big five" in teamwork? *Small Group Research*, *36*(5), 555–599. https://doi.org/10.1177/1046496405277134

Salmon, P. M., Stanton, N. A., Walker, G. H., Baber, C., Jenkins, D. P., McMaster, R., & Young, M. S. (2008). What really is going on? Review of situation awareness models for individuals and teams. *Theoretical Issues in Ergonomics Science*, *9*(4), 297–323. https://doi.org/10.1080/14639220701561775

Salmon, P. M., Stanton, N. A., Walker, G. H., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*, *37*(2), 225–238. https://doi.org/10.1016/j.apergo.2005.02.001

Serfaty, D., Entin, E. E., & Johnston, J. H. (1998). Team coordination training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 221–245). American Psychological Association. https://doi.org/10.1037/10278-008

Sexton, J. B., & Helmreich, R. L. (2000). Analyzing cockpit communications: The links between language, performance, error, and workload. *Human Performance in Extreme Environments: The Journal of the Society for Human Performance in Extreme Environments*, *5*(1), 63–68. https://doi.org/10.7771/2327-2937.1007

Shaw, M. E. (1976). *Group dynamics: The psychology of small group behavior*. McGraw-Hill.

Stout, R. J., Cannon-Bowers, J. A., & Salas, E. (2017). The role of shared mental models in developing team situational awareness: Implications for training. In *Situational awareness* (pp. 287–318). Routledge.

Sulistyawati, K., Wickens, C. D., & Chui, Y. P. (2009). Exploring the concept of team situation awareness in a simulated air combat environment. *Journal of Cognitive Engineering and Decision Making*, *3*(4), 309–330. https://doi.org/10.1518/155534309x12599553478791

Svensson, J., & Andersson, J. (2006). Speech acts, communication problems, and fighter pilot team performance. *Ergonomics*, *49*(12-13), 1226–1237. https://doi.org/10.1080/00140130600612671

Thornton, R. C. (1992). *The effects of automation and task difficulty on crew coordination, workload, and performance*. Old Dominion University. (Unpublished doctoral dissertation).

van den Oever, F., & Schraagen, J. M. (2021). Team communication patterns in critical situations. *Journal of Cognitive Engineering and Decision Making*, *15*(1), 28–51. https://doi.org/10.1177/1555343420986657

Weingart, L. R. (1997). How did they do that? The ways and means of studying group processes. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 19, pp. 189–240). JAI Press.

Zaheer, S., Albert, S., & Zaheer, A. (1999). Time scales and organizational theory. *The Academy of Management Review*, *24*(4), 725–741. https://doi.org/10.2307/259351

Zalesny, M. D., Salas, E., & Prince, C. (1995). Conceptual and measurement issues in coordination: Implications for team behavior and performance. *Research in Personnel and Human Resources Management*, *13*, 81–115.

Heikki Mansikka is a retired Air Force fighter pilot (LtCol). He is a researcher at the Systems Analysis Laboratory, Department of Mathematics and Systems Analysis, Aalto University, Finland. Heikki is also an Adjunct Professor at the Department of Military Technology, National Defence University, Finland. His research concentrates on human and team performance in air combat.

Kai Virtanen is a professor in the joint professorship of Operations Research with Systems Analysis Laboratory, Department of Mathematics and Systems Analysis, Aalto University, Finland, and the Department of Military Technology, National Defence University, Finland. His research interests include decision analysis, dynamic optimization, simulation optimization, and human performance in complex systems.

Don Harris is Professor of Human Factors at Coventry University, UK. He is a fellow of the Institute of Ergonomics and Human Factors and a chartered psychologist. Don Harris is heavily involved in the design and development of the next generation of advanced pilot interfaces and the certification of civil flight decks.