11 Team Performance

Mental Performance N+1

Deanna Beech, Heikki Mansikka, Don Harris, Kai Virtanen, and Nathalie Pattyn

"No man one is an island."

– John Donne

John Donne's poignant statement from his 1624 work, Meditation, has been frequently cited to underscore the idea that no individual can thrive in isolation. Achieving a fulfilling life and realizing our potential necessitates the aid and guidance of others. Through interactions with friends and mentors, we not only absorb knowledge but also refine it, propelling ourselves toward excellence. In essence, our personal achievements are a testament to the cumulative support and involvement of everyone who has aided us on our journey.

Those that share in supporting us toward our goals, that share this goal with us, can be described as "being on our team". A team is not just a group of associated people. A team is "a distinguishable set of two or more people who interact, dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission" (Salas, Dickinson, Converse, & Tannenbaum, 1992, p. 4). From this generally accepted definition it is clear that even individuals who are performing what appears to be a very individual effort, such as a writer or a pole vaulter, are impacted by the team around them that support their effort toward the shared objective. The impact on the individual is intensified as the interdependence of the individuals increases as is the case with a football team or a specialized military unit.

When a team is performing well, individual members are motivated and engaged. They are receptive to trying new skills and open to learning from others. This not only benefits individual performance, but it also provides social and emotional support to buffer the individual against the intense stress that is present at elite performance levels (Eisenberger et al., 2007; Eisenberger & Cole, 2012). The opposite is also true. When a team is performing poorly, it can have a significant negative impact on the individual. Team members can become unmotivated to work together or even engage in open frustration or hostility, leading to a decrease in mental and physical performance and an increase in stress levels. Studies have even shown that individuals in difficult, negative, or high-conflict team conditions may experience burnout and other mental health issues (Burke et al., 2017; DeChurch, et al., 2010; Gully et al., 2002, Salas et al., 2012; Salanova et al., 2013). As such, it is important to prioritize team performance to maintain and improve individual mental performance. The following chapter focuses on the contributions of team dynamics to performance, team performance theory, and team skills, along with how to assess and intervene to improve team performance.

DOI: 10.4324/9781003378969-14

Compounding stress

Team-based stress and individual performance

Much research has focused on understanding the interplay between the physiological reactions, cognitions, and behaviours that contribute to elite performance. Physiologically, a compelling body of research points to the fact that our actions and reactions are a "set of collaborative processes that strategically deploy resources to preserve functionality in an unpredictable and dynamically changing environment" (Kiely et al., 2018, p. 13). This constant adaptation to the world around us has been described as allostasis. What appears to be homeostasis, also referred to as allostasis, is really the neurobiological imperative to "sensitively pre-empt and respond to emerging challenges by orchestrating multi-level system-wide coordinated compensations" (Kiely et al., 2018, p. 13).

When too much adaptation is required of the individual's system an "allostatic load" is created (McEwen, 2000). For a more detailed description of the physiological and psychological mechanisms subtending this, we refer to Chapter 2 and its section about stress. If this allostatic load is chronic (continually being required to adapt) the chronic over-stimulation of the autonomic nervous system can, over time, result in negative outcomes such as difficulties with sleep, concentration, and poor working memory (Giuliano et al., 2017). Without awareness this trajectory can even lead to physiological illnesses such as hypertension, cardiovascular disease, reduced testosterone, and the development of type II diabetes (Fisher et al., 2009). The psychological costs are also well documented and can include difficulties with insomnia, substance abuse, low self-esteem, reduced energy and motivation, emotional reactivity and anger, hypervigilance, depression, anxiety, sexual dysfunction, cognitive impairments, and combat stress reactions (Frueh et al., 2020). It is clear, failure to effectively incorporate recovery in our performance routines can degrade our ability to reach ideal performance states, and it reduces well-being and longevity in the team (we refer to Chapters 2, 3, and 5 for a more detailed discussion).

Theories of team performance

Our understanding of the various individual and team-based skills that are present in a high-performing team has evolved in an explosion of research on this topic over the last century (Mathieu et al., 2017). However, the starting point is often credited to a foundational series of studies conducted by Harvard University researchers at the Hawthorne Works electrical plant between 1924 and 1933. Originally intended to probe the link between work conditions and employee productivity, these experiments revealed unexpected outcomes. It was found that regardless of whether work conditions were enhanced or degraded, productivity surged. Elton Mayo, who spearheaded the research, discerned that the uptick in productivity wasn't tied to the specific environmental modifications (Gale, 2004). Instead, increased performance stemmed from the workers' awareness that they were under observation which led them to desire to meet or exceed the researchers' expectations. This phenomenon, now coined the "Hawthorne effect", underscores a seminal principle in team performance research: as the act of observation or scrutiny impacts team performance, by aligning these observations with performance metrics and goal achievement, team performance is generally enhanced (Marks et al., 2001; Mathieu et al., 2019; Stewart et al., 2023).

Since these early studies the evolution of team performance research has been described as branching into three distinct trajectories with "very limited cross-pollination" (Mathieu et al., 2017, p. 453; McGrath, 1997; Arrow, McGrath & Berdahl, 2000; Steiner, 1974). In their retrospective analysis for the Journal of Applied Psychology's centennial, Mathieu et al. (2017) delineated these trajectories as the Individualistic Orientation, Group Orientation, and Task Contingency Approach. As makes sense, the research focus of the Individual Orientation branch was on the individual's attitudes and behaviours of the team and its members. From this approach "groups were considered as social influences on individual-level processes" (p. 454). The Group Orientation trajectory focuses on small-group interaction dynamics, often employing complex Interaction Process Analysis (IPA) to yield intricate visualizations of group dynamics after a comprehensive data examination. The third, and perhaps most influential trajectory – the Task Contingency Approach, emphasizes the consistent properties that are found across groups with a primary focus on the factors that enhance task outcomes. It was this perspective that led Hackman and Morris to introduce the Inputs–Process–Outputs (IPO) model in 1975, ushering in a wave of research designed to consider the interplay of team-based skills. They conceptualized inputs as encompassing member characteristics (such as their knowledge, skills, abilities, and values), along with group-level determinants like size and structure, and external influences like stress and task attributes. These inputs, they argued, underwent transformation through group interactions, culminating in performance outputs. This framework was pivotal in shaping subsequent team performance research and mirror our understanding of what has been identified as spontaneously "emerging states" that develop as teams operate. Emergent states reflect a synthesis of team members' emotions, interactions, and behaviours – reflecting elements like trust and team cohesion. As Marks and his team described in 2001, emergent states can be understood as the "cognitive, motivational, and affective states of teams [that are…] dynamic in nature and vary as a function of team context, inputs, processes, and outcomes" (p. 357).

The IPO model was advanced to reflect the increased evidence that the team's processes and emergent states are continuous variables that mediate the inputs brought to the team as the team goes to create the required output. This modification was dubbed the Input–Moderators–Output–Input (IMOI) model by Ilgen et al. (2005), and it was further developed by Mathieu et al. (2008) by drawing attention to the team's Affect, Behaviours, and Cognitions (ABCs) which are seen as central to productive outputs (see Figure 11.1, Grossman et al., 2019, p. 247).

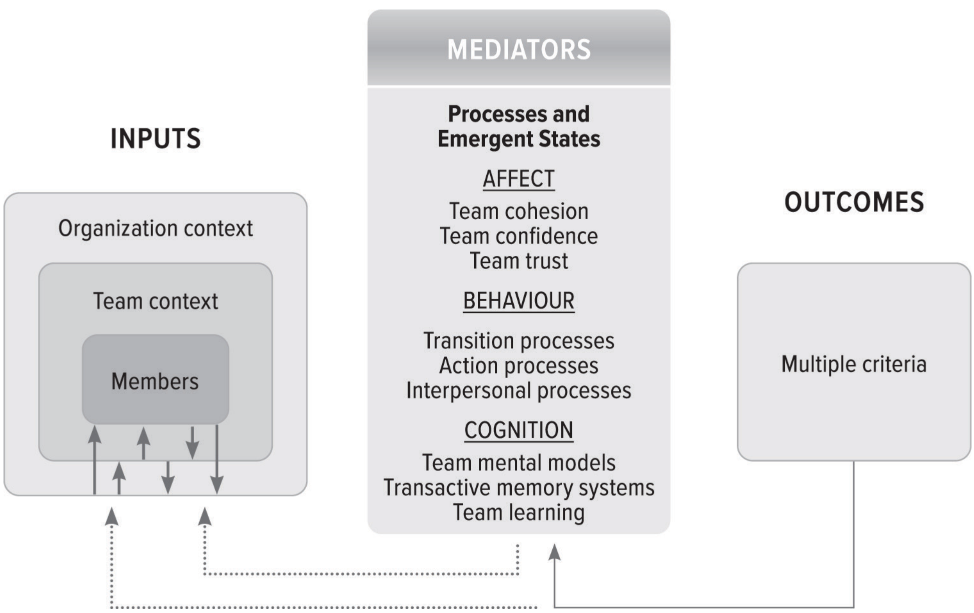

Figure 11.1 Input–Mediator–Output–Input model, originally adapted by Grossman et al., (2019) from the work of Ilgen et al., (2005) and Mathieu et al., (2008) to emphasize the role of affect, behaviour, and cognition in team processing.

Building on foundational research from the IPO/IMOI models, Salas and colleagues proposed their ground-breaking research that began to identify the interactive nature of team processes in the "Big 5 in Teamwork" (2005a). They proposed that team performance takes place through a series of ever present and interactive emergent states (2005a). The model emerged from a thematic analysis of the variables that demonstrated the greatest benefits to team effectiveness. They called the model the "Big 5" to draw attention to the five key Teamwork Components (leadership, mutual performance monitoring, team orientation, back-up behaviours, and adaptability). In addition to these emergent states, their model also identified three Coordinating Mechanisms (shared mental models, communication, and mutual trust). According to the "Big 5" model, each of these factors is important for effective team performance, and each can be developed and improved through training and other interventions. Since 2005 this framework has been instrumental in shaping strategies for enhancing team cohesion, flexibility, and overall performance across diverse settings and continues to influence the conceptualizations of team performance.

Ten years later Salas and his colleagues (2015b) took another look at the literature and expanded their framework. They suggested that there was now sufficient meta-analytic data to support the presence of nine key elements of teamwork – dubbed the "9 C's" (see Figure 11.2, Salas, E., 2015b, p. 602). This framework provides a richer understanding of the various nuances of team performance by integrating the literature into six core Processes and Emergent States

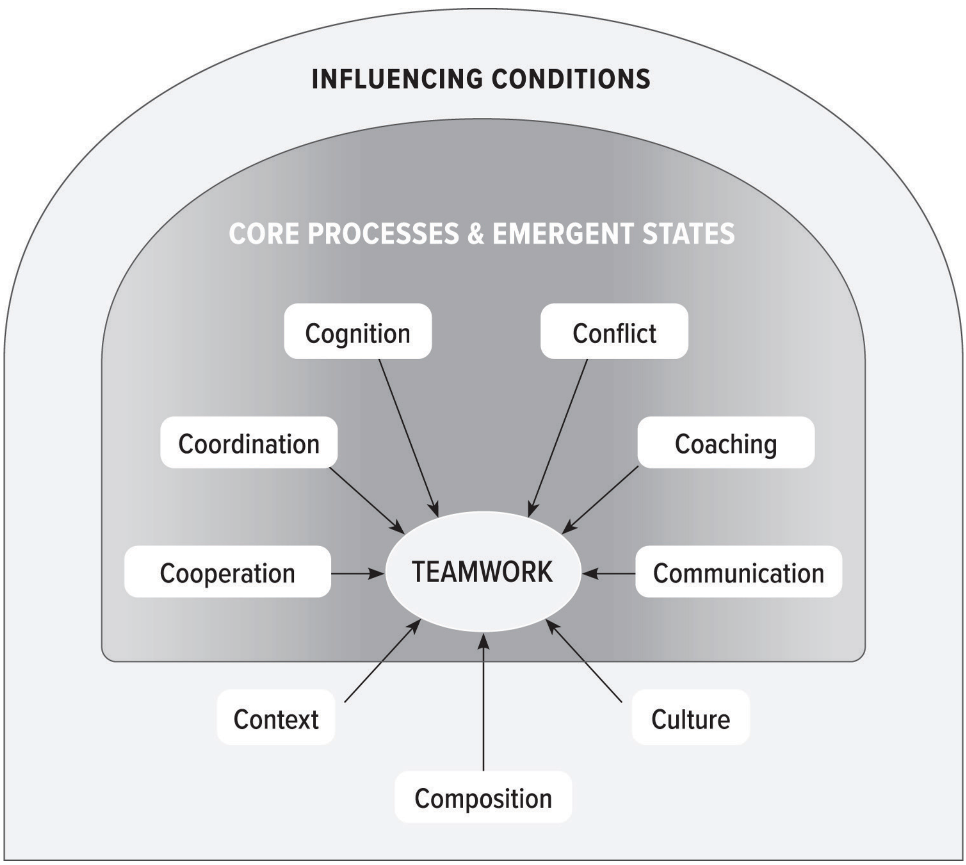

Figure 11.2 Heuristic of the critical considerations of teamwork to illustrate the 9 C's of teamwork as described by Salas, E. (2015).

(cognition, coordination, cooperation, communication, coaching, and conflict) along with three key Influencing Conditions (team context, composition, and culture) (Salas, E., 2015b, p. 602). Each of the processes and emergent states comprise related variables providing an integration of a vast amount of research in this field. For example, the emergent state of Cognition includes the concepts of shared team knowledge, as well as shared understandings (mental models) of the team's roles, norms, and goals. Coordination connotes more than the "behavioral mechanisms necessary to perform a task and transform team resources into outcomes" (Salas, E., 2015b, p. 606); it also involves planning and communication as well as anticipating upcoming needs and adjusting to the circumstances. Cooperation is described as a "motivational driver" of teamwork as it engages the "attitudes, beliefs, and feelings of the team that drive behavioral action" (Salas, E., 2015b, p. 604). Communication involves all the interactive behaviours and thought processes that go into receiving information to form and re-form the team's attitudes, behaviours, and cognitions. Coaching involves the enactment of leadership behaviours to "establish goals and set direction" in a way that empowers the team to reach their goal (Salas, E., 2015b, p. 603). The inclusion of Conflict in these emergent states addresses the inevitable differences of opinion and breakdowns in communication that lead to strife between teammates. Defined as "perceived incompatibilities in the interests, beliefs, or views held by one or more of team member" (Jehn, 1995, p. 257). Conflict, no matter what type, has been found to have a "strong negative correlation with team performance as well as team member satisfaction" (Salas, E., 2015b, p. 605). However, the current understanding of the effect of conflict on team performance is more complicated as relationship conflict consistently has a strong negative correlation to performance outcomes, while task-based conflict has a curvilinear relationship to performance when relationship conflict is low. The implication is that when a team can resolve relationship difficulties, they are able to effectively disagree about ideas related to tasks in ways that improve the eventual solution to the problem by incorporating different perspectives (Caesens et al., 2019).

The three key influencing elements reflect the "factors that have an impact on the core teamwork processes and emergent states"; specifically the team composition, context, and culture (Salas, E., 2015b, p. 610). Team composition is considered critical to the team's success as it is necessary that the team is composed of members with the knowledge, skills, and abilities (KSAs) to achieve the goal. The context in which the team functions serves to "shape the very nature in which team members interact with one another" (Salas, E., 2015b, p. 611). This includes aspects such as whether the team engages with each other in person or virtually, the resources available to the team, as well as the organizational climate within which the team must function. Culture refers to the assumptions held by the people in the organization about how to relate to each other and the environment in which they work. It is a "driving force for [team] member values, norms, and behavior" (Salas, E., 2015b, p. 613). Salas and his colleagues propose this framework to concisely organize a vast body of information pertaining to team performance in a way that makes the information manageable for those trying to enhance team performance.

Recently, team performance research has been working to effectively reflect the interactive nature of team variables, through the development of more dynamic conceptualizations such as the Complex Adaptive Systems (CAS; Ramos-Villagrasa et al., 2018) and small complex systems (Mathieu et al., 2019). These conceptualizations suggest that teams are always in an evolving and adapting state, capturing the creative multilevel expression present in the team's dynamics. For example, the model proposed by Mathieu and his colleagues (2019) organizes team interactions into three primary categories and three overlapping regions. The primary categories include the team's *structural features* (task scope and complexity, interdependence, and knowledge management systems), the *compositional features* (member attributes, diversity, and fault-lines/subgroups), and the *mediating mechanisms* of member interactions (team processes, information sharing/communication, emergent states and conflict). The overlapping regions suggest primary interactions

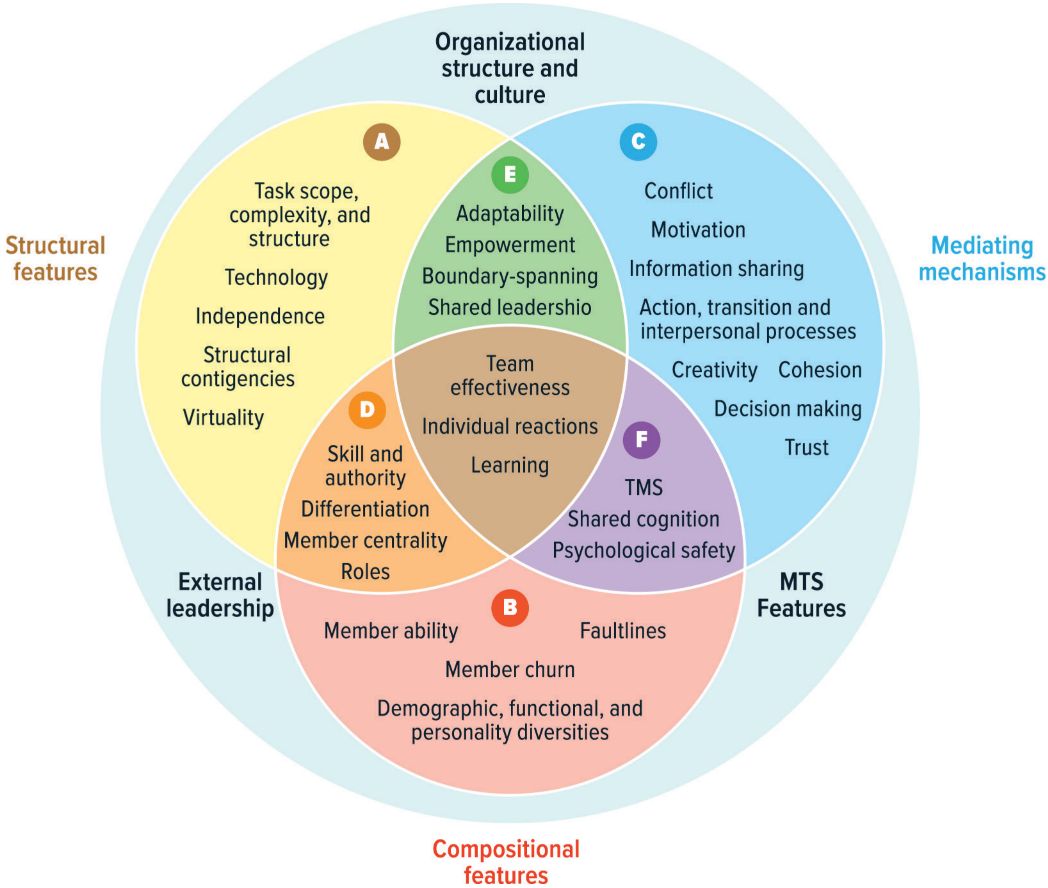

Figure 11.3 Co-evolving team compositional and structural features, mediating mechanisms, external influences, and outcomes. Adapted from Mathieu et al. (2017). Abbreviations: MTS, multiteam systems; TMS, transactive memory system.

between the categories. They proposed that the interplay of these dynamics results in the teams' effectiveness and learning and can even account for individual team member reactions (see Figure 11.3, Mathieu et al., 2017, p. 19).

Further, it has been proposed that when the dynamics are optimal, and a team is in a high-performance moment they experience a state of "flow" that is like that of an individual's flow experience (van den Hout, Davis & Weggeman, 2018). As is true of the subjective experience of the individual who achieves the flow state (Csikszentmihalyi, 1990, 2014; Jackson & Csikszentmihalyi, 1999), to reach this state as part of a team is intrinsically rewarding and enhances the team members' commitment to and satisfaction with the group which perpetuates the teams' continued engagement in high performance. Sawyer pioneered the application of this concept to groups and teams by describing "a collective state that occurs when a group is performing at the peak of its abilities" (Sawyer, 2003, p. 167). Early studies focused on the antecedents to achieving flow in a team (Aubé et al., 2014) and found consistent support for the correlation between flow and high performance of a team (Landhaüßer and Keller, 2012).

Expanding on this, van den Hout and his colleagues conducted a review of this literature and proposed a conceptualization of team flow (2018). They proposed that "team flow be defined as a shared experience of flow derived from an optimized team dynamic during the execution of interdependent personal tasks" (p. 400). They identify seven prerequisites and four characteristics of team flow. The prerequisites include a collective ambition, common goal, aligned personal goals, high skill integration, open communication, safety, and mutual commitment. The characteristics experienced when the group is in a flow state include a sense of unity, sense of joint progress, mutual trust, and holistic focus.

Thus, the principle that no individual operates in isolation is foundational to understanding team dynamics and performance. Teams are more than mere collections of individuals; they represent intricate networks of interdependent members, whose performance can be enhanced or hindered by the team's functioning. Effective team performance promotes individual growth, provides emotional support, and increases resilience against stress. In contrast, poorly performing teams can contribute to individuals becoming unmotivated, poor performing, and burnt out. Our understanding of team functioning has evolved from early seminal research like the Hawthorne studies to more complex dynamic models that elucidate the multifaceted nature of team dynamics. Current research emphasizes the interplay between team structure, composition, and interactions, aiming to harness the synergistic potential of teams to reach optimal performance, where collective participation enhances both individual well-being and team performance.

Assessing team performance

How to know if the team has a problem?

Informal team assessment

As we delve deeper into the crux of team dynamics, the crucial next question becomes – *How to know if the team has a problem?* Answer: actively observe and ask.

When a team has developed high-performing teamwork skills, you can "see" it. The team acts like a cohesive unit with seamless collaboration and communication, team members work fluidly together, proactively supporting each other, and share a deep understanding of their collective goals and norms. They exhibit a strong team identity, seek to work with each other, and show that they enjoy their interactions. Team members willingly offer needed information and feel safe to explore even unpopular viewpoints or differences in opinions trusting that they will continue to be accepted by the team as it seeks to find harmonious solutions and maintain positive interpersonal relationships. Their motivation is aligned with team objectives, and they engage in collective problem-solving, valuing each member's contribution. Shared accountability for outcomes, a clear understanding of individual roles, and supportive leadership are evident, promoting an environment of collective effort. Decision-making is inclusive, ensuring all voices are heard, and the team regularly engages in joint celebrations of successes and reflections for further performance development, reinforcing their unity and commitment to shared success.

Conversely, in low-performing teams, these elements manifest differently. Collaboration is disjointed, communication sporadic and unclear, and there is an evident lack of mutual support and trust. Conflicts go unresolved or may be handled in a way that further damages relationships. The group identity is weak, with members focusing more on individual goals than on the team's objectives. Leadership is often perceived as unsupportive or inconsistent, and decision-making processes are often exclusive, leaving team members feeling undervalued or unheard. Social interactions appear perfunctory or individualized, further eroding the sense of team unity. This stark contrast in behaviours and attitudes helps to highlights the importance of observation in identifying a team's performance level (see Addendum A for a worksheet to organize team observations).

Hopefully, with observations in hand, we can fill in the "what" things are working great, and "why" other things are not by directly asking. With the right questions, in the right setting, the problems are usually volunteered with little prompting. Conducting one-on-one conversations in a secure and private manner to foster honest communication about the team's challenges we can shed light on the difficulties beneath the surface of everyday interactions. A straightforward inquiry like, "How are things on your team?" or "How's the team doing?" can elicit the needed insights. Guaranteeing anonymity for the individual team members by ensuring that all data will be presented in an aggregate format will facilitate a sense of safety in the discussion that may lead to richer disclosure of the team's inner workings.

Another useful structured format for starting these conversations is provided in publicly available work that resulted from an extensive research project conducted by Google. Titled "Project Aristotle" and led by Abeer Dubey, Google set out to determine the attributes of their high-performing teams. This initiative was rooted in the hypothesis that understanding the dynamics behind successful teams could not only enhance collaboration within Google but also provide a blueprint for other organizations striving for excellence. The research was expansive. It spanned several years beginning in 2012 and involved data collection from more than a hundred teams across the company. Their results, published by Charles Duhigg in an article of The New York Times Magazine (2016) and on their website, suggested that the key to effective teams isn't so much who is on the team, but how the members interact, structure their work, and view their contributions. Specifically, the research identified five core dynamics of successful teams: (1) Psychological safety, where team members feel safe to take risks and be vulnerable with each other; (2) Dependability, with peers reliably completing quality work; (3) Structure and clarity, characterized by clear roles, plans, and goals; (4) Meaning, the sense that work has personal significance to each member; and (5) Impact, the belief that one's work makes a difference. Their findings emphasized the importance engaging with teams in support of creating a trusting environment where open communication, dependability, and meaningful work align to produce a significant impact which fundamentally shifts the focus from individual to collective team performance (Duhigg, 2016).

When accessible, supplementary data pertaining to individual team members – encompassing personality assessments, 360-degree feedback, and aptitude evaluations – can enhance a practitioner's comprehension of potential contributions by these members to the challenges currently faced by the team. This information can guide practitioners towards targeted discussions and interventions that may prove advantageous in addressing these issues. For an example of this approach see the case study below titled *Transitioning from Individual to Team Training*.

Formal team assessments

Formal team assessments can be a vital component for any organization striving to ensure their teams are operating at peak efficiency. By methodically evaluating a team's dynamics, performance, and overall effectiveness, practitioners can identify not only areas in need of enhancement but also enhance their strengths. This holistic approach involves selecting appropriate assessment tools, designing the assessment process, collecting and analyzing data, and then sharing the results with stakeholders to get buy-in for a plan of action. Ultimately, such assessments serve as a compass for continuous improvement, aligning resources and goals, identifying difficulties and needed resources, and paving the way for superior team performance and collaboration.

A simple Google search will immediately provide numerous organizations that offer team assessment questionnaires. It is important to vet these sites as most of these questionnaires are not validated instruments. However, several well-researched instruments are available and can be administered for a fee due to their proprietary nature. The benefit of this approach is that the practitioner does not need to be an expert at team performance or assessment and can still obtain actionable information that would support tailored training. Many of these organizations also offer training programs for more in-depth use of the results. One of the most comprehensive of these measures is the Optimizing Team Performance Profile – OTP Profile (Beech, 2023). This instrument is based on an extensive review of the evidence-based team skills that are present in high-performing Special Forces units and assesses the team's Leadership Engagement (Organizational Identification, Leadership Trust, Coaching, and Empowerment), their Mission Focus (Shared Purpose, Shared Vison, Goal Direction, and Roles & Expectations), the team's Action Processes (Monitoring, Joint Action, Communication, and Reasoning), and their Relational Elements (Team Cohesion, Emotional Intelligence, Trust & Safety, and Conflict Repair). The Denison Organizational Culture Model – Denison Model, was developed based on a line of research examining high-performing organizations (Denison & Ko, 2006). This instrument breaks team performance into Adaptability (creating change, customer focus, organizational learning), Mission (strategic direction & intent, goals & objectives, and vision), Consistency (core values, agreement, coordination & integration), and Involvement (empowerment, team orientation, and capacity development). The Team Diagnostic Survey (Wageman et al., 2005) focuses on six team conditions that are broken into three Task Processes (Effort, Strategy, and Knowledge & Skills) and three Criteria of Effectiveness (Quality of Group Process, Member Satisfaction, and Task Performance). The instrument provides an assessment of a team's structure, support, and leadership, along with various indicators of the members' work processes and their emotional responses to the team and its work that is grounded in team research and theory. Other measures of team performance that are focused on specific industries can also be found, such as those for educational institutions like DORA by Algozzine et al. (2012) and medical settings like STAT by Reid et al. (2012).

For those adept at managing survey data, Mathieu et al. (2020) have provided team practitioners with a valid open-source instrument that allows the team practitioner to gather general data on team-based skills. They conducted an extensive review of the literature and identified three critical areas of team processes: transition processes, action processes, and interpersonal processes. Transition processes "occur prior to or between performance episodes and have a dual focus whereby members reflect on and interpret previous accomplishments as well as prepare for the future" (p. 3). This includes questions focused on mission analysis, goal specification, and strategy formulation and planning. The action processes "describe the behaviors that members engage in while working toward goal accomplishment" (p. 4). Included in this section are the team variables of monitoring progress towards goals, systems monitoring, and team monitoring and backup behaviours. The final section, Interpersonal processes includes the team's ability to manage conflict, motivate and build confidence, and affect management. The scale was shown to be valid as a 10-item survey, a 30-item survey, or a comprehensive 50-item survey. Choosing the level of survey would depend on the time requirements and the level of detail needed. The full survey is available in their article (Mathieu et al., 2020). For an example of how to use this method for team assessment and intervention please see the case study below titled *Trouble with Boundary Spanning*.

Team assessment with task analysis

For those with more advanced skill in organizational psychology or human resource management, incorporating an evidence-based understanding of the steps involved in the taskwork that is necessary to meet team objectives (task analysis) can greatly benefit the team assessment by allowing interventions to be more targeted to specific outcomes. Task analysis entails dissecting the team's work into individual tasks, documenting each step, and identifying the requisite KSAs to accomplish them. This detailed breakdown greatly benefits the development of the training program and formal team performance measures can be added to provide a clear picture of the team's current functioning and a baseline from which to judge the effectiveness of the interventions. Those interested in a deeper dive into the methods and benefits of task analysis can refer to "A Guide to Task Analysis" by Kirwan and Ainsworth (1992). The interplay between task analysis and team assessments is exemplified in the case study below titled *Developing High Performing Flight Crews*.

Together, these informal and formal team assessment methods form a tiered approach to identifying the strengths and developmental needs of teams moving toward higher levels of performance. This allows practitioners the flexibility to consider the level appropriate level of interventions given the team's needs, the resources available, and the practitioner's level of experience.

Team training interventions

Extensive research, including meta-analyses by McEwan et al. (2017), Delisa et al. (2010), Lacerenza et al. (2018), and Hughes et al. (2016), indicates that team training interventions yield significant, positive effects on both teamwork and outcome performance. Developing a consultation with a team or organization involves a systematic approach. First, it would be essential to ensure that all stakeholders have "bought in" to the process, and then a thorough team needs assessment is conducted to identify the specific skills and capabilities that need to be improved. This is followed by engaging in detailed discussions with both the team leaders and members to ensure alignment between the training objectives and the stakeholders' expectations, as suggested by Lacerenza et al. (2018).

The next step is to design the training program utilizing a blend of delivery methods, including informational presentations, demonstrations, and practice exercises. Providing a combination of training methodologies is recommended as the most effective approach (Salas et al., 2012). In the program development it is also recommended that the material include a focus on the theoretical framework of team performance to support the transfer of insights to application, as discussed by Hughes et al. (2016) and Lacerenza et al. (2018). For a detailed guide to team training program development see the *Team Training Essentials: A research-based guide* (Salas, E., 2015a). Additionally, while there is an overwhelming number of team training programs available on the internet, they are not all equal. Below is a review of several modalities that have reached the level of well-codified programs with consistent, robust findings for improved team performance: Leadership training, Team coordination and adaptation training, Team building, Crew resource management, and Team debriefing.

Leadership training

Leadership training stands as a cornerstone intervention across organizations, deeply woven into the fabric of operational success. Its pivotal role is not just a matter of consensus, it's also supported by a substantial body of empirical research. Multiple meta-analyses have identified robust findings that support the positive relationship between good leadership and good team performance (Dunst et al., 2018; Fischer & Sitkin, 2023), while bad leadership correlates with poor performance (Burns, 2017; Fischer & Sitkin, 2023; Schyns & Schilling, 2013). Literally hundreds of articles and books have been published on this topic, but if you're looking for a practical guidance on how to develop and implement leadership training programs consider *Leadership Training Design, Delivery, and Implementation: A Meta-Analysis* (Lacerenza et al., 2017) and *Unlocking Human Potential through Leadership Training & Development Initiatives* (Day et al., 2021).

Team coordination and adaptation training

Team coordination and adaptation training aims to enhance team dynamics by teaching members how to streamline their collaborative efforts. The primary focus is on optimizing coordination strategies and minimizing the need for communication while improving efficiency and effectiveness (Salas et al., 2005b). It is suggested that the focus of this training be on the learning cycles embedded within specific team activities to foster such skills. However, a notable problem with this approach is the tendency to provide instructional feedback at an individual level rather than addressing the team as a whole. This individual-centric feedback overlooks the interdependent nature of team processes and fails to reinforce collective behaviours and shared understandings that are critical for team adaptation and synchrony. It's essential for such training to incorporate collective feedback mechanisms that reflect the team's performance as a unit, ensuring that the training outcomes align with the goals of enhanced team coordination and adaptability (Beech et al., 2023).

Team building

Training in team building focuses on cultivating "improvement within a team, providing individuals closely involved with the task with the strategies and information needed to solve their own problems" (Lacerenza et al., 2018, p.13). Team building centres on four basic approaches to team process: goal setting, interpersonal-relationship management, role clarification, and problem-solving.

Interventions focused on goal setting build a context and help to identify components that move the team toward the goal. It is also important that the goal be just challenging enough to focus the team's energy and attention (Vashdi et al., 2013). Maintaining an ongoing dialog about the team's objectives fosters the development of a shared mental model of the objective which, as described above, is one of the most robust variables in team efficacy. It has been proposed that this is accomplished through four specific mechanisms: direct attention and effort toward the goals, energizing the group to meet the goals, fostering effective persistence to keep moving toward the goals, and effecting actions through discovering, sharing, and using knowledge related to goal achievement (Lacerenza et al., 2018).

Interpersonal-relationship management is the process of directly and openly discussing affective interaction to develop a solid trust in the members of the team. As described above, trust enables the team to overcome uncertainty and accept vulnerability which facilitates coordination and communication. Trust does not mean the absence of conflict, rather it enables team members to overcome their uncertainty in expressing their views and accept the vulnerability of not always knowing or being right (DeJong et al., 2016). With solid trust the team can engage in healthy conflict as they problem-solve to reach the objective (Seitchik, 2019).

Role clarification is also thought to develop the teams' shared mental model and coordination by identifying who does what to get to the goal (Salas et al., 2015a). Also, rather than creating rigidity, role clarification allows for a better understanding of individual contributions which fosters the ability for other team members to provide backup behaviours and monitor workload distribution.

Problem-solving emphasizes the involved planning and action of the team members in identifying way to achieve their goals (Klein et al., 2009). This brainstorming provides the team members the opportunity to step back and identify specific processes, outcome levels, and resources needed for their work, which improves motivation and commitment to goal obtainment (Salas et al., 2005a).

Klein and colleagues (2009) conducted a comprehensive review and meta-analysis on team building and found that all of the four components significantly contributed to team process improvements, with goal-setting activities and role clarification showing the greatest benefits. Given the specific focus on processes related to task completion this type of team training can be particularly beneficial for ad hoc teams and newly formed teams.

Crew resource management (CRM) training

CRM was born out of the findings and recommendations from the National Transportation Safety Board's investigation into the preventable crash of United Airlines flight 173 in 1978 (NTSB, 1978). They concluded that the crash was caused by poor communication and situational awareness.

Extrapolating from this report others have also identified the contributions of rigid leadership style in which the captain failed to accept input from junior officers (Jedick, 2014). In response, training in CRM was developed and United Airlines, the airline responsible for the crash, was the first to adopt it. Shortly thereafter all other major airlines began training in CRM and it became mandated by the Federal Aviation Administration, National Aeronautics Space Administration, and Military Aviation. The concept then spread to other industries such as health care (TeamSTEPPS see O'Dea et al., 2014; Epps et al., 2015) and emergency response services (Griffith et al., 2015).

While early meta-analyses into the effectiveness of this approach found questionable results (O'Conner et al., 2008), the training has matured. More recent meta-analyses find that with CRM "positive and significant medium-sized effects were found for teamwork interventions on both teamwork and team performance" (McEwan et al., 2017, p. 1), with a large effect on knowledge and behaviours (O'Dea et al., 2014). Specifically, CRM has been described as:

> A flexible, systematic method for optimizing human performance in general, and increasing safety in particular, by (1) recognizing the inherent human factors that cause errors and the reluctance to report them, (2) recognizing that in complex, high risk endeavours, teams rather than individuals are the most effective fundamental operating units and (3) cultivating and instilling customized, sustainable and team-based tools and practices that effectively use all available resources to reduce the adverse impacts of those human factors (Marshall, 2009, p.22).

The content of the training will vary depending on the team context and specific needs of the intended team, but it usually includes the core variables of teamwork, leadership, situational awareness, decision-making, communication, and personal limitations (Salas et al., 2006). For those interested in establishing a CRM training program the chapter on *The Design, Delivery and Evaluation of Crew Resource Management Training* by Shuffler, Salas, and Xavier (2010) in the book *Crew Resource Management* provides an excellent guide; and Schuermann and Marquardt's 2016 work adds additional support with lessons learned and successful factors for CRM training as identified by expert CRM trainers.

Team debriefs

Team debriefing emerges as a critical mechanism in the learning and development landscape, particularly for collective settings where experience is a shared. Eddy et al. (2013) have positioned debriefing not merely as a reflective practice but as a powerful catalyst for expediting a team's learning curve. This is not hyperbole; Tannenbaum and Cerasoli (2013) quantify this advantage, citing a performance uptick of 20–25% when debriefing is employed effectively. The essence of debriefing lies in its ability to foster a robust shared mental model within the team, clarify roles and responsibilities, reinforce effective strategies, and pinpoint areas needing improvement – all through the prism of active learning (Lacerenza et al., 2018).

The debriefing process entails a collaborative analysis of events or performance periods to dissect outcomes, identify both strengths and weaknesses, and develop actionable strategies for future endeavours. This methodology, with its roots in military applications dating back to the 1970s, has been refined through extensive research, including insights from Morrison and Meliza (1999), to establish evidence-based best practices. A conducive environment characterized by trust, safety, and a tolerance for conflict is pivotal for candid dialogue and constructive feedback, as suggested by Arafeh et al. (2010). Further emphasis on teamwork, alongside task work, ensures a holistic approach to improvement (Reyes et al., 2018).

The structure and leadership of the debriefing session are equally important. While there is value in team-led debriefs, especially when the leader is well-versed in debriefing (Eddy et al., 2013), facilitator-led sessions often yield greater benefits (Fanning & Gaba, 2007). Structured, balanced

1. Set the stage (30 to 60 seconds)

- Explain why you are conducting a debrief and what the team will be discussing.
- “This is a quick opportunity to learn from our experience.
  Let’s look at how we handled this [situation, project, event, meeting, shift]: what we did well or could improve.”
- “Let’s consider how we worked as a team, in addition to any technical issues”
- If there are any boundaries or “non-negotiables,” let the team know what’s off limits.Identify our main tasks?

Basic assumption: “We’re all competent and well intentioned people who want to do our best. This is about getting better at what we do.”
Identify the key challenges that we expect to face?

2. Ask the team for their observations (5-20 seconds)

- What **happened**?
- What did we do **well**? What **challenges** did we face?
- What should we do differently or focus on +?
- What could **help us** be more effective? Anything we need?

3. Add your observations/recommendations and confirm understanding (5-10 minutes)

- Reinforce their observations, or if you noticed something different, share your view of what happened or needs to happen in the future.
- Be sure any feedback you provide is clear, actionable, and focuses on the work, not personal traits.

4. Summarize any agreed-upon actions or focus for the future (5 minutes)

- Be clear about who will do what, when...and how this will help the team.
- Specify when and how you will follow up to assess progress (e.g., next debrief?).

Tip 1: Ask the team for their perceptions first. Then if possible, acknowledge one thing that you could have done differently or that you will focus on in the future. This will make it easier for team members to voice their own observations or concerns.

Tip 2: Tip: If the team doesn’t discuss teamwork, ask “how well did we work together as a team?” Perhaps ask one or two specific questions such as:

HOW WELL DID WE...	**HOW CLEAR WERE OUR...**
• Communicate/share info	• Roles/assignments
• Monitor/provide backup	• Goals/priorities
• Coordinate with “outsiders”	
• Speak up/challenge one another	
• Ask for/offer help	
• Handle conflict	
• Share/allocate resources	
• Prepare/plan	

Figure 11.4 Quick Team Debrief Outline. Originally published by Reyes, Tannenbaum, and Salas (2018, p.50). Adapted from www.gOEbase.com. Permission granted.

discussions that methodically address both the highs and lows of performance, as advocated by Lacerenza et al. (2018) and Reyes et al. (2018), culminate in a more productive debrief. Documentation of the outcomes serves as a vital reference point for continuous improvement.

Considering the substantial benefits for the relatively minimal investment of time, averaging around 18 minutes (Tannenbaum & Cerasoli, 2013), the case for integrating team debriefing into routine practice is compelling. For organizations looking to implement or refine debriefing procedures, Reyes, Tannenbaum, and Salas (2018, p. 50) offer an invaluable resource in their guide, *Team Development: The Power of Debriefing* (see Figure 11.4). This guide lays out a pragmatic framework for debriefing, encompassing preparatory considerations, a detailed debriefing protocol, and strategies for post-debriefing follow-up, making it an essential tool for any team committed to continuous growth and performance excellence.

A strong body of evidence supports the conclusion that interventions designed for training teams are fundamental in enhancing both teamwork and overall performance. Such interventions begin with a needs assessment, followed by aligning training objectives with stakeholder expectations, and employing varied training methods for effective learning. Leadership training is central to team goal setting and outcome measurement, and team coordination and adaptation training emphasizes collaborative efficiency, shifting towards collective feedback to enhance team synchrony. CRM has been successfully extended to various industries, improving teamwork and performance through core teamwork principles. Team building activities are tailored to improve problem solving, goal setting, role clarification, and interpersonal relationships, and have been shown to be particularly effective for new teams. Lastly, team debriefing acts as a critical learning tool, significantly enhancing performance through structured, reflective practices that are documented for continuous improvement.

Case study: transitioning from individual to team training

A pilot project with a special operations forces team

This case study focuses on the implementation of a pilot project conducted with the Belgian Special Forces Group. The fundamental purpose of the program was to connect various divides: bridging the divide between physical and mental training; closing the gap between the curative/preventive medical approach and an approach focused on enhancing performance; and linking individual training with team training. The following summary will focus on the third objective, transitioning the operator's focus from individual effort to excellent teams, toward an identity as a team striving for excellence. For a full review of the development and implementation of this comprehensive Human Performance Program in a tier 1 Special Forces Unit see Pattyn et al. (2022).

Special Operations Forces (SOF) operators and elite athletes both demonstrate a wholehearted dedication to their respective professions. Historically, the process of selecting and training operators has been focused on cultivating their skills and abilities with the goal of achieving individual excellence. Therefore, transitioning from a solely individual-focused approach to performance management, while maintaining high personal standards, necessitated a change in mindset. Our human performance program aimed to take this additional stride.

To steer the program development process our team adopted the Intervention Mapping (IM) approach, as the benefits of this approach had been highlighted by Mattie et al. (2020) in their development of a human performance program for Canadian Special Forces. This method, delineated by Bartholomew Eldridge et al. (2016), involves six steps: needs assessment, setting objectives, designing the program using evidence-based methods, pilot testing, planning for sustainable implementation, and program evaluation. To this we added two additional aspects that we felt were essential – the incorporation of a multidisciplinary team approach that brings together all subject-matter experts in the human performance domain and a focus on co-creation with end-users, ensuring ethical and practical alignment.

This process was then broken into two stages, methods and results, and each of these was subdivided into two steps. The methods stage was broken into a needs assessment step and a program objective step. The first step of the results stage identified the program design contributions of each discipline including the specific assessment tools, training methods, and outcomes measures that would be utilized in the program, and the final step was implementation of the pilot project (see Table 11.1, Pattyn et al., 2022, p. 5).

Table 11.1 Overview of the four steps of the Intervention Mapping methodology

Methods	
STEP 1 Needs assessment	• Establish a multidisciplinary expert and stakeholders team to design the program. • Determine the current need based on real-life participant observation and analysis.
STEP 2 Determine program objectives	• Setting-up the program within a holistic approach regarding health and performance. • Define an individualized tailor-made approach to customize the whole support. • Address physical activity, nutrition and sleep needs to facilitate healthy lifestyle choice and performance improvement. • Support injury prevention and healthy coping mechanisms.
Results	
STEP 3 Program design according to each area of expertise	• Physiotherapy ◦ Identify body regions discomfort and potential musculoskeletal injuries through a first screening questionnaire. ◦ Provide an overall whole body functional movements assessment. ◦ Offer a detailed assessment for specific injuries involving lower back, cervical or lower/upper limb dysfunctions. • Physical training ◦ Define a detailed individualized physical performance assessment. ◦ Provide a specific, validated and practical test battery ◦ Create an evaluation tool to be used by the PTI, the operator and the physiotherapists. ◦ Provide individualized physical training programs. ◦ Adapt specific nutrition and hydration knowledge to the particular constraints of the population. • Performance psychology ◦ Determine the specific psychometry assessment need. ◦ Specify the most adequate validated trait and state assessment tools. ◦ Provide a customized individual feedback. ◦ Dispense a team workshop to provide feedback and determine possible interventions
STEP 4 Implementation in a pilot project	• Conceive a modular training program about the impact of human factors on the individual and team functioning. • Integrate an evolution from individual functioning to team functioninig; and from participant operator receiving expert advice to autonomous actor of their own performance management. • Distribute the program throughout the year, according to the modular built-up principles discussed earlier: ◦ Four weeks at the unit (January – April – June – December) ◦ Two deployment periods (3 weeks/3 months) with embedded experts.

Adapted from the original (Patton et al., 2022).

Step 1: Needs assessment

To ensure a comprehensive focus on all aspects involved in human performance, professionals from across the fields of clinical medicine, physiotherapy, physical training, nutrition, and performance psychology were brought together to create the program design team. We capitalized on our psychologist's specialization working therapeutically with networks of individuals, resulting in our approach being rooted in psychotherapeutic systems thinking.

To initiate our project, we arranged multiple coordination sessions where all collaborators prioritized their involvement despite other commitments. The first critical step involved each specialist gaining an in-depth understanding of others' fields, recognizing the distinct contributions of each discipline to the program. Specialists demonstrated how their expertise enhanced human performance, detailing their assessment techniques, feedback processes, and intervention strategies. This exchange aimed to achieve two primary objectives: firstly, to develop a comprehensive grasp of the program as a whole, identifying synergies and effective strategies for guiding operators in each domain; and secondly, to showcase the concrete impact of each expert's involvement, underscoring the collaborative essence of the program. We acknowledged that while expanding the scope of expertise beyond their individual interventions might initially appear demanding for our specialists, it is essential to recognize that performance is a holistic concept. Addressing all the "pillars of health" is fundamental to optimizing client performance effectively.

Step 2: Determine program objectives

As the primary approach for the program development process, we adopted a non-hierarchical co-creation approach, actively involving both specialists and the client. This process included unit representatives, team leaders, and key stakeholders in collaborative discussions to foster program relevance, acceptance, and long-term sustainability (Gergen, 2008; Jorgensen, 1989; McTaggart, 1991; McIntyre, 2007; Spradley, 2016). An important aspect of this stage was the deliberate shift of focus from individual clients to the team as the primary beneficiary in aligning with the program's ultimate objectives.

The involvement of clients and stakeholders is crucial in this context. Their buy-in is essential for several reasons. First, it ensures that the program is responsive to the actual needs and expectations of both the individuals and the team as a whole. When clients and stakeholders are involved in the program's design, they are more likely to feel a sense of ownership and commitment to its success. This sense of ownership is vital for the program's acceptance and longevity. Additionally, stakeholders provide valuable insights into the broader organizational context and goals, ensuring that the program aligns with strategic objectives. Their support is also crucial in providing necessary resources, facilitating implementation, and championing the program within the organization. Gaining buy-in and participation from clients and stakeholders as a component of developing the human performance program increases the likelihood of successful implementation, sustained engagement, and the achievement of desired outcomes (Mattie et al., 2020; Pattyn et al., 2022).

Step 3: Program design

Psychological assessments were to first serve as the basis for developing greater individual awareness and then would provide insights into interpersonal styles as they are shared with the team. This process was broken into a three-tiered approach. In the first-tier individual assessments were conducted and the results were mapped using psychometric measures. The second tier focused on providing personalized feedback to each individual. Finally, the third tier involved integrating the assessment results at the team level, which included a psychoeducational approach aimed at fostering self-regulation within the team.

In the study of human performance, the definition or specification of what one should assess is the most essential challenge, whether for individual or crew performance, as Hollnagel (1998) described more than two decades ago. Measurements must be possible, trustworthy, relevant, and valid. Few practical measurements meet all of these conditions, so this continues to be a challenge for those looking to assess team performance. Additionally, the quality of the expert who provides comments and uses the information is often disregarded when analyzing assessment outcomes. A trained systemic psychotherapist and a performance psychologist individually interviewed each participant, followed by a joint feedback session with both professionals. This interdisciplinary collaboration, though uncommon in psychology, was deemed essential based on our previously established criteria.

After consideration of various psychological instruments, including those previously completed by the operators, the psychometric tools chosen for effective team feedback and interventions included the NEO-PI-R, EQi, and the MBTI. The Revised NEO Personality Inventory (NEO-PI-R), was central to this methodology, offering an in-depth analysis of five major personality traits and facets underpinning these traits (Young & Schinka, 2001). To measure emotional intelligence we used the Bar-On Emotional Quotient Inventory (EQi), a tool assessing emotional intelligence across a range of competencies including intrapersonal, interpersonal, stress management, adaptability, and general mood (Bar-On, 1997, 2004). Alongside these, the Myers's-Briggs Type Inventory (MBTI; Boyle, 1995) was incorporated, despite wide skepticism of the its usefulness among psychologists, typified in the theoretical criticism of the MBTI outlined by Stein and Swan (2019). The researchers selected the MBTI primarily because it was familiar to the operators from a previous international training, and in part to leverage its popularity to promote self-awareness and bridge communication gaps between different assessment dimensions.

Considering the team's small and egalitarian structure, a unique "third person" assessment technique was implemented. This involved team members completing the MBTI for each other, aiming to explore the differences between self-perception and how others perceive them. This strategy proved crucial in examining social desirability and authenticity, key elements in maintaining effective team dynamics and coping mechanisms.

The concept of social desirability in assessments, often viewed as a bias, was addressed through the dual categories of self-deception and impression management, as defined by Paulhus (1984). While impression management involves a conscious effort to project a positive image, self-deception is characterized by respondents believing their positive self-reports. Understanding the importance of these aspects in various contexts, including military settings, the researcher included the Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984, 1991) in the evaluation process. This measure aimed to raise participant awareness of their behavioural tendencies.

Step 4: Implementation blueprint

The implementation of the pilot program was designed to be structured in a modular manner, consisting of six distinct training times over the course of a year. The initial training was scheduled for the only time in the year when the entire team operates together in a standard work environment, commonly known as the "administration and logistics" weeks. We selected this time to start the program, ensuring the availability of all team members for the training. There were also two distinct training programs during deployments that were spread over the one-year period. The first deployment training involved a three-week mountain training session, during which the physical training instructor served as the embedded expert. The second period encompassed a three-month operational deployment, during which the medical doctor assumed the role of the embedded expert for a duration of one month. During these trainings the focus began to shift in emphasis from the individual to the collective entity, specifically the team. Table 11.2 presents a comprehensive summary of the team members' timetable and organization, excluding any consideration of the preliminary work and collaboration among the specialists (Pattyn et al., 2022, p.19). Over the course of the year, the process transitioned from a model in which experts provided guidance to a collaborative approach involving the team, which was informed by emerging insights and accumulated experience.

Table 11.2 Blueprint

Administration and logistics weeks			
BLOCK 1	1 week (Jan)	Education (team classroom sessions)	1. introduction to the program and goalsetting (1 h) 2. exercise physiology and training principles (4 hrs) 3. information processing and learning processes (4 hrs)
		Individual assessment	1. initial medical interview 2. psychometry tools 3. individual intake interview with the clinical psychologist 4. VO2 max testing at the sports physiology laboratory 5. individual physiotherapy screening: questionnaire and consultations
		Team Intervention	first team training session with PTI to illustrate training principles (half day)
BLOCK 2	1 week (Apr)	Education (team classroom sessions)	1. nutrition basics (4 hrs) 2. communication and team cognition (2 hrs) 3. sleep and fatigue management for optimal performance (2 hrs)
		Individual assessments	1. full physical assessment with PTI 2. repeat sleep aspects of psychometry
		Individual Intervention/ Feedback	individual consultation with physiotherapist and PTI to discuss customized training program based on the assessments of block 1.
		Workshop/Practical exercise	1. nutrition: analysis of the different types of field rations used by the unit 2. exercise on determination of metabolic needs in function of different types of settings and activities (based on real exercises/deployments) 3. sleep and fatigue: scheduling examples based on observational data from the mission of the previous year
		Feedback	individual interview with the psychologists regarding the psychometry results from block 1.
		Team intervention	1. introducing the concept of team performance management and the team assessments 2. group workshops around personality types, behavioural preferences, and team dynamics
BLOCK 3	1 week (Jun)	Individual intervention	1. individual follow-up with physiotherapist and PTI on customized training program 2. individual pre-deployment interview with clinical psychologist
		Team intervention/ workshop	1. how to implement the Human Performance Program on deployment. 2. team cognition, performance and human error: how to reframe error analyses (with real-cases examples). 3. team training session with PTI

(*Continued*)

Table 11.2 (Continued)

Administration and logistics weeks			
BLOCK 4	1 week (Dec)	Individual assessment	repeat of the full physical assessment to evaluate impact of deployment
		Individual intervention	follow-up with physiotherapist and PTI on customized training program follow-up consultation with ad hoc experts based on individual needs
		Team intervention/ workshop:	1. debriefing on human performance aspects on deployment: physical activity, nutrition, sleep and fatigue 2. education refresher regarding nutrition and sleep (2 x 2 hrs) based on feedback during deployment
Deployment periods			
Mountain training period	3 weeks (Feb)	• Mixed education/intervention with PTI: • Injury prevention and recovery applied to a technical and tactical setting. • Physical activity as a means (technical), an end (tactical), and a recovery resource (mountaineering activity during the free week-end). • Emphasis on the importance of managing physiological resource spending and acceptable pain thresholds depending on the context. • Illustration of nutrition choices depending on the type of activity.	
Operational deployment	3 months (Aug-Nov)	• Interventions: • Managing nutrition in a resource-constrained environment, based on the previous lectures and workshops; • Adapting sustained operations schedule to the team set-up in terms of chronotype and sleep need; • Individual physical training schedules depending on available time and space. • Availability of the experts (PTI, MD, Physiotherapists, Psychologists) for reach back guidance and support.	

To facilitate the paradigm shift our team workshops drew an initial comparison between Chris Hadfield's memoir detailing his experiences as an astronaut (Hadfield, 2015) and the professional trajectory of an operator. In his narrative, Hadfield provides a lucid account of the transformation in his cognitive framework, wherein he transitions from a competitive and individualistic fighter pilot to a member of a space crew. This shift in perspective leads him to see that his survival prospects in the space environment are contingent upon the collective efficacy of the crew, rather than solely relying on his own capabilities. Hadfield employs the terms "how to be a zero" to depict this transformation in mindset: his transition from striving to be the standout individual in any given system, to endeavouring to function as a highly adaptable and efficient component within a significantly intricate apparatus.

The program built on the insights gained in the individual psychological feedback sessions by facilitating discussion of the information collectively with team members sharing with each other their strengths and weaknesses. In this process the teams were able to begin developing mental maps of how to support each other and utilize each other's strengths. The conclusion of the program involved an interactive classroom activity in which we critically analyzed authentic instances of performance evaluations and human errors from past courses and deployments, drawing upon our first-hand participant observations from embedded experts. This encompassed evaluations of performance in physical or tactical difficulties within training programs, as well as insights derived from post-deployment after-action reports. This phase facilitated the demonstration that individuals who may have appeared to be “high performers” were, in fact, reliant on the system (team, unit) in which they operated to achieve genuine high performance. Further, the team’s active engagement in arriving at this understanding demonstrated their ability to apply the concepts covered in the activities beyond our human performance program to real-world scenarios, an essential step for demonstrating the success and effectiveness of the workshop.

Results achieved

The primary positive outcome for the team entailed the enhanced quality of the team dynamics, insights, and self-awareness, thereby instilling a sense of fortitude through the identification of distinct strengths and flaws, as well as the strategic means to rectify them. The successful application of the selected knowledge and behavioural skills was demonstrated through the effective transfer of information across individual interventions, team workshops, and classroom sessions to real-life settings. Nevertheless, the participants also expressed a significant degree of dissatisfaction over the perceived “culture clash” within the unit, which presented challenges in effectively implementing their acquired knowledge. This underscores the necessity for the successful implementation of such a program to encompass all levels of end users and management in decision-making and program development. Given the scale of an organization such as Defence, certain decision-making procedures cannot be effectively decentralized and efforts such as this must be viewed as an iterative process. This type of program could perhaps be facilitated within the context of a sports team setting, with a higher autonomy, or in a more decentralized organization.

Case study: trouble with boundary spanning

The Officer in Charge (OIC) of an Intelligence Department at a military unit sought help to address problems with burnout that were plaguing his team. Military intelligence is a specialized branch of the armed forces focused on the collection, analysis, and dissemination of information that is vital for military decision-making. They play a critical role in all military operations. Personnel in military intelligence are tasked with gathering data through various means, analyzing this information, and then providing the analysis to the command as part of their planning processes. Their work requires that they collaborate with other teams, the command group, and external agencies to provide a comprehensive understanding of potential threats and security issues related to the mission.

To get an understanding of what was going on in this intelligence team we began with the administration of the 30-item Team Process Survey (Mathieu et al., 2019) and included four open-ended questions. Specifically, which elements of your team’s performance are key to your current success? What areas of teamwork does your team need to improve to optimize performance? What type of resources does your team need to perform better? And, is there another question/issue that should be asked about how this team works together? The survey is on a five-point Likert scale and results under 3 were considered a weakness while above 4 were considered a strength. The results are indicated below (see Table 11.3 below). The outcomes from the open-ended questions were analyzed to determine the team dynamics they represented. The findings aligned with the survey results and offered insights directly from the team members, expressed in their own words.

As is evident, the team was having significant difficulty identifying their priorities. Material from the open-ended questions suggested that they specifically had a hard time prioritizing which actions were most important, as they felt that it was their responsibility to respond to all requests that came to the department. They did not feel empowered to say no to any unit member who requested their support. This resulted in an unmanageable workload that led to all the team members putting in extreme hours. They provided good backup behaviour to each other in their effort to cover everything and showed good ability to manage their relationships with each other, but the stress was clearly evident.

Table 11.3 Intel "Team Process Survey" results

	Transition Processes: 3.16	
	Mission Analysis:	3.36
1	Identify our main tasks?	3.36
2	Identify the key challenges that we expect to face?	3.45
3	Determine the resources that we need to be successful?	3.55
	Goal specification:	2.79
6	Set goal for the team?	2.55
7	Ensure that everyone on our team clearly understand our goals?	2.82
8	Link our goal with the strategic direction of the organization?	3.0
	Strategy Formulation and Planning:	3.24
9	Develop an overall strategy to guide our team activities?	2.73
10	Prepare contingency ('if then') plans to deal with uncertain situations?	3.73
11	Know when to stick with a given working plan, and when to adopt a different one?	3.27
	Action Processes: 3.16	
	Monitoring Progress Toward Goals:	3.22
12	Regularly monitor how well we are meeting our team goals?	2.36
13	Use clearly defined metrics to assess our progress?	2.18
14	Seek timely feedback form stakeholders (e.g. customers, top, management, other organizational units) about how well we are meeting our goals?	2.91
	Systems Monitoring:	3.36
15	Monitor and manage our resources (e.g. equipment, manpower, etc. …)?	3.18
16	Monitor important aspects of our work environment (e.g. inventories, equipment and process operations, information flows)?	3.27
17	Monitor events and conditions outside the team that influence our operations?	3.64
	Team Monitoring and Backup:	3.82
18	Develop standards for acceptable team member performance?	4.00
19	Balance the workload among our team members?	3.18
20	Assist each other when help is needed?	4.27
	Coordination:	3.48
21	Communicate well with each other?	3.36
22	Smoothly integrate our work efforts?	3.64
23	Coordinate our activities with one another?	3.45
	Interpersonal Processes: 3.62	
	Conflict Management:	3.97
24	Deal with personal conflicts in fair and equitable ways?	3.82
25	Show respect for one another?	4.45
26	Maintain group harmony?	3.64
	Motivating and Confidence Building:	3.67
29	Take pride in our accomplishments?	3.18
30	Develop confidence in our team's ability to perform well?	3.91
31	Encourage each other to perform our very best?	3.91
	Affect Management:	3.21
32	Share a sense of togetherness and cohesion?	3.73
33	Manage stress?	2.73
34	Keep a good emotional balance in the team?	3.18

A one-day workshop was developed to address these concerns. In the workshop the team was broken into work groups to discuss how to communicate with other teams and clients when their requests are not aligned with the command's priorities and would draw away resources from primary tasks. Each work group gave a report of their discussion, and then the team came to a shared understanding of how they will collectively communicate with the command group, other teams, and unit members. They agreed to a common phrasing that they would use to set a clear boundary that would help others recognize when what they were asking for was not focused on the units strategic priorities. As a result, the team was able to feel empowered to address their situation which improved morale and reduced their perceived stress level.

In this case summary the Team Process Survey and four open-ended questions were used to gather data on the difficulties faced by a military Intelligence Department facing intense stress. The results were used to develop a one-day workshop in which the team developed their own way of addressing their situation which empowered the team, boosting morale and reducing stress levels.

Case study: developing high-performing flight crews

Traditionally, the performance of a team is associated with taskwork and the products of taskwork in the sense of how well the team achieves the objectives of its tasks. However, when trying to comprehensively understand team performance, it is appropriate to look at more than simply the outcome of the team's task. Focusing only on output performance will provide a narrow view about the performance of team members and the team as a whole. This approach potentially misses the complexity of the task environment, behaviour of the team members during the task, as well as team processes and emergent states leading to the output.

In air combat, teams operate in a constantly changing, probabilistic environment and decisions are often made with incomplete information. As a result, a team might achieve the desired result in almost any decision-making activity even with incorrect decisions. On the other hand, an undesirable outcome can be reached after making all the right decisions. In both cases, the team members could have good or poor perceptions of each other's actions and knowledge about the environment when conducting the task; their cognitive resources could be completely depleted, or their mental workload may be low. Considerations of this type, some of which are also related to teamwork (i.e., how the team does a task) must be taken into account in order to draw a holistic view of team performance. This following case study discusses how to obtain a comprehensive estimate for team performance by using task performance, normative performance, team situation awareness, and mental workload as supplementary measures for the team's output performance. These metrics are illustrated with examples, and their measurement practices are introduced. This section also considers effects of explicit coordination on teamwork, i.e., the communication of team members, as well as implicit coordination based on their shared knowledge, or team situation awareness. Affective concepts, such as cooperation and psychological safety, as well as some concepts related to teamwork, such as backup behaviour and mutual performance monitoring, are not addressed. The measures discussed in this section are summarized in Figure 11.5.

Output performance (OP)

The most traditional way of assessing team performance is to measure a team's output performance (OP) after its action period. A sales department's OP can be evaluated by reviewing the sales at the end of the quarter, or the output performance of a football team can be assessed by simply looking at the scoreboard at the end of the second half. OP is the most used measure in air combat when evaluating the performance of a flight. A flight is the standard team used in air combat. It comprises

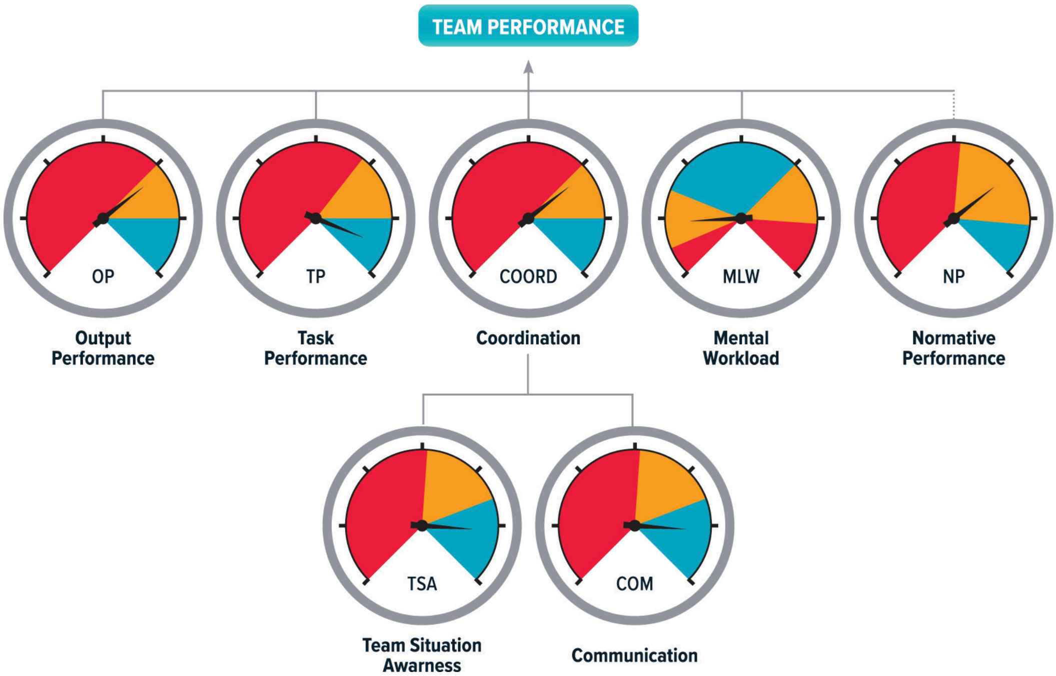

Figure 11.5 Air Combat Team Performance. Summary of measures used to evaluate team performance. The measures include output performance (OP), task performance (TP), coordination (COORD), team situation awareness (TSA), communication (COM), mental workload (MWL) and normative performance (NP).

four aircraft arranged in two sections with two aircraft (lead and wing) in each. The most typical OP measure used is the ratio between the number of enemy aircraft shot down to the number of friendly aircraft lost. While it is the measure of performance most directly related to the objectives of the mission, it only provides a relatively narrow insight concerning the team's performance and it potentially misses the complexity of air combat. It does not reflect the flight's taskwork using tactical operating procedures, the team's competency and the applicability of the aircraft systems utilized (Mansikka et al., 2021a). One way to gain a better understanding about a team's performance is to evaluate its task performance.

Task performance (TP)

Task performance (TP) can be best explained with an example. Imagine that the national football team of England plays two practice games against France. The manager of England decides to play the first game using one set of tactics and the second game using another, whereas France uses the same tactics in both games. From all other aspects, both teams are assumed to perform similarly in both games. The final score in both games is 0-0. Does this mean that both the English national team tactics were equally good in both matches? Not necessarily, because the final score itself does not say much about what happened during the game. It could be that in the first game England was constantly just inches away from scoring a goal, whereas France may have dominated the second game but were just unlucky. To gain an insight about team performance and the dynamics of the game, we need other indicators in addition to the OP.

One way to do this is to measure the team's TP. When TP is measured, the team's task must be divided into recognizable and measurable phases or stages. Using the same football example again, TP could be described as a hierarchical construct where during the first phase the English team must gain possession of the ball. In the next phase, the team must move the ball into France's defensive zone. Then, in the third phase, the English team must pass the ball and position the players such that the French defensive line is broken and an attempt on goal is made. Finally, if the English team successfully completes the fourth phase and a goal is scored, this is reflected in both teams' OP.

Any TP measurement should start by identifying the hierarchical phases or stages impacting the OP of the team. There is no global way to describe the progress of teams' taskwork. In air combat, the flight's taskwork can be described as the management of two parallel processes called the live-chain and kill-chain (Joint Chief of Staff 2013). On one hand, the flight tries to shoot down the enemy fighters, but on the other hand it tries to deny the enemy from killing members of the friendly flight. Both friendly and enemy pilots can engage a number of opposing aircraft at the same time. Every friendly pilot has a kill-chain against each enemy aircraft. The management of a kill-chain can be assigned to a single pilot, or the management of a single kill-chain can be a joint effort between several flight members. However, at the same time, each pilot has only one live-chain.

TP describes how close or how far the team is from achieving its overall goal, such as destroying the enemy aircraft while staying alive. In air combat, the kill-chain can be divided into phases such as find, target, engage, and assess. During the find phase, the flight attempts to find the enemy aircraft with its radar. Once found, the enemy aircraft is targeted. This means that the follow-on management of the kill-chain is assigned to a flight member or a group of flight members. The responsible flight member (or members) manoeuvres into a position where they can launch weapons against the targeted enemy aircraft. As the launch parameters are satisfied, a weapon is launched and the engage phase begins. The engage phase continues until the weapon hits or misses the enemy aircraft. At this point, the flight's kill-chain progresses to the assess phase. During the assess phase, the flight evaluates whether a satisfactory effect on the enemy has been achieved or does its kill-chain need further management? It is important to note that TP is a hierarchical construct, where the preceding phase must be satisfied for the following phase to be reached. As such, the phases associated with TP can move dynamically back and forth. For example, in air combat, the friendly team can lose track of the enemy aircraft in the middle of the target phase. If this happens, the kill-chain against that enemy aircraft reverts to the find phase.

In air combat, the live-chain is essentially about denying the opposing team's kill-chain from progressing. Due to the equilibrium of the chains, the live-chain has phases which mirror the phases of the kill-chain. The phases of the live-chain can be described as deny-find, deny-target, deny-engage, and deny-assess. Each time the phase of the kill-chain changes, it also changes the opposing side's live-chain. From the perspective of TP, any force-on-force setup, such as air combat, is a zero-sum game: both teams have similar kill- and live-chains, where a gain in a one team's kill-chain results in an equal loss in the other team's live-chain and vice versa.

As noted earlier, teams with different pilots, aircraft systems, or tactics might achieve equivalent OP but this does not mean their TPs are similar. If TPs are dissimilar but OPs are similar, the flight who has kept its members' live-chains more intact will have maintained a greater survival margin. In the same vein, a flight whose kill-chains have progressed further will have been closer to achieving the desired effect on the enemy. In summary, it is more informative to evaluate TP based on the progression of the kill- and live-chains than just OP alone (Mansikka et al., 2021a).

When the phases in the TP measurement involve cognitive characteristics, such as situation awareness, it is essential to differentiate the machine phases from the human-machine phases. For example, an aircraft's sensor may have detected an enemy aircraft, but the pilot may have allocated his/her attention elsewhere and thus be completely unaware of the changed machine phase. If the phase changes can be unambiguously observed from the system with which the team members interact, the measurement of the machine TP can be evaluated in real time by simply logging the system status or events associated with the chain changes. However, the measurement of the human-machine TP can be laborious and time-consuming. In air combat, the task situation can be highly dynamic and chaotic as the multiple kill-chains and the live-chains of individual pilots move constantly back and forth. Real-time evaluation of the phases of the kill- and live-chains can seriously disrupt the execution of the primary task. It is therefore essential that the human-machine TP is evaluated after the task. To do this the flight members must be able to recall what happened during the task execution and to verbalize how they understood the kill- and live-chains to have evolved. To avoid the TP measurement from turning into a memory recall task, it is recommended to replicate or reconstruct the progression of the tasks and thereby assist the flight members in recalling how the tasks evolved from the perspective of TP. With relatively little practice, trained pilots can usually provide highly accurate assessment of their TP phases.

Compared to OP, TP provides a far more diagnostic measure of team performance. With TP, it is possible to differentiate the competencies of teams, the procedures they follow, and the effectiveness of the tools and systems they use, even when these differences cannot be identified from OP. In addition, TP can reveal if the team performance was closer to being good or closer to being bad as evaluated by the team's OP.

Normative performance (NP)

Normative performance (NP) describes the level of adherence to the team's tactics, techniques, and procedures (TTPs), i.e., how accurately they are followed during a team's task execution (Mansikka et al., 2021d). While NP is not a measure of team performance per se, it plays a critical role when the utility of TTPs, the competence of teams, or the applicability of a team's tools and systems are evaluated and compared. NP also considers the impact that non-adherence of TTPs has on task accomplishment, i.e., OP.

While the concept of NP was originally developed for the assessment of air combat TTPs, the principles of NP can be extended beyond air combat. For example, let us consider the previous football example again. As discussed, the head coach of England tried two different tactics. But let us also assume that the head coach was stuck in traffic on the M25 motorway around London and could not see the game. Could s/he still draw conclusions about the effectiveness of those two tactics based on OP and TP alone? Probably not. This is because without the knowledge of the team's NP, the head coach would not know whether the team had followed the tactics they were supposed to follow. As this example highlights, the concept of NP can be used in many different domains. However, for the sake of clarity, the rest of the section deals only with the adherence of air combat TTPs.

For NP measurement to be possible, the directed TTP must be documented in sufficient detail, and it must be possible to identify possible TTP adherence violations during or after its execution. The NP measurement starts by identifying the core tasks a flight must undertake to be successful in air combat. Depending on the depth and breadth of the analysis, a large number of tasks may be identified.

Almost every task in air combat is somehow regulated and could thus be included in the assessment of NP. However, as it would be impractical to assess the team's adherence to every possible task it in which it engages, the number of tasks evaluated with NP measures must be reduced to a manageable level (Mansikka et al., 2021b). Subject matter experts (SMEs) can assist in identifying the most relevant tasks. SMEs can also help in reducing the number of nominated tasks by combining and grouping them into meaningful units and removing possible duplicates. This process may require several iterations. The selection of the most relevant tasks can be conducted by rating the tasks based on their impact on the flight's OP. A decision-tree format based on the Cooper-Harper scale (Cooper & Harper 1969) can be useful when doing this – especially when each rating in the decision tree is associated with a verbal description. Figure 11.6 (adapted from Mansikka et al., 2021b) illustrates a Cooper-Harper type tool for rating the tasks based on their importance for the flight's air combat mission. In this figure, the importance of the tasks ranges from 1 (low importance) to 5 (high importance). Based on the resulting ratings, the tasks can be shortlisted such that the NP measurement captures a desired number of tasks with known importance to the flight's OP.

NP is measured by comparing the flight's task execution with that described for each sub-task. If the automatic measurement of NP is not possible (as is often the case) the NP measurement can best be conducted post-task. Post-task NP measurement requires that the flight's task execution can be recorded or otherwise tracked and reviewed. It can also be done by using observer ratings or the flight members can conduct the NP measurement by themselves. Usually SME observers

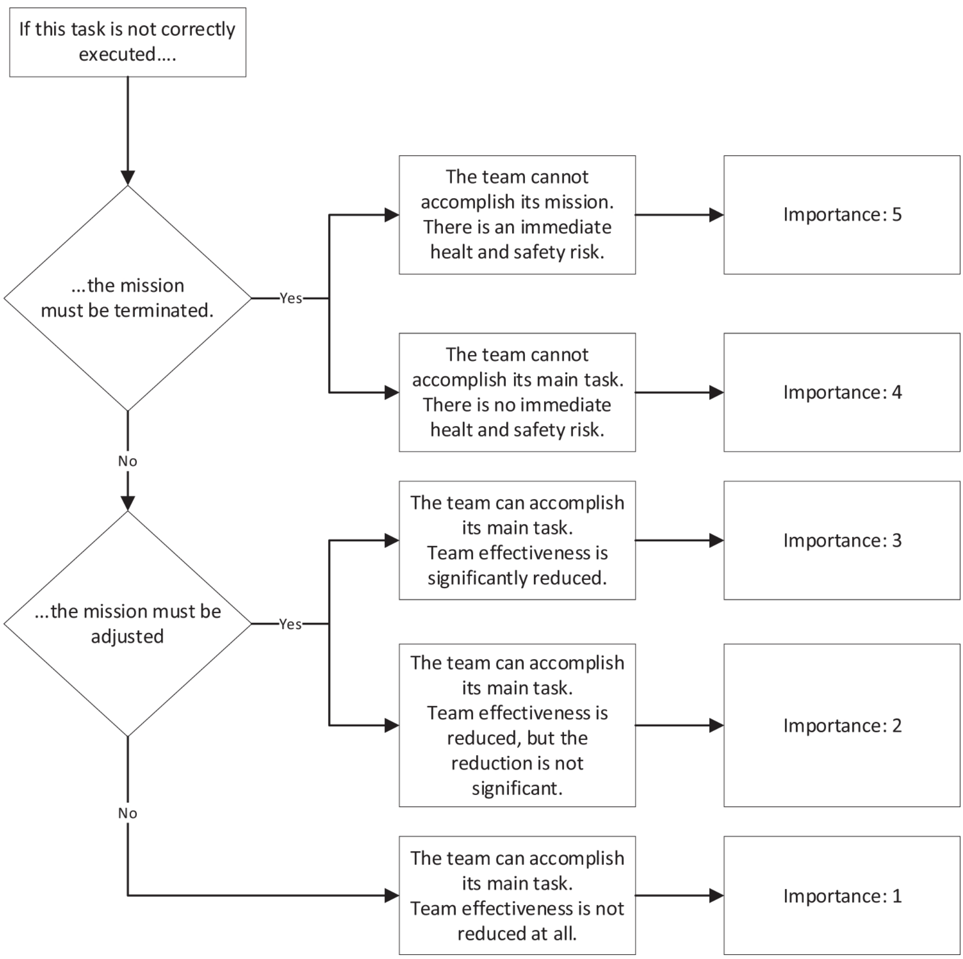

Figure 11.6 Cooper-Harper type tool for rating the tasks based on their importance for the flight's air combat mission (adapted from Mansikka, H., Virtanen, K., Mäkinen, L. and Harris, D. (2021b). Normative Performance Measurement in Simulated Air Combat. Aerospace Medicine and Human Performance, 92(11), 908-912).

are preferred, as the measurer must be capable of evaluating how the degree of adherence or non-adherence affects the flight's overall task accomplishment.

To enable scoring of NP, questions tapping the execution correctness of each sub-task must be generated. For example, an adherence question evaluating an aircrew's engine start-up procedure could read as: "Did the crew follow the established standard operating procedure for engine start-up?" As the team's NP is reviewed, each time a nominated sub-task is performed, an associated adherence question is introduced and the NP regarding that sub-task is scored (Mansikka et al., 2021b). The NP score is based on the level of process or procedure adherence and the impact that a possible nonadherence has on the overall task accomplishment. While there is no global scale for NP, a scale ranging from 0 (low NP) to 3 (high NP) has proven to be practical (Mansikka et al., 2021b). Within the scale, each score is associated to a verbal description as follows:

- 0 = Did not adhere to an established process or procedure. The non-adherence had a negative impact on the overall task accomplishment. The negative impact was significant;
- 1 = Did not adhere to an established process or procedure. The non-adherence had a negative impact on the overall task accomplishment. The impact was not significant;
- 2 = Did not adhere to the process or procedure. The non-adherence had no impact on the overall task accomplishment;
- 3 = Did adhere to the process or procedure.

Once all the shortlisted sub-tasks have been assigned a NP score, an overall NP score for each flight member can be formed by averaging the individual scores. The level of an acceptable NP must be determined on a case-by-case basis.

To summarize, whenever a new work process, procedure or tool is introduced, one should always make sure that the new workflow is being followed and executed accordingly before drawing any conclusions concerning whether it is effective or not. Unless satisfactory NP is verified, all other team performance measures may have little or no meaning as it could be that the team never did what it was supposed to do.

Mental workload (MWL)

Mental workload (MWL) characterizes the demands imposed by the tasks on limited mental resources when the desired performance is to be maintained (Vicente et al., 1987; Parasuraman et al., 2008). From the perspective of performance, too low or too high workload is not desirable and the optimal level of cognitive resource expenditure lies somewhere between the two extremes.

The available cognitive resources define the portion of such resources to achieve a desired level of performance. The variations of task demand cause variations in the amount of resources required to satisfy that demand. From the perspective of an individual, the task demand can be managed by, for example, either lowering or increasing the personal desired level of performance, i.e., varying the amount of effort invested on the task, and/or by changing the strategy used for executing the task. However, in air combat and many other safety critical tasks, the satisfactory level of performance cannot be lowered and if the task demand increases, the predefined level of performance must be maintained by investing more resources and greater voluntary effort. Unfortunately, once the pilot's cognitive resources have been depleted, performance will begin to deteriorate regardless of the amount of effort invested. In many dynamic tasks, working memory is often the most performance-limiting cognitive resource.

An unbalanced MWL has a negative impact on both individual and team performance. A team has several ways to manage its members' MWL. First, a team may try to organize and allocate its tasks such that the team members are not exposed to an undesirable level of task demands. Second, a team may attempt to lower the MWL experienced by team members by changing the teams' operating strategy. Third, assuming that overload is a problem, the team could release team members with too low cognitive capacity and replace them with individuals with higher capacity. For a team to be aware whether MWL is an issue, the MWL of team members must be measured.

MWL cannot be measured directly. Most empirical measures can be categorized either as behavioural (or performance based), subjective, or physiological. Behavioural measures can be broadly divided into primary and secondary task measures (Paas et al., 2003). When secondary task measures are used, an operator is given both a primary task and a secondary task. The operator is expected to maintain his/her performance on the primary task and to use any excess cognitive capacity on the secondary task (Casner & Gore, 2010). Variations in the secondary task performance are assumed to reflect the amount of spare capacity – and hence the MWL caused by the primary task (Verwey & Veltman, 1996; Ogden et al., 1979; Brown, 1978). The secondary task measures are, by their nature, disruptive and cannot be safely utilized in high-risk environments when the MWL is high (O'Donnell et al., 1986; Casali & Wierwille, 1984). The primary task measures are based on the assumption that the pilot's primary performance is related to MWL. The sensitivity of the primary task measures decreases as MWL moves towards the high or low extremes. All secondary task-based measures are unique to the scenario and the secondary task used. There is no MWL metric that can be generalized from such measures.

Physiological measures of MWL are based on the physiological changes caused by arousal, effort, and general activation level. As different tasks cause different physiological responses, not all physiological measures are sensitive to MWL in all tasks. A wide range of measures tapping physiological changes in either the central nervous system or the peripheral nervous system have been developed. These include measures such as electroencephalographic (EEG) activity (Berka et al., 2007), event-related brain potentials (ERPs) (Kramer et al., 1987), electrooculographic (EOG) standing potentials (Ryu & Myung, 2005), functional near-infrared spectroscopy (fNIRS) (Hamann & Carstengerdes, 2022), heart rate (HR) (Reimer & Mehler, 2011), heart rate variability (HRV) (Mehler et al., 2011), electrodermal activity (EDA) (Setz et al., 2010), and electromyography (EMG) (Roman-Liu et al., 2013).

HR and HRV are probably the most widely used physiological measures of MWL. HR measures the heart's beat-to-beat sinus rhythm or interval, whereas HRV measures the differences between these intervals (Sampson & McGrath, 2015; Houghton & Gray, 2014; Ivanov et al., 1999). While physiological measures allow objective, continuous, and real-time monitoring of the pilot's state, which does not intrude on performance, their sensitivity may become limited when used during a flight. For example, pupillary diameter may be affected not only by variations in the information processing demands, but also by variations in the eye's fixation distance or ambient lightning. In a similar fashion, cardiological responses can be affected by blood pressure variations, body temperature, and arterial pressure. In fighter aviation, factors like extreme cockpit temperatures, exposure to direct sunlight and high G-loads can generate physiological responses which can, if not properly controlled, be falsely interpreted as MWL responses. In addition, the instrumentation used for physiological MWL measuring is often bulky, disruptive, and requires expensive aircraft integration and flight approval. Physiological responses are also unique to the individual and once again, there is no generalizable metric of MWL. All MWL measures are relative to the various task conditions. While HR and HRV have been successfully used to measure MWL in a simulator environment (Mansikka et al., 2019a; 2016a; 2016b), recent developments in smart rings and other wearable technologies may open new possibilities for physiological MWL measurement during a flight as well (see Kinnunen et al., 2020; Kinnunen & Koskimäki, 2018; Stone et al., 2021).

Subjective measures of MWL, on the other hand, are easy to employ in simulated and real fighter missions, and they generally enjoy high face validity with wide operator acceptance. However, subjective measures have been criticized for their inherent tendency to generate time-related errors as the data collection is typically conducted after the activity. In addition, subjective measures typically rely on a MWL rating scale of some sort. This mechanism is a source of bias as the subjects must memorize their past sensations and arrange them to a rating scale (Annett, 2002). In addition, the number of different task features and the phasing of high and low task demand events can affect subjectively sensed MWL (Wierwille et al., 1985). Subjective, post-task measures have been criticized as it is often unclear if the rating reflects average workload for the task, peak workload during the task, or is a measure of the user's experience of MWL during the latter stages of the task. Furthermore, the potentially negative outcomes of unbalanced workload may limit the willingness to report MWL honestly (Mansikka et al., 2019a).

NASA-TLX (Task Load Index) is the most widely used subjective MWL measurement method. It is a multidimensional method, where MWL is assessed across six dimensions. Despite its popularity, it has some serious challenges. For example, the original NASA-TLX weighting technique based on pairwise comparisons of MWL dimensions' importance does not allow directly expressing two or more dimensions as equally important. Also, if pairwise comparisons are conducted consistently, there exists only one possible importance order for the dimensions. Finally, with consistently conducted pairwise comparisons, a weight of 0.33 is artificially forced on the most important dimension and the least important dimension is given a zero weight. Thus, in practice, the contribution of one of the dimensions is not considered at all. These challenges can be overcome by determining weights of load dimensions with the Swing weighting method (see Virtanen et al., 2021 for details of how to apply Swing weighting with NASA-TLX).

In summary, both subjective and physiological MWL measures may be effective and sensitive in certain situations but highly ineffective or unreliable in others. While Mansikka et al. (2019a) have demonstrated HR/HRV and subjective measures to be equally sensitive when assessing flight members MWL in an air combat environment, the final selection of the type of MWL measure(s) should be assessed on a case-by-case basis.

In flight, no obtrusive measures or ones that require divided attention on the part of the pilot should be used. These may be acceptable in a simulator, but care needs to be taken that they do not impinge on the flight task being evaluated. In general, most physiological measures associated with the collection of ECG, respiration rate, or blood pressure can be used at any time. However, there can often be problems associated with locating and securing the equipment safely in the aircraft itself, especially in smaller types. In general, EEG, EOG, and electrodermal responses are impractical to collect in flight. To help improve reliability of MWL measures, particularly intra-rater reliability, consideration should be given to performing a series of workload ratings on a number of reference tasks (of varying levels of workload) to help "calibrate" the pilots, prior to making the workload assessments on the tasks of interest. This also helps to avoid "errors of severity" or "errors of leniency" when making assessments.

(Team) situation awareness (T)SA, communication (COMM) and coordination

Various definitions of situation awareness (SA) have been proposed. All relate to the acquisition of task specific information for integration into a dynamic mental model to support a pilot's dynamic decision-making processes. SA is not "achieved"; it is constantly being revised and updated during a flight in light of new task-relevant information. Endsley's (1988; 1995) three-level model of SA is perhaps the most frequently cited theory. In this SA is defined as "…the perception of the elements in the environment within a volume of time and space [SA level 1], the comprehension of their meaning [SA level 2], and the projection of their status in the near future [SA level 3]" (Endsley, 1995; p. 36). At level 1, SA is based on the perception of the basic building blocks of data obtained from the environment, cockpit systems, and tactical communications, etc. At the second level, obtained data are integrated to form a holistic understanding of the situation, aiming to achieve the comprehension of all significant elements within it: data start to become information (Ackoff, 1989). Level 3 SA addresses the pilot's understanding of the tactical situation and how it will evolve in the near future. Ensley's model is essentially hierarchical in nature, with each level building upon the preceding lower level; poor SA at a lower-level results in low SA at subsequently higher levels (e.g., Endsley and Garland, 2000). SA forms the basis of pilot decision-making.

In a flight, each pilot will have their own SA, however, the flight also has a collective team SA (TSA). TSA is more complex than individual SA. There are many definitions of TSA. Endsley (1995) defined TSA as "the degree to which every team member possesses the SA required for his or her responsibilities" (p. 39). Salas et al. (1995) suggested that "TSA is at least in part the shared understanding of a situation among team members at one point in time" (p. 131). Wellens (1993) defined it as "the sharing of a common perspective between two or more individuals regarding current environmental events, their meaning and projected future status" (p.6). Salmon et al. (2008) suggested that TSA comprised the SA of individual members; their shared SA (the elements of a common mental model combined with an appreciation of individual responsibilities), and the emergent, "common picture" which was the combined SA of the whole team.

A flight uses its TSA to understand and predict the progress of an air combat engagement, and hence to select appropriate TTPs to enable the execution of kill- or live-chains (Rouse & Morris, 1986; Mansikka et al., 2023b). TSA builds upon the theory of transactive memory (Wegner, 1985) and shared cognition (Rogers, 1997). A transactive memory system (TMS) has two basic components: the knowledge resident in each team member, and a set of transactive processes concerning what members collectively understand about the knowledge held by others. The knowledge possessed by team members can be highly variable and redundant. Rogers (1997) further illustrated the generic properties of cognition in teams, describing how members interact, allowing them to pool their cognitive resources and share knowledge through both implicit and formal communication, building upon their prior knowledge of each other.

The measurement of TSA in a fast-moving, dynamic context such as air combat poses several difficulties. Techniques which require pausing activities to make SA measurements (e.g., Cooke et al., 1997; Bolstad & Endsley, 2003; Sulistyawati et al., 2009) are not always possible, especially during a live exercise. Salmon et al. (2009) criticized the validity of this approach, suggesting it was unclear if it was SA or recall memory being assessed. Sulistyawati et al. (2009) also used combat performance measures to assess TSA effectiveness, but this can be a misleading as there is often a dissociation between SA and OP, TP or both (see Mansikka et al., 2019b).

Self- or peer-appraisal assessment techniques (e.g., Weigl et al., 2020) do not intrude on TP, but have been criticized as they may reflect pilot confidence or knowledge rather than TSA (Lichacz, 2006; Prince et al., 2007). Fowlkes et al. (1994), Salas et al. (1995), Bolstad and Endsley (2003), and Gorman et al. (2006) all argue that TSA is an emergent state, rather than a product of teamwork. As a result, the best method to assess TSA is by examining teamwork behaviours. Rosenman et al. (2018) developed a measurement approach based upon post-task probes of SA. The score was based upon the response accuracy to these questions, and the TSA metric was derived from averaging the pairwise agreement for each dyad in the team.

To address some of these measurement issues Mansikka et al. (2021c; 2023a) developed a TSA measurement technique based upon Endsley's freeze-probe SAGAT (Situation Awareness Global Assessment Technique; Endsley, 1988) combined with a shortened form of the critical decision-making (CDM) interview approach (Crandall et al., 2006). During a post-sortie debrief, an instructor pilot (IP) would reconstruct the mission using cockpit video, simulator flight trajectories, sensor tracks, weapon simulations, etc. At significant decision points the post-sortie replay would be paused, and a set of questions based upon relevant attributes relating to different levels of SA would be asked. With the help of deepening probes, the pilots were assisted to compare their expectations with the manner in which the simulated combat situation evolved.

In contrast to other measures of TSA, Mansikka et al. (2021c; 2023a) derived two related metrics: TSA accuracy (TSA ACC) and TSA similarity (TSA SIM). TSA ACC is a measure of how closely the flight's collective knowledge is aligned with the ground truth and TSA similarity (TSA SIM) is a measure of the similarity of team members' SA (Mansikka et al., 2021c).

To assess TSA ACC, each pilot's SA accuracy for a relevant attribute at a decision point was established. To do that, the accuracy of the pilot's cognitive model of the situation was compared to the simulation ground truth. The SA accuracies were scored, and scores were aggregated to provide an overall SA accuracy score for a single pilot (see Mansikka et al. 2021c; 2023a for further details). Separate scores were calculated for each SA level. By summing the individual pilots' SA accuracy scores at each SA level, level 1-3 TSA ACC indexes for the flight were obtained. TSA SIM was determined by making pairwise comparisons of level 1-3 SA accuracy at each decision point between all members of the flight (Mansikka et al., 2023a). A higher score was associated with a higher similarity.

For most purposes, the evaluation of TSA ACC is adequate but taken alone it does not portray a complete picture of the shared situational knowledge possessed by the flight. If TSA ACC is high, the flight's TSA is closely aligned with the ground truth, and in such a situation, the SA of each flight member must also be very similar. If TSA ACC is low, however, flight members can have similar or dissimilar SA.

It was found that a flight's TP showed a curvilinear relationship with TSA ACC (Mansikka et al., 2021c; 2023a). As TSA accuracy increased, there were diminishing returns in TP. Gains in TP decreased disproportionately with increases in TSA. Furthermore, the strongest predictors of OP or TP (or both) were level 1 TSA measures. Low TSA SIM was found to have a negative impact on combat performance both offensively and defensively (Mansikka et al., 2023a). These findings were consistent with Endsley's hierarchical model of SA.

The main challenge in achieving high TSA is the coordination of the flight members (Mansikka et al., 2023b; 2023c). If both aspects of TSA are high (ACC and SIM) the pilots can anticipate the actions of other flight members without communication. Explicit coordination depends upon active communication, whereas implicit coordination relies on the flight coordinating its members' actions without such verbal efforts. Implicit coordination is based on members' shared knowledge about each other, the task, and the task environment. In so doing this enables flight members to anticipate each other's actions without need for overt communication (Entin & Serfaty, 1999; Rico et al., 2008; Stout et al., 2017). Team members require common knowledge about the task situation during the action phase while they are updating their knowledge by ongoing situation assessment. Implicit coordination enables a flight to rapidly synchronize its activities.

Team training has been found to facilitate the formation of TMSs as well as TSA. For stable groups, such as a flight, the effect of team training also extends beyond the task for which they are initially trained (Lewis et al., 2005). On the other hand, communication facilitates the encoding, storage, retrieval, and update of information drawn from individual memory system components. Peltokorpi and Hood (2019) observed that the frequency of communication decreases over time as teams increase in familiarity. Mansikka et al. (2023a) examined the relationship between within-flight communications, TSA and OP. During simulated engagements TSA ACC nonlinearly increased with a concomitant decrease in the number of SA-related communication acts. In a similar vein, when TSA started to deteriorate, additional communication as a means to recover TSA was observed not to be effective; increased communication was associated with poorer TSA. There was less communication in successful engagements as measured by OP compared to unsuccessful engagements. In highly time-pressured, high-workload, and extremely dynamic situations such as air combat, attempts at explicit coordination may be counterproductive. In such cases, implicit coordination based upon the common knowledge developed during team training is the better option.

Dissociation of SA, MWL, and OP

Both Vidulich and Wickens (1986) and Yeh and Wickens (1988) noted that MWL can become dissociated from performance, particularly if the task is resource limited. The relationship between SA and performance (OP and TP) can also be complex and unclear (Sulistyawati et al., 2009; Mansikka et al., 2019b). A pilot's awareness of the task demands, which is predicated upon their SA, will partially determine their MWL. Mansikka et al. (2019b) observed that when pilots had low awareness of the tactical situation, they also exhibited a combination of low workload and poor OP and/or TP: they were not aware of the need to invest more cognitive effort to enhance their SA. Furthermore, in a highly dynamic and uncertain environment such as air combat, both success and failure can occasionally be a product of chance (Mansikka et al., 2021a).

In summary

Focusing only on the outcome of a team's task, i.e., the product of taskwork, may give a biased estimate for the performance of a team and provide little explanation for the observed level of performance. A much richer picture of team performance can be obtained by also considering taskwork processes and teamwork, including the coordination mechanisms of team members. Such a holistic performance evaluation can be carried out with the five-dimensional measurement approach discussed. It offers a way to explain why a specific performance level has, or has not, been achieved in an understandable, transparent, and traceable manner. This kind of insight can be utilized in many ways, for example, when identifying means to improve the performance of a team. It should be noted that dependences between the performance metrics discussed, as well as the utility of any metric depends on the context in which the team performance is evaluated. All of them might not be used in all cases but it is important to acknowledge the availability of these complementary measurement techniques when assessing performance of teams undertaking complex tasks.

Conclusion

Team performance, as we have seen, is more than mere aggregates of individual performance. Teams are complex networks where each member's performance, well-being, and development are inextricably linked to the collective functioning of the group. The synergy that emerges from effective team dynamics transcends individual capabilities, leading to enhanced performance, resilience, and fulfilment.

From the foundational work at the Hawthorne Works to the innovative approaches in special operations forces and flight crews, the evolution of team performance research and application demonstrates a continuous quest to understand and optimize the interplay of individual skills, team processes, and environmental factors. The progression of our understanding of linear IPO models to a more dynamic conceptualization of teams as complex adaptive systems reveals a deepening appreciation of the nuanced interrelations that define team performance.

Real-world applications further illustrate a practical process for developing team training programs. These case studies highlight the critical role of team assessments, both informal and formal, in diagnosing and enhancing team performance. They show how tailored interventions, based on a deep understanding of team dynamics and individual contributions, can significantly improve both team effectiveness and member satisfaction.

In summary, this chapter illuminates the complexities and rewards of effective team performance. It underscores the importance of understanding and nurturing the intricate dynamics within teams to harness their full potential. The profound truth that no one is an island finds its embodiment in the realm of teams, where the collective interplay of skills, knowledge, and emotions shapes not only the outcomes of team endeavours but also the personal growth and resilience of each member. As we continue to explore and apply these insights, we contribute to a world where teams not only achieve their goals but also foster environments of collaboration, innovation, and mutual support.

References

Aubé, C., Brunelle, E., & Rousseau, V. (2014). Flow experience and team performance: The role of team goal commitment and information exchange. *Motivation and Emotion*, 38(1), 120–130.

Ackoff, R. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16(1), 3–9.

Annett, J. (2002). Subjective rating scales: Science or art? *Ergonomics*, 45(14), 966–987.

Arafeh, J. M., Hansen, S. S., & Nichols, A. (2010). Debriefing in simulated-based learning: facilitating a reflective discussion. *The Journal of Perinatal & Neonatal Nursing*, 24(4), 302–309.

Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems: Formation, coordination, development, and adaptation*. Sage Publications.

Bar-On, R. (1997). *The Emotional Intelligence Inventory (EQ-i): technical manual*. Toronto, Canada: Multi-Health Systems.

Bar-On, R. (2004). *The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description and summary of psychometric properties*. In G. Geher (Ed.), Measuring emotional intelligence: Common ground and controversy (pp. 115–145). Nova Science Publishers.

Boyle, G. J. (1995). Myers-Briggs type indicator (MBTI): Some psychometric limitations. *Australian Psychologist*, *30*(1), 71–74.

Beech, D. (2023). The development and validation of the Optimizing Team Performance Profile — OTP Profile. Retrieved on 19 Nov 2023 from www.otpprofile.com/profile-validation-and-research

Beech, D., Gist, R., & Corneliussen, J. (2023). Optimizing Team Performance. In R. Hauffa., & N. Pattyn (Eds.), *Optimizing Human Performance in NATO SOF Personnel Through Evidence-Based Mental Performance Programming.* NATO HFM-308 Working Group.

Berka, C., Levendowski, D., Lumicao, M., Yau, A., Davis, G., Zivikovic, V., Olmstead, R., Tremoulet, P., & Craven, P. (2007). EEG correlates of task engagement and mental workload in vigilance and memory tasks. *Aviation, Space and Environmental Medicine*, 78(5), Supplement 1, 231–244.

Bolstad, C., & Endsley, M. (2003). Measuring shared and team situation awareness in the army's future objective force. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47(3), 369–373.

Brown, I. (1978). Dual task methods of assessing workload. *Ergonomics*, 21(3), 221–224.

Burke, M. J., Salvador, R., Smith-Crowe, K., Chan-Serafin, S., Smith, A., Kataria, N., & ... Halvorsen-Ganepola, M. D. K. (2017). The dread, fatigue, and prediction of team performance: An integrative model. *Journal of Occupational Health Psychology,* 22(3), 363–372. doi:10.1037/ocp0000053

Burns Jr, W. A. (2017). A descriptive literature review of harmful leadership styles: Definitions, commonalities, measurements, negative impacts, and ways to improve these harmful leadership styles. *Creighton Journal of Interdisciplinary Leadership*, 3(1), 33–52.

Caesens, G., Stinglhamber, F., Demoulin, S., De Wilde, M., & Mierop, A. (2019). Perceived organizational support and workplace conflict: The mediating role of failure-related trust. *Frontiers in Psychology*, 9, 2704.

Casali, J., & Wierwille, W. (1984). On the measurement of pilot perceptual workload: A comparison of assessment techniques in a simulated flight task emphasizing communications load. *Human Factors*, 25(6), 623–641.

Casner, S., & Gore, B. (2010). *Measuring and evaluating workload: A primer. NASA technical memorandum 2010-216395*. San Jose, CA: NASA Ames Research Center.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper & Row.

Csikszentmihalyi, M. (2014). *Flow and the foundations of positive psychology*. Springer.

Cooke, N., Stout, R., & Salas, E. (1997). Broadening the measurement of situation awareness through cognitive engineering methods. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 41(1), 215–219.

Cooper, C., & Harper, R. (1969). *The use of pilot rating in the evaluation of aircraft handling qualities*. (NASA TN-D-5153). Moffett Field, CA: NASA Ames Research Center.

Costa, P. T., Jr., & McCrae, R. R. (2008). The Revised NEO personality inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE handbook of personality theory and assessment*, Vol. 2. Personality measurement and testing (pp. 179–198). Sage Publications, Inc. https://doi.org/10.4135/9781849200479.n9

Crandall, B., Kein, G., & Hoffman, R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*, MIT Press, Cambridge, CA.

Day, D., Bastardoz, N., Bisbey, T., Reyes, D., & Salas, E. (2021). Unlocking human potential through leadership training & development initiatives. *Behavioral Science & Policy*, 7(1), 41–54.

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

de Jong, B., Dirks, K. & Gillespie, N. (2016). Trust and team performance: A meta-analysis of main effects, moderators and covariates. *Journal of Applied Psychology*, 101(8). DOI: 10.1037/apl0000110

Delise, L. A., Allen Gorman, C., Brooks, A. M., Rentsch, J. R., & Steele-Johnson, D. (2010). The effects of team training on team outcomes: A meta-analysis. *Performance Improvement Quarterly*, 22(4), 53–80. https://doi.org/10.1002/piq.20068

Denison, D. R., Janovics, J., Young, J., & Cho, H. J. (2006). Diagnosing organizational cultures: Validating a model and method. *Documento de trabajo. Denison Consulting Group*, *1*(1), 1–39.

Donne, J. (1624). Devotions upon emergent occasions: No. 17. Meditation. In J. M. Dent & Co. (Eds.), *Collected poems of John Donne* (pp. 404). (Originally published in 1624).

Donne, J. (2001). *The complete poetry and selected prose of John Donne*. Modern library.

Duhigg, C. (2016). What google learned from its quest to build the perfect team. *The New York Times*. Retrieved on 8 November 2023 from www.nytimes.com/2016/02/28/magazine/what-google-learned-from-its-quest-to-build-the-perfect-team.html

Dunst, C. J., Bruther, M. B., Hamby, D. W., Howse, R., & Wilkie, H. (2018). Meta-analysis of the relationships between different leadership practices and organizational, teaming, leader, and employee outcomes. *Journal of International Education and Leadership*, 8(2), n2.

Eddy, E. R., Tannenbaum, S. I., & Mathieu, J. E. (2013). Helping teams to help themselves: Comparing two team-led debriefing methods. *Personnel Psychology*, 66(4), 975–1008. https://doi.org/10.1111/peps.12041

Eldredge, L. K. B., Markham, C. M., Ruiter, R. A., Fernández, M. E., Kok, G., & Parcel, G. S. (2016). *Planning health promotion programs: An intervention mapping approach*. John Wiley & Sons.

Eisenberger, N. I., & Cole, S. W. (2012). Social neuroscience and health: Neurophysiological mechanisms linking social ties with physical health. *Nature Neuroscience*, 15(5), 669–674. https://doi.org/10.1038/nn.3086

Eisenberger, N. I., Taylor, S. E., Gable, S. L., Hilmert, C. J., & Lieberman, M. D. (2007). Neural pathways link social support to attenuated neuroendocrine stress responses. *Journal of Personality and Social Psychology*, 93(1), 7–19. https://doi.org/10.1037/0022-3514.93.1.7

Endsley, M. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, 32(2), 97–101.

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

Endsley, M., & Garland, D. (2000). *Situation awareness: Analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Entin, E., & Serfaty, D. (1999). Adaptive team coordination. *Human Factors*, 41(2), 312–325.

Epps, H. R., & Levin, P. E. (2015). The TeamSTEPPS approach to safety and quality. *Journal of Pediatric Orthopaedics,* 35, S30–S33.

Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 2(2), 115–125. https://doi.org/10.1097/SIH.0b013e3180315539

Fischer, T., & Sitkin, S. B. (2023). Leadership styles: A comprehensive assessment and way forward. *Academy of Management Annals*, 17(1), 331–372.

Fisher, J., Young, C., & Fadel, P. J. (2009). Central sympathetic overactivity: Maladies and mechanisms. *Autonomic Neuroscience*, 148(1), 5–15. https://doi.org/10.1016/j.autneu.2009.02.003

Fowlkes, J., Lane, N., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The TARGETs methodology. *Military Psychology,* 6(1), 47–61.

Frueh, B.C., Madan, A., Fowler, J. C., Stomberg, S., Bradshaw, M., Kelly, K., Weinstein, B., Luttrell, M., Danner, S. G., & Beidel, D. C. (2020). "Operator syndrome": A unique constellation of medical and behavioral health-care needs of military special operation forces. *International Journal of Psychiatry in Medicine*, 55(4), 281–295. https://doi.org/10.1177/0091217420906659

Gale, E. A. (2004). The Hawthorne studies—a fable for our times?. *Qjm*, 97(7), 439–449.

Gergen, K. J., & Gergen, M. M. (2008). Social construction and research as action. *The Sage handbook of action research: Participative inquiry and practice*, 159–171. Ed Bradbury, Thousand Oaks, Sage.

Giuliano, R., Gatzke-Kopp, L. M., Roos, L. E., & Skowron, E. A. (2017). Resting sympathetic arousal moderates the association between parasympathetic reactivity and working memory performance in adults reporting high levels of life stress. *Psychophysiology*, 54(8), 1195–1208. https://doi.org/10.1111/psyp.12872.

Gorman, J., Cooke, N., & Winner, J. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, 49(12–13), 1312–1325.

Griffith, J., Roberts, D., & Wakeham, R. (2015). A Meta-Analysis of crew resource management/incident command systems implementation studies in the fire and emergency services. *Journal of Aviation/Aerospace Education & Research*, 25(1), 1–25. https://doi.org/10.15394/jaaer.2015.1647

Grossman, R., Friedman, S., & Suman, K. (2019). Teamwork Processes and Emergent States. In E. Salas, R. Rico, & J. Passmore (Eds.). *The Wiley Blackwell handbook of the psychology of team working and collaborative processes* (pp. 245–269). Wiley Blackwell Publishing.

Gully, S. M., Incalcaterra, K. A., Joshi, A., & Beaubien, J. M. (2002). A meta-analysis of team-efficacy, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology*, 87(5), 819–832. https://doi.org/10.1037/0021-9010.87.5.819.

Hadfield C. (2015). *An astronaut's guide to life on Earth*. MacMillan.

Hamann, A., & Carstengerdes, N. (2022). Investigating mental workload-induced changes in cortical oxygenation and frontal theta activity during simulated flights. *Scientific Reports*, 12(1), 6449.

Hollnagel, E. (1998). Cognitive reliability and error analysis methodology. London: Elsevier.

Houghton, A., & Gray, D. (2014). *Making sense of the ECG: A hands-on guide*. CRC Press, Boca Raton, FL.

Hughes, A. M., Gregory, M.E., Joseph, D.L., Sonesh, S.C., Marlow, S.L. (2016). Saving lives: a meta-analysis of team training in healthcare. *Journal of Applied Psychology*, 101, 1266–304

Ilgen, D., Hollenbeck, J., Johnson, M., & Jundt, D. (2005). Teams in organizations: From input-process-output models to IMOI models. *Annual Review of Psychology*, 56, 517–543.

Ivanov, P., Amaral, L., Goldenberger, A., Havlin, S., Rosenblum, M., Struzik, Z., & Stanley, H. (1999). Multifractality in human heartbeat dynamics. *Nature*, 399(6735), 461–465.

Jackson, S. A., & Csikszentmihalyi, M. (1999). *Flow in sports*. Human Kinetics.

Jedick, R. (2014). United Airlines 173 — The need for CRM. *Go flight Medicine*. https://goflightmedicine.com/united-airlines-173/

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, 40(2), 256–282.

Joint Chief of Staff. (2013). *Joint Publication 3-60: Joint Targeting*. Washington, DC: US Joint Chief of Staff.

Jorgensen, D.L. (1989). *Participant Observation. A Methodology for Human Studies*. Sage Publications.

Kiely, J. (2018). Periodization theory: Confronting an inconvenient truth. *Sports Medicine (Auckland*), 48(4), 753–764. https://doi.org/10.1007/s40279-017-0823-y

Kinnunen, H., & Koskimäki, H. (2018). The HRV of the ring-comparison of nocturnal HR and HRV between a commercially available wearable ring and ECG. *Sleep*, 41(1), 120.

Kinnunen, H., Rantanen, A., Kenttä, T., & Koskimäki, H. (2020). Feasible assessment of recovery and cardiovascular health: Accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG. *Physiological Measurement*, 41(4), 1–9.

Kirwan, B., & Ainsworth, L. K. (Eds.). (1992). *A guide to task analysis: the task analysis working group*. CRC press.

Klein, G. (2009). *Streetlights and shadows: Searching for the keys to adaptive decision making*. Cambridge, MA: MIT Press.

Kramer, A, Sirevaag, E., & Braune, R. 1987. A psychological assessment of operator workload during simulated flight missions. *Human Factors*, 29(2), 145–160.

Lacerenza, C. N., Marlow, S. L., Tannenbaum, S. I., & Salas, E. (2018). Team development interventions: Evidence-based approaches for improving teamwork. *American Psychologist*, 73(4), 517–531. https://doi.org/10.1037/amp0000295

Lacerenza, C. N., Reyes, D. L., Marlow, S. L., Joseph, D. L., & Salas, E. (2017). Leadership training design, delivery, and implementation: A meta-analysis. *Journal of Applied Psychology*, 102(12), 1686–1718. https://doi.org/10.1037/apl0000241

Landhaüßer, A., & Keller, J. (2012). Flow and its affective, cognitive, and performance-related consequences. In S. Engeser (Ed.), *Advances in flow research* (pp. 65–85). Springer Science.

Lewis, K., Lange, D., & Gills, L. (2005). Transactive memory systems, learning, and learning transfer. *Organization Science*, 16(6), 581–598.

Lichacz, F. (2006). An examination of situation awareness and confidence within a distributed information sharing environment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(3), 344–348.

Mansikka, H., Simola, P., Virtanen, K., Harris, D., & Oksama, L. (2016a). Fighter pilots' heart rate, heart rate variation and performance during instrument approaches. *Ergonomics*, 59(10), 1344–1352.

Mansikka, H., Virtanen, K., & Harris, D. (2019a). Comparison of NASA-TLX scale, modified Cooper–Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. *Ergonomics*, 62(2), 246–254.

Mansikka, H., Virtanen, K., & Harris, D. (2019b). Dissociation between mental workload, performance, and task awareness in pilots of high performance aircraft. *IEEE Transactions on Human-Machine Systems*, 49(1), 1–9.

Mansikka, H., Virtanen, K., & Harris, D. (2023a). Accuracy and similarity of team situation awareness in simulated air combat. *Aerospace Medicine and Human Performance*, 94(6), 429–436.

Mansikka, H., Virtanen, K., & Harris, D. (2023c). What we got here, is a failure to coordinate: Implicit and explicit coordination in air combat. *Journal of Cognitive Engineering and Decision Making*, 17(3), 279–293.

Mansikka, H, Virtanen, K., Harris, D., & Jalava M. (2021a). Measurement of team performance in air combat — have we been underperforming? *Theoretical Issues in Ergonomics Science*, 22(3), 338–359.

Mansikka, H., Virtanen, K., Harris, D., & Järvinen, J. (2023b). Team performance in air combat: A teamwork perspective. *International Journal of Aerospace Psychology*, 33(4), 232–246.

Mansikka, H., Virtanen, K., Harris, D., & Salomäki, J. (2021d). Live–virtual–constructive simulation for testing and evaluation of air combat tactics, techniques, and procedures, Part 1: Assessment framework. *Journal of Defense Modeling and Simulation*, 18(4), 285–293.

Mansikka, H., Virtanen, K., Harris, D., & Simola, P. (2016b). Fighter pilots' heart rate, heart rate variation and performance during an instrument flight rules proficiency test. *Applied Ergonomics,* 56, 213–219.

Mansikka, H., Virtanen, K., Mäkinen, L., & Harris, D. (2021b). Normative performance measurement in simulated air combat. *Aerospace Medicine and Human Performance*, 92(11), 908–912.

Mansikka, H., Virtanen, K., Uggeldahl, V., & Harris, D. (2021c). Team situation awareness accuracy measurement technique for simulated air combat- Curvilinear relationship between awareness and performance. *Applied Ergonomics,* 96, 103473.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review*, 26(3), 356–376.

Marshall, D. (2009). *Crew resource management: from patient safety to high reliability*. Safer Healthcare Partners, LLC.

Mathieu, J., Gallagher, P., Domingo, M. & Klock, E. (2019). Embracing Complexity: Reviewing the Past Decade of Team Effectiveness Research. *Annual Review of Organizational Psychology and Organizational Behavior.* 6, 17–46.

Mathieu, J., Hollenbeck, J., van Knippenberg, D. & Ilgen, D. (2017). A century of work teams in the journal of applied psychology. *Journal of Applied Psychology*, 102 (3), 452–467.

Mathieu, J. E., Luciano, M. M., D'Innocenzo, L., Klock, E. A., & LePine, J. A. (2020). The development and construct validity of a team processes survey measure. *Organizational Research Methods*, *23*(3), 399–431.

Mathieu, J., 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, 410–476. http://dx.doi.org/10.1177/0149206308316061.

Mattie, C. P., Guest, K., Bailey, S., Collins, J., & Gucciardi, D. F. (2020). Development of a mental skills training intervention for the Canadian Special Operations Forces Command: An intervention mapping approach. *Psychology of Sport and Exercise*, 50, 101720.

McEwan, D., Ruissen, G., Eys, M., Zumbo, B. & Beauchamp, M. (2017). The effectiveness of teamwork training on teamwork behaviors and team performance: a systematic review and meta-analysis of controlled interventions. *PLoS ONE*. 12(1), e0169604. doi:10.1371/journal. pone.0169604

McEwen, B.S. (2000). Allostasis and allostatic load: Implications for neuropsychopharmacology. *Neuropsychopharmacology*, 22, 108–24.

McGrath, J. E. (1997). Small group research, that once and future field: An interpretation of the past with an eye to the future. *Group Dynamics: Theory, Research, and Practice*, 1, 7–27. http://dx.doi.org/10.1037/1089-2699.1.1.7

McIntyre, A. (2007). *Participatory action research*. Sage.

McTaggart, R. (1991). Principles for participatory action research. *Adult Education Quarterly*, 41(3), 168–187.

Mehler, B., Reimer, B., & Wang, Y. (2011). A comparison of heart rate and heart rate variability indices in distinguishing single task driving and driving under secondary cognitive workload. Proceedings of the Sixth International Driving Symposium on Human Factors in Driver Assessment, *Training and Vehicle Design*, 590–597.

Morrison, J. E., & Meliza, L. L. (1999). Foundations of the after-action review process (Special Report 42). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. https://apps.dtic.mil/dtic/tr/fulltext/u2/a368651.pdf

National Transportation Safety Board (1978) *Aircraft accident report: United Airlines*, Inc. NTSB-AAR-79-7. Retrieved from www.ntsb.gov/investigations/ AccidentReports/Reports/AAR7907.pdf

O'Connor, P., Campbell, J., Newon, J., Melton, J., Salas, E., & Wilson, K. A. (2008). Crew resource management training effectiveness: Ameta-analysis and some critical needs. *The International Journal of Aviation Psychology*, 18(4), 353–368.

O'Dea, A., O'Connor, P., & Keogh, I. (2014). A meta-analysis of the effectiveness of crew resource management training in acute care domains. *Postgraduate Medical Journal*, 90(1070), 699–708. https://doi.org/10.1136/postgradmedj-2014-132800

O'Donnel, R., Eggemeier, F., & Thomas, F. (1986). Workload assessment methodology. In Boff, K., Kaufman, L., and Thomas, J. (Eds.), *Handbook of Perception and Human Performance* Vol 2: 42, 1–49. Wiley Interscience, New York, NY.

Ogden, G., Levine, J., & Eisner, J. (1979). Measurement of workload by secondary tasks. *Human Factors*, 21(5), 529–548.

Paas, F., Tuovinen, J., Tabbers, H., & van Gerven, P. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2008). Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making*, *2*(2), 140–160.

Pattyn, N., Van Cutsem, J., Lacroix, E., Van Puyvelde, M., Cortoos, A., Roelands, B., Tibax, V., Dessy, E., Huret, M., Rietjens, G., Sannen, M., Vliegen, R., Ceccaldi, J., Peffer, J., Neyens, E., Duvigneaud, N., & Van Tiggelen, D. (2022). Lessons from special forces operators for elite team sports training: How to make the whole greater than the sum of the parts. *Frontiers in Sports and Act Living*, 4, 780767. doi: 10.3389/fspor.2022.780767

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46(3), 598–609. https://doi.org/10.1037/0022-3514.46.3.598

Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). Academic Press. https://doi.org/10.1016/B978-0-12-590241-0.50006-X.

Peltokorpi, V., & Hood, A. (2019). Communication in theory and research on transactive memory systems: A literature review. *Topics in Cognitive Science* 11(2019), 644–667.

Prince, C., Ellis, E., Brannick, M., & Salas, E. (2007). Measurement of team situation awareness in low experience level aviators. *International Journal of Aviation Psychology*, 17(1), 41–57.

Ramos-Villagrasa, P., Marques-Quinteiro, P., Navarro, J., & Rico, R. (2018). Teams as complex adaptive systems: Reviewing 17 years of research. *Small Group Research*, 49(2), 135–176. https://doi.org/10.1177/1046496417713849

Reid, J., Stone, K., Brown, J., Caglar, D., Kobayashi, A., Lewis-Newby, M., ... & Quan, L. (2012). The simulation team assessment tool (STAT): Development, reliability and validation. *Resuscitation*, 83(7), 879–886

Reimer, B., & Mehler, B. (2011). The impact of cognitive workload on physiological arousal in young adult drivers: A field study and simulation validation. *Ergonomics*, 54(10), 932–942.

Reyes, D. L., Tannenbaum, S. I., & Salas, E. (2018). Team development: The power of debriefing. *People & Strategy*, 41(2), 46–52.

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.

Rogers, Y. (1997). A brief introduction to distributed cognition. Retrieved from http://yvonnerogers.com/wp-content/uploads/2014/07/dcog-brief-intro.pdf

Roman-Liu, D., Grabarek, I., Bartuzi, P., & Choromanski, W. (2013). The influence of mental load on muscle tension. *Ergonomics*, 56(7), 1125–1133.

Rosenman, E., Dixon, A., Webb, J., Brolliar, S., Golden, S., Jones, K., Sachita, S., Grand, J., Kozlowski, S., Chao, G., & Fernandez, R. (2018). A simulation-based approach to measuring team situational awareness in emergency medicine: a multicenter, observational study. *Academic Emergency Medicine*, 25(2), 196–204.

Rouse, W., & Morris, N. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100(3), 349–363.

Ryu, K., & Myung, R. (2005). Evaluation of mental workload with combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *International Journal of Industrial Ergonomics*, 35(1), 991–1009.

Salanova, M., Llorens, S., Cifre, E., & Martínez, I. M. (2013). We need a hero! Toward a validation of the healthy and resilient organization (HERO) model. *Group & Organization Management*, 38(5), 569–598.

Salas, E. (2015b). *Team training essentials: A research-based guide*. Routledge.

Salas, E., Bowers, C. A., & Cannon-Bowers, J. A. (1995). Military team research: 10 years of progress. *Military Psychology*, 7(2), 55–75.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). *Toward an understanding of team performance and training*. CRC Press.

Salas, E., Priest, H., & DeRouin, R. (2005a) Team building. In N. Stanton (Ed.), *Handbook of human factors and ergonomics methods*. (pp. 465–470). CRC Press.

Salas, E., Shuffler, M. L., Thayer, A. L., Bedwell, W. L., & Lazzara, E. H. (2015a). Understanding and improving teamwork in organizations: A scientifically based practical guide. *Human Resource Management*, 54(4), 599–622.

Salas, E., Sims, D., & Burke, C. (2005b). Is there a "Big Five" in Teamwork?. *Small Group Research*, 36(5), 555–599.

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest*, 13(2), 74–101. https://doi.org/10.1177/1529100612436661

Salas, E., Wilson, K. A., Burke, C. S., Wightman, D. C., & Howse, W. R. (2006). Crew resource management training research, practice, and lessons learned. *Reviews of Human Factors and Ergonomics*, 2(1), 35–73.

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

Salmon, P., Stanton, N., Walker, G., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, 39(3), 490–500.

Sampson, M., & McGrath, A. (2015). Understanding the ECG Part 2: ECG basics. *British Journal of Cardiac Nursing*, 10(12), 588–594.

Sawyer, R. K. (2003). *Group creativity music, theater, collaboration*. Erlbaum Associates.

Schuermann, V., & Marquardt, N. (2016). Adaptation of crew resource management training in high-risk industries. *International Journal of Safety and Security Engineering*, 6(2), 341–350.

Schyns, B., & Schilling, J. (2013). How bad are the effects of bad leaders? A meta-analysis of destructive leadership and its outcomes. *The Leadership Quarterly*, 24(1), 138–158.

Seitchik, M. (2019). The Goldilocks approach to team conflict: How leaders can maximize innovation and revenue growth. *The Psychologist-Manager Journal*, 22(1), 37–45. https://doi.org/10.1037/mgr0000082

Setz, C., Armrich, B., Schumm, J., La Marca, R., Troster, G., & Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable EDA device. *Information Technology in Biomedicine*, 14(2), 410–417.

Shuffler, M. L., Salas, E., & Xavier, L. F. (2010). The design, delivery and evaluation of crew resource management training. In Kanki, B., Helmreich, R.,; Anca, J. (Eds) *Crew resource management* (pp. 205–232). Academic Press.

Spradley, J. P. (2016). *Participant obse*rvation. Waveland Press.

Stein, R, Swan, A. B. (2019). Evaluating the validity of Myers-Briggs Type Indicator theory: A teaching tool and window into intuitive psychology. *Social and Personality Psychology Compass*, 13, e12434. https://doi.org/10.1111/spc3.12434.

Steiner, I. D. (1974). Whatever happened to the group in social psychology? *Journal of Experimental Social Psychology*, 10, 94–108. http://dx.doi.org/10.1016/0022-1031(74)90058-4

Stewart, V.R., Snyder, D.G., Kou, C.Y. (2023). We hold ourselves accountable: A relational view of team accountability. *Journal of Business Ethics*, 183(3), 691–712. doi: 10.1007/s10551-021-04969-z. Epub 2021 Nov 18. PMID: 34812211; PMCID: PMC8600914

Stone, J., Ulman, H., Tran, K., Thompson, A., Halter, M., Ramadan, J., & Hagen, J. (2021). Assessing the accuracy of popular commercial technologies that measure resting heart rate and heart rate variability. *Frontiers in Sports and Active Living*, 37.

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

Sulistyawati, K., Wickens, C., & Chui, Y. (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.

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors*, 55(1), 231–245. https://doi.org/10.1177/0018720812448394

van den Hout, J., Davis, O. & Weggeman, M. (2018). The conceptualization of team flow. *The Journal of Psychology*, 152(6), 388–423, DOI: 10.1080/00223980.2018.1449729

Vashdi, D. R. (2013). Teams in public administration: A field study of team feedback and effectiveness in the Israeli public healthcare system. *International Public Management Journal*, 16(2), 275–306.

Verwey, W, & Veltman, H. (1996). Detecting short periods of elevated workload: A comparison of nine workload assessment techniques. *Journal of Experimental Psychology*, 2(3), 270–285.

Vicente, K., Thornton, D., & Moray, N. (1987). Spectral analysis of sinus arrythmia: A measure of mental effort. *Human Factors*, 29(2), 171–182.

Vidulich, M., & Wickens, C. (1986). Causes and dissociation between subjective workload measures and performance. *Applied Ergonomics*, 17(4), 291–296.

Virtanen, K., Mansikka, H., Kontio, H., & Harris, D. (2022). Weight watchers: NASA-TLX weights revisited. *Theoretical Issues in Ergonomics Science*, 23(6), 725–748.

Wageman, R., Hackman, J. R., & Lehman, E. (2005). Team diagnostic survey: Development of an instrument. *The Journal of Applied Behavioral Science*, 41(4), 373–398.

Weigl, M., Catchpole, K., Wehler, M., & Schneider, A. (2020). Workflow disruptions and provider situation awareness in acute care: An observational study with emergency department physicians and nurses. *Applied Ergonomics*, 88, 103155.

Wegner, D. M., Giuliano, T., & Hertel, P. T. (1985). Cognitive interdependence in close relationships. In W. J. Ickes (Ed.), Compatible and incompatible relationships (pp. 253–276). New York: Springer-Verlag.

Wellens, A., 1993. Group situation awareness and distributed decision making: from military to civilian applications. In: Castellan Jr., N.J. (Ed.), Individual and Group Decision Making: Current Issues. Lawrence Erlbaum Associates Inc, Hillsdale, NJ, pp. 267–291. https://doi.org/10.4324/9780203772744.

Wierwille, W. W., Rahimi, M., & Casali, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human factors*, *27*(5), 489–502.

Yeh, Y. Y., & Wickens, C. D. (1988). Dissociation of performance and subjective measures of workload. *Human factors*, 30(1), 111–120.

Young, M. S., & Schinka, J. A. (2001). Research validity scales for the NEO-PI-R: Additional evidence for reliability and validity. *Journal of personality Assessment*, 76(3), 412–420.