Using Physiological Measures to Assess and Diagnose Team Performance:

An Interrogative Approach

Maha Khalid[1], Lila Berger[2], Eunyoung Kim[3], Jonathan Sussman-Fort[4], Jeffery Craighead[4], &

Eduardo Salas[1]

1. Department of Psychological Sciences, Rice University, Houston, TX, United States
2. Department of Human Factors and Behavioral Neurobiology, Embry-Riddle Aeronautical University, Daytona Beach, FL, United States
3. Ewha Womans University, Seoul, South Korea
4. Soar Technology, Orlando, FL, United States

Author Note

Maha Khalid https://orcid.org/0000-0003-3517-8960

Lila Berger, PhD https://orcid.org/0009-0002-8852-5129

Eunyoung Kim

Jonathan Sussman-Fort, PhD https://orcid.org/0009-0004-4455-9575

Jeffrey Craighead, PhD https://orcid.org/0000-0003-0033-4740

Eduardo Salas, PhD https://orcid.org/0000-0002-5448-4958

Correspondence concerning this article should be addressed to Maha Khalid, Department of Psychological Sciences, Rice University, 6100 Main Street, Sewall Hall 211H, Houston, TX 77005, USA. Email: mk127@rice.edu

Abstract

Physiological measures have the potential to capture the dynamic nature of teamwork as it unfolds. Despite this potential, relatively few studies have applied physiological data in team contexts, leaving a gap in theory and practice. To address this gap, this work examines the following question: *What are the emerging principles for using physiological measures to assess and diagnose team performance?* To explore this question, we present an interrogative approach that structures the key considerations for applying physiological measures in team research. Specifically, this approach organizes considerations related to the purpose, context, personnel, timing, measurement selection, and implementation of using physiological measures to assess team performance. Building on prior guidance, we propose a set of emerging principles that synthesize best practices while introducing new considerations specific to the use of physiological data in team contexts. These principles provide practical guidance for researchers and practitioners seeking to leverage physiological measures, while emphasizing that their effective use depends on rigorous methodology, ethical safeguards, and integration with other validated assessment approaches.

Keywords: *team, team performance, team dynamics, physiological dynamics, measurement, assessment*

Using Physiological Measures to Assess and Diagnose Team Performance: An Interrogative Approach

Over a century ago, Wilhelm Wundt (1904) outlined two challenges in *Principles of Physiological Psychology* that remain relevant today: the problem of method and supplement. In the problem of method, Wundt emphasized the need for experimental approaches capable of capturing the fleeting nature of mental processes, rather than relying solely on passive self-observation. The problem of supplement, in turn, called for connecting physiological measurements to psychological phenomena, recognizing that conscious experience must be connected to, but not reduced to, physical processes. Together, these challenges reflect the difficulty of capturing dynamic psychological processes, an issue that remains particularly relevant to the study of teamwork.

Significant gaps persist in quantitatively capturing team interactions as dynamic, multilevel, and continuously evolving processes (Salas et al., 2015). Addressing these gaps requires measurement approaches that bridge theory and practice by translating abstract team concepts into observable indicators of teamwork. Over the years, team science researchers have called for theories and methods that align with the need to capture the dynamic nature of team interactions (Kozlowski, 2015). To measure the dynamic nature of teamwork, the rate of measurement must keep pace with, or surpass, the speed at which the dynamic properties of teaming change; otherwise, important elements may be missed or inaccurately captured (Kozlowski & Chao, 2018). Tools that can capture team members' attitudes, behaviors, and cognitions in a real-time while remaining pragmatic, relevant, and unobtrusive are considered the holy grail of team performance measurement (Cooke, 2015; Kazi et al., 2021; Rosen et al., 2012; Salas et al., 2017).

The need for real-time team measurement becomes especially critical in high-stakes environments, where breakdowns in teamwork can contribute to safety incidents, performance degradation, and system failure (Leonard & Frankel, 2011; Salas et al., 2020). Extensive research supports this connection, showing that workplace safety is shaped not only by individual behavior, but also by the quality of teamwork, with safety often compromised when team performance breaks down under pressure (Hughes et al., 2016). Understanding how teamwork influences safety and performance requires viewing team performance as an ongoing process shaped by teamwork and taskwork (Campbell, 1990; Fernández Castillo et al., 2023; Rosen et al., 2013). Teamwork refers to the execution of interpersonal processes that either facilitate or hinder overall team functioning and outcomes (Rosen et al., 2013). In contrast, taskwork relates to the individual efforts and technical activities performed by each member to achieve specific team goals (Rosen et al., 2013).

This positions team performance as an important linking mechanism that connects team members' actions to broader safety outcomes. If teamwork is considered the catalyst for safety, then accurately measuring it becomes essential for identifying strengths, gaps, and opportunities for improvement. Traditionally, teamwork has been assessed using self-report and observational methods, which provide valuable insights into team processes and behaviors (Rosen et al., 2013). However, these approaches may be limited in their ability to capture the dynamic, moment-to-moment nature of team interactions as they unfold in real time (Rosen et al., 2013). As such, advancing the science of teamwork may require complementing traditional self-report and observational methods with objective and real-time measures that can capture the dynamic nature of team interactions as they unfold.

At the same time, measurement is not a “one-stop shop,” as no single approach can adequately capture the complexity of team performance (Salas et al., 2017). As teams operate across multiple levels (e.g., individual, team, organization) and evolve over time, there is no universal measurement approach that can capture the ever-changing life cycle of teams (Rosen et al., 2012, as cited in Salas et al., 2017). Relying on a single method can increase the risk of bias, making it important to draw on multiple sources of data to build a more complete picture of performance. Within this context, physiological measures present a promising complementary approach, as physiological data can provide objective indicators of interaction patterns and behavioral dynamics, often with minimal disruption to the team’s natural workflow (Salas et al., 2017). *Physiological measures* are tools that provide objective data on bodily functions (e.g., heart rate, skin conductance, and eye-tracking) and may offer a valuable means of capturing team dynamics (American Psychological Association, n.d.).

Rather than replacing self-report or observational measures, this perspective supports the continued evolution of how team performance is conceptualized and measured. However, the application of physiological measures to team performance remains underdeveloped, with limited guidance on how these measures should be selected, implemented, and interpreted in team contexts. To address this gap, this work provides a practical roadmap for researchers and practitioners seeking to use physiological measures to assess and diagnose teamwork. More broadly, it seeks to address the following question: *What are the emerging principles for using physiological measures to assess and diagnose team performance?* To examine this question, we adopt an interrogative approach that organizes the purpose, context, personnel, timing, measurement selection, and implementation of using physiological measures to assess team performance by considering: (1) why physiological measures should be used, (2) where such measurements should be made, (3) who should be measured and who should interpret the data, (4) when physiological indicators should be measured, (5) what content needs to be captured by the measurement system, and (6) how can physiological measures be used.

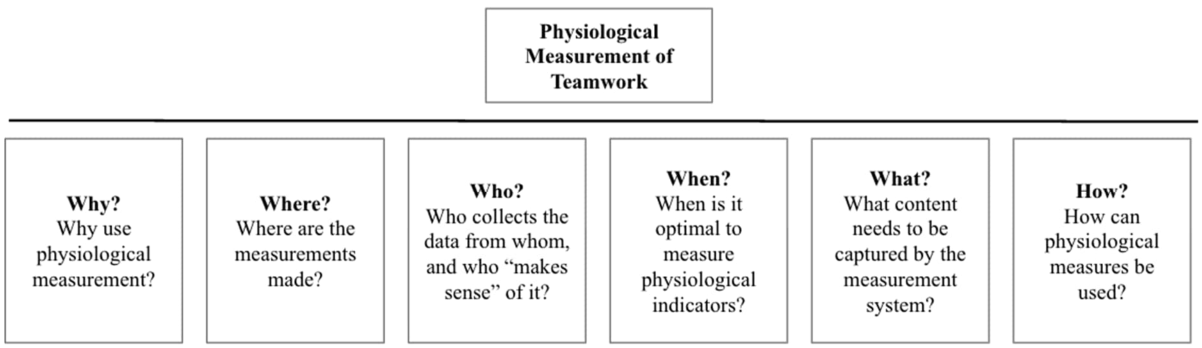
Collectively, these questions form an interrogative approach that structures the assessment of team performance across the dimensions of why, where, who, when, what, and how, with an emphasis on the role of physiological measures. Through this approach, we ground our exploration in the practical relevance and theoretical insights of team performance. At the same time, important methodological and interpretive challenges remain, particularly in translating individual-level physiological signals into meaningful team-level insights (for more information, see the *Limitations* section). To organize this discussion, we first present the Method section, followed by the emerging principles, and conclude with limitations, ethics, and broader research and practical implications.

Method

The *interrogative approach* used in the current work is rooted in previous work on team performance measurement (e.g., Rosen et al., 2013; Wildman et al., 2011; Figure 1). As this approach offers a structured framework for examining the purpose, context, content, timing, and application of performance measurement, it serves as a useful guide for the present discussion of physiological measures in team contexts. In the current work, the interrogative approach is adapted to physiological measurement. As such, the *why* of performance measurement addresses the reasoning for obtaining physiological data, whereas the *where* outlines where physiological measurements can be made. *Who* highlights the participants and those responsible for interpreting the data, and *when* describes when the measurements should be collected. Finally, *what* describes the content captured by the measurement system, and *how* determines how the information gathered from these measurements should be used. The dimensions of the interrogative approach (Figure 1) provide a valuable framework for examining the application of physiological measures to team performance.

Figure 1

Interrogative Approach of Measuring and Assessing Teamwork



Note. Adapted from Rosen et al., 2013; Wildman et al., 2011

Building on this framework, the present work extends the review by Kazi and colleagues (2021), who synthesized empirical studies employing physiological measures and situated them within an existing framework of teamwork (i.e., the Input-Mediator-Output-Input model). Their review indicated a notable gap in the literature, identifying only 32 studies that have examined physiological data in task-performing teams, highlighting the limited integration of such measures within team performance research. The work by Kazi and colleagues (2021) provides a valuable catalog of the existing empirical literature while offering a retrospective account of how physiological measures have been applied in team settings. The present work extends and integrates this foundational knowledge toward the prospective selection, implementation, and interpretation of physiological measures for team performance assessment.

A growing body of research has also demonstrated the predictive potential of physiological data for assessing team performance through data-driven and machine learning approaches (e.g., Li et al., 2025; Zahmat Doost et al., 2025). While these studies demonstrate what physiological measures can do, particularly in the context of performance prediction and classification, the present work complements this line of research by addressing how such measures can be selected and applied in team contexts. This guidance is reflected in a set of emerging principles. The term “emerging” is used to reflect the evolving nature of this area of research and the intent to provide flexible, evidence-informed guidance rather than prescriptive rules. As such, the emerging principles are designed to serve as concise, practical recommendations, providing researchers and practitioners with actionable direction for selecting and using physiological measures to inform team performance measurement.

To synthesize existing evidence, a literature search was conducted using a structured approach informed by prior review methodologies in team physiological research. First, an initial pool of relevant studies was identified through backward reference screening of Kazi and colleagues (2019), who synthesized empirical work on physiological measurement in team contexts. Second, this initial set of studies was expanded through a structured keyword-based search conducted in Google Scholar. For each physiological modality (i.e., HRV, EEG, EDA, fMRI, fNIRS, breathing, eye-tracking/pupillometry, and endocrine/hormonal measures), Boolean search strings were developed using the following logic: (1) team*, and (2) modality-specific terms (e.g., “EEG,” “electrodermal activity”). Third, the search process was supplemented by the authors’ domain knowledge in neuroscience and human factors.

Following the literature search, the interrogative framework adapted from prior team performance research and physiological measurement research (i.e., Kazi et al., 2021; Rosen et al., 2013; Wildman et al., 2011) served as the organizing structure for synthesizing the literature. The reviewed literature, including the prior synthesis and the additional modality-specific studies identified through the structured search, was used to inform the physiological considerations associated with each interrogative dimension (i.e., why, what, who, where, how, and when). This synthesis also informed Table 2 and the development of the emerging principles presented throughout this work.

As mentioned, the 11 emerging principles are grounded in and extend existing guidance on team performance measurement. Specifically, they draw on established recommendations (e.g., event-based assessment, triangulation across multiple data sources, and the need to account for multilevel and temporal dynamics in team processes). As such, they serve to organize and clarify prior best practices by situating them within an interrogative approach. Simultaneously, the present work introduces additional considerations that have been underdeveloped in the literature, particularly regarding the use of physiological measures in team contexts (e.g., the importance of establishing meaningful baselines for interpreting physiological indicators). Thus, these emerging principles consolidate and structure existing measurement guidance while extending it to address the unique methodological and interpretive challenges associated with using physiological data to measure and assess team performance.

Emerging Principles Informed by the Interrogative Approach

WHY: Why use physiological measurement?

Having a clear and explicitly articulated reason for why team performance measurement is needed ensures that measurement efforts are purpose-driven rather than arbitrary (Rosen et al., 2013). This helps align the selection of methods with specific goals, prioritize which aspects of teamwork are important to assess, and provides a framework for managing the inherent tradeoffs among different measurement approaches (Rosen et al., 2013). Without a well-defined “why,” physiological data collection can be unfocused, inefficient, or misaligned with practical or theoretical objectives, making it difficult to link physiological indicators to underlying cognitive processes (Wildman et al., 2010).

The motivation for collecting team performance measurement data varies and may include generating new research insights, providing feedback to teams, developing team training or other interventions, evaluating performance, and planning for the future (Rosen et al., 2013). Physiological measurement, in particular, offers unique advantages by providing objective, real-time data that can enhance these efforts. For instance, in high-risk environments such as healthcare, aviation, and the military, physiological data may provide early indicators of changes in workload, coordination, and stress that impact team safety and performance (Dias et al., 2019; Orsanu et al., 2005; Wilson et al., 2007).

A needs analysis can help determine whether moment-to-moment data are necessary to address specific team performance goals, and what resources are required for implementation (Salas, 2015). This process can help clarify the “why” for using physiological measurement and identify which aspects of team performance it is best suited to capture. Additionally, a needs analysis can help determine why physiological data may provide meaningful insight beyond traditional outcome metrics, such as error rates and task completion times.

Identifying what should be measured is only one part of the challenge. Differences in task structure, measurement procedures, and standards for analysis and reporting make it challenging to compare findings across studies, even when they target the same construct (Kazi et al., 2021). This sheds light on the importance of consistent approaches that improve the reliability and comparability of findings. Equally important is construct clarity. Construct clarity ensures that the data collected accurately reflect the targeted teamwork process. Without it, measurement tool selection, data interpretation, and alignment with theory are all compromised. When integrating physiological measures, behavioral annotations can be superimposed onto the data to mark the occurrence of specific team events (Sharma et al., 2019; Yang et al., 2019), helping bridge behavioral observation with underlying constructs.

Emerging Principle 1: Conduct a needs assessment to determine whether physiological data are necessary, feasible, and relevant.

Emerging Principle 2: Clearly define the purpose of collecting moment-to-moment physiological measurements and identify the specific team constructs the tool(s) are intended to capture.

WHAT: What content needs to be captured by the measurement system?

In addition to grounding physiological content in theory, triangulation with other data sources can help strengthen validity (Rosen et al., 2013). *Triangulation* refers to capturing different aspects of performance from different vantage points; in other words, using multiple measurement approaches to assess team performance (Campbell & Fiske, 1959; Edmondson & McManus, 2007; Rosen et al., 2008). Triangulation is important when capturing physiological data, as researchers cannot make assumptions about physiological data in isolation; rather, it is best to capture multiple types of data (e.g., behavioral observations, objective mission metrics) in conjunction with physiological measures to form a more complete snapshot of a team's performance (Fowlkes et al., 1998, as cited in Rosen et al., 2013). That said, although it is possible to collect data from a multitude of different sources, it is not feasible to collect *all* potential measures in every study (Salas et al., 2017). Therefore, researchers and practitioners should prioritize triangulation when feasible, considering the resources available to them, while selecting complementary measures that best align with the specific team processes and performance outcomes of interest.

Construct clarity extends beyond defining team processes to ensuring precision in the terminology used to measure them. In the context of physiological assessment, terms such as *physiological measures*, *biological markers,* and *biometrics* are often used interchangeably, despite representing distinct concepts in the literature. To support consistent application in team research, Table 1 provides definitions of these commonly used terms.

Table 1

Clarifying Constructs: Definitions of Physiological Measures and Other Related Terms

Construct	Definition	Example Measure	Citation
Physiological measure	"Physiological measures: any of a set of instruments that convey precise information about an individual's bodily functions…Studies using these tools typically obtain measurements before and after the introduction of a stimulus condition as a way to document an individual's response to that stimulus."	"Heart rate, skin conductance, skin temperature, cortisol level, palmar sweat, and eye tracking."	American Psychological Association. (n.d.)
Biological marker (Biomarker)	"A biological marker, or biomarker, is defined as a measurable indicator of some biological state or condition… The National Institutes of Health defines a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention"	"Blood, urine, or tissues to examine normal biological processes, pathogenic processes, or responses to therapeutic interventions."	Strimbu, K., & Tavel, J. A. (2010)
Biometric	"Biometrics refers to the automated recognition of individuals based on their biological and behavioral	"Individual's unique physical and behavioral traits, such as fingerprints,	Jain, A. K., & Kumar, A. (2012)

	characteristics.”	facial features, iris patterns, or voice, used primarily for identification and access control.”

Determining what content should be captured by the measurement system also requires identifying how many team behaviors can be meaningfully observed and measured within a given assessment period. For physiological measures to be interpretable, they must be anchored to specific behavioral events. Without this grounding, it is difficult to know what a physiological response reflects. To adequately capture team performance, assessments should include multiple task-relevant (otherwise called “event-based”) events that allow specific teamwork behaviors (e.g., leadership, coordination) to be observed across team members, ensuring that all members have sufficient opportunities to demonstrate the desired behaviors (Smith-Jentsch et al., 1998). Research recommends providing two or three opportunities for members to exhibit the desired teamwork competencies within a given session (Smith-Jentsch et al., 1998). In contrast, physiological measures may require hundreds of trials to achieve an adequate signal-to-noise ratio for reliable detection. Similarly, studies have suggested that even subject matter experts (SMEs) cannot discriminate between more than five team performance behaviors at one time (Salas et al., 2017; Smith-Jentsch et al., 1998). Focusing on three to five team performance behaviors is sufficient, as evaluating more than this often results in redundancy, because more than five behaviors tend to exhibit high correlations with one another (Salas et al., 2017; Smith-Jentsch et al., 1998).

To ensure validity and mitigate the limitations of any single measurement tool, researchers and practitioners should attempt to collect both quantitative and qualitative measures of team performance (Salas et al., 2017). This is especially important when using physiological measures, which capture underlying states but cannot, on their own, reveal the behavioral or cognitive processes driving them. Triangulating physiological data with self-report, behavioral observation, or performance outcomes strengthens the basis for drawing valid inferences about team processes and performance.

Emerging Principle 3: The content captured by the measurement system should be supplemented with additional metrics of team performance (i.e., triangulation).

Emerging Principle 4: Ensure that your subject matter experts measure only three to five team performance behaviors during a performance window, as measuring more than five behaviors can reduce each construct's potential to provide distinct and actionable information due to their potential correlation.

WHO: Who collects the data from whom, and who "makes sense" of it?

Determining the best individuals for data collection, observation, and analysis is a critical consideration when designing a measurement system (Rosen et al., 2013). Collecting physiological data tends to require some form of training to promote improved data quality and measurement precision. For example, collecting electroencephalography (EEG) data typically demands a higher level of technical expertise and specialized knowledge compared to more straightforward methods such as heart rate monitoring due to the complexity of the equipment.

Beyond data collection, researchers have emphasized the importance of using SMEs throughout the measurement process (Smith-Jentsch et al., 1998). SME input is particularly valuable when deciding which teamwork constructs and behaviors are observable and should be measured or rated (Smith-Jentsch et al., 1998). These decisions are best made *a priori* to ensure that raters have a clear understanding of what exactly they are assessing and what desired responses look like for each team member (Fowlkes et al., 1998). Critically, SMEs should have sufficient familiarity with the operational context in which the team is performing (Rosen et al., 2013). For instance, a rater assessing surgical team performance should understand basic surgical terms, procedures, and the operational context in order to make informed judgments about observed behaviors (Rosen et al., 2013).

When SMEs are unavailable to observe and rate team performance, research has demonstrated that individuals can be trained to do so effectively (Fowlkes et al., 1999). Regardless of who collects the data, raters should not be fellow team members, as research has documented a general inability for individuals to accurately self-monitor their own performance, as well as the potential for bias when rating fellow team members (Dunning et al., 2004; Holzbach, 1978; Thornton, 1980, as cited in Harris & Schaubroeck, 1988). This limitation extends to self-report measures, which may be influenced by recall bias, social desirability, and limited awareness of one's own team behaviors (Rosen et al., 2013). Although physiological measures can provide a more objective and complementary source of information, these data are not inherently self-explanatory and still require interpretation by individuals with sufficient expertise in teamwork, taskwork, and the operational context.

Emerging Principle 5*:* Ensure that subject matter experts possess a thorough understanding of relevant teamwork behaviors, taskwork, and the operational context before measuring and assessing physiological measures.

Emerging Principle 6*:* When triangulating, minimize the use of self-assessments and peer evaluations to reduce the risk of bias in team performance assessments when feasible.

WHERE: Where are the measurements made? ?

As with other measurement approaches, team physiological measurements can be taken in two major contexts: the laboratory and the field. Laboratory settings are common, as they provide many benefits over field studies, such as greater experimental control, the ability to constantly monitor participants, and access to measurement modalities that may be less feasible to collect in the field or operational environments (e.g., MRI or EEG measurements; Kusserow et al., 2012, as cited in Smets, De Raedt, & Van Hoof, 2018). However, laboratory settings also introduce limitations and challenges, such as the inability to generalize results to real-world environments (i.e., decreased ecological validity; Gerin et al., 1998; Wilhelm & Grossman, 2010). For example, data from pilots performing a task in a flight simulator may not accurately generalize to pilots performing the same task during live flight operations (Hays et al., 1992).

Moreover, studies conducted in the laboratory may present different physiological results due to psychological factors such as the "white coat effect," a phenomenon where individuals exhibit elevated physiological responses (e.g., increased heart rate) when being observed (Gerin et al., 1998; Wilhelm & Grossman, 2010). The limited generalizability of laboratory studies represents one reason researchers have begun to move away from laboratory data collection toward more accurate, real-world operational settings for physiological data collection (Smets, De Raedt, & Van Hoof, 2018). Collecting physiological data in the field presents a more generalizable and *in-situ* understanding of team performance (Cascio & Aguinis, 2008, as cited in Ganster, Crain, & Brossoit, 2018; Wilhelm & Grossman, 2010).

In-situ team performance research allows for a more comprehensive understanding of how dynamic, real-world factors influence a wide variety of factors related to team performance at the physiological level (Cascio & Aguinis, 2008). However, as with laboratory settings, physiological measurement in the field is not without its limitations, as researchers are often limited by the nature of the task the team is performing, as well as other factors such as motion artifacts and sensor interference that can lower the signal-to-noise ratio (Smets, De Raedt, & Van

Hoof, 2018). These challenges become problematic when the task involves significant physical movement or takes place in environments where certain measurement techniques, such as MRI or wired EEG, are either impractical, unsafe, or disruptive to team performance (Bullock et al., 2021). As each type of environment presents its own unique strengths and limitations, it is critical that the environment be selected based on whether it best suits the task and type of measure being collected.

Regardless of the environment, unobtrusiveness of the measurement should be prioritized to avoid potential performance effects (Kazdin, 1979). Webb and colleagues (1966) define unobtrusive measures as those that can be collected without needing the participant's active involvement and without influencing the behavior being measured. This is important in team research, as measurement systems that alter how individuals move, communicate, or interact may compromise the validity of the data by capturing reactions to the measurement system rather than the natural team processes of interest (Webb et al., 1966). Certain measurement techniques, such as MRI, involve heavy machinery, equipment, or complex wiring that may interfere with participants' ability to complete tasks as they normally would. Importantly, obtrusiveness is evaluated on a case-by-case basis, as the obtrusiveness of a measurement system can be subjective (Knight, 2018). For instance, what one individual may perceive as a non-interfering method of data collection, another may view as intrusive or distracting.

Emerging Principle 7: Select the environment for physiological measurement based on the specific task participants are performing to support the generalizability of the results.

Emerging Principle 8: When possible, employ unobtrusive measures, keeping in mind that what is considered obtrusive can vary depending on individual perception.

HOW: How can physiological measures be used?

Simply collecting physiological data from individuals and averaging them does not automatically yield valid team-level insights (Tesluk et al., 1997). To meaningfully use physiological measures for evaluating teams, two overarching considerations need to be taken into account: theoretical justification and statistical support. To aggregate individual-level data to infer team-level traits, there must be a theoretical justification for doing so (Nijstad et al., 2025; Tesluk et al., 1997). In addition to theoretical support, there must be statistical evidence to validate the aggregation (Tesluk et al., 1997).

Before drawing team-level inferences, it is important to establish that individual physiological responses reflect meaningful patterns of alignment within the team rather than isolated individual variation. Depending on the team construct of interest, these patterns may be represented through aggregation, synchrony, coupling, or other relational metrics that preserve the interdependent nature of team performance (Kazi et al., 2021; Stevens & Galloway, 2014). Simply averaging physiological data across team members may obscure meaningful differences in how individuals respond, adapt, or coordinate during task performance (Nijstad et al., 2025; Tesluk et al., 1997). Many physiological signals, such as heart rate variability, exhibit nonlinear, fractal-like patterns over time (Sammito et al., 2024). Traditional linear analysis techniques may overlook critical aspects of this complexity. Therefore, appropriate nonlinear and dynamic methods may be needed to capture the intricate temporal relationships that emerge in the data (Müller et al., 2017).

Finally, Table 2 provides selected physiological measurement methods, including a description of each method, how it is analyzed, the teamwork variables that have been examined with respect to each method's strengths and weaknesses.

Emerging Principle 9: Select analysis techniques that account for the complex, nonlinear nature of physiological signals rather than relying solely on traditional linear methods.

364 **Table 2**

365 *Selected Physiological Methods for Teamwork Research*

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
Heart Rate Variability (HRV)	HRV refers to the variation in time intervals between successive heartbeats that reflects autonomic nervous system activity (Shaffer & Ginsberg, 2017)	Time domain: SDNN, RMSSD, and related indices (Peabody et al., 2023; Shaffer & Ginsberg, 2017) Frequency domain: Low-frequency to high-frequency, spectral power measures, etc. (Peabody et al., 2023; Montague et al., 2014; Shaffer & Ginsberg, 2017) Non-linear: Entropy, Poincaré plots, and related complexity measures (Hemakom et al., 2017; Shaffer & Ginsberg, 2017)	HRV has been linked to team cohesion, mutual interdependence, higher comprehension, and coordination (Fusaroli et al., 2016, as cited in Kazi et al., 2021; Järvelä et al., 2014, as cited in Kazi et al., 2021; Mitkidis et al., 2015, as cited in Kazi et al., 2021)	Strengths: Noninvasive, cost-effective, and scalable for real-time monitoring of multiple participants; sensitive to changes in autonomic regulation and affective states (McCraty, 2017) Weaknesses: Sensitive to physical movement, circadian rhythms, and other individual differences, which may introduce variability in interpretation (Henning et al., 2001)
Electroencephalography (EEG)	EEG records electrical activity of the brain via scalp electrodes (Stevens et al., 2012; Algumaei et al., 2023)	Time domain: Less commonly applied in team-based EEG studies and often used for event-related potential (ERP) analyses	Team synchrony, workload, communication, expertise, competition, team task demands, objective task performance, and team	Strengths: High temporal resolution, real-time neural tracking, and well-suited for examining dynamic cognitive and attentional

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
		(Kazi et al., 2021) Frequency domain: Band power across beta, alpha, theta, and delta frequency bands (Stevens et al., 2012)	roles (Algumaei et al., 2023; Gorman et al., 2016, as cited in Kazi et al., 2021; Stevens et al., 2012, as cited in Kazi et al., 2021; Stevens et al., 2014)	processes (Algumaei et al., 2023) Weaknesses: Sensitive to physiological artifacts such as movement and nonphysiological sources of electrical noise; requires complex processing, offers limited portability in field settings, and may require larger sample sizes to make meaningful inferences (Algumaei et al., 2023; Gorman et al., 2016).
Electrodermal Activity (EDA)	EDA reflects sympathetic nervous system activity through skin conductance, including both tonic (SCL) and phasic (SCR) components. In this context, EDA encompasses related measures such as galvanic skin response (GSR) and is recorded using electrodes placed on the palms or wrists (Johannessen et al., 2020;	Time domain: Amplitude and frequency of phasic SCR signals (Johannessen et al., 2020) Frequency domain: Rarely applied due to low-frequency dynamics of EDA (Johannessen et al., 2020) Non-linear: Occasionally used in multimodal	Workload, task performance-related responses, induced emotion (positive versus negative), and competitiveness (Ahonen et al., 2018, as cited in Kazi et al., 2021; Montague et al., 2014, as cited in Kazi et al., 2021; Sariñana-González et al., 2017, as cited in Kazi et al., 2021)	Strengths: Real-time monitoring of arousal, noninvasive, and relatively low-cost (Ahonen et al., 2018; Montague et al., 2014) Weaknesses: Sensitive to external temperature and movement; slow return to baseline may reduce temporal precision during rapidly changing tasks

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
	Kazi et al., 2021; Van Der Mee et al., 2021).	modeling of arousal variability (Guastello et al., 2016; Neubauer et al., 2020)		(Christopoulos et al., 2019)
Functional Magnetic Resonance Imaging (fMRI)	fMRI detects brain activity by measuring changes in blood oxygenation (blood-oxygen-level-dependent [BOLD] signal), enabling the assessment of neural responses during collaborative tasks (Algumaei et al., 2023; Xie et al., 2020)	Time domain: BOLD signal amplitude and timing during task phases (Kerr et al., 2020) Frequency domain: Intersubject correlation (ISC) for neural coupling (Xie et al., 2020) Non-linear: Not typically applied due to fMRI's low temporal resolution Other: Region-of-interest (ROI) analysis; functional connectivity modeling; triadic inter-brain synchrony (Algumaei et al., 2023; Kazi et al., 2021)	Team cooperation, shared cognition, mutual prediction, empathy, and group performance during realistic tasks, joint attention, and decision-making (Algumaei et al., 2023; Xie et al., 2020)	Strengths: High spatial resolution; reveals task-specific neural correlates of collaboration (Kazi et al., 2021; Kerr et al., 2020). Weaknesses: Low temporal resolution (unless combined with EEG), high cost, and constrained movement which may limit ecological validity (Algumaei et al., 2023).
Functional Near-Infrared Spectroscopy	fNIRS measures cortical activity by detecting hemodynamic responses	Time domain: Hemodynamic response amplitudes and time-to-	Team coordination, joint attention, engagement, shared cognitive load, and	Strengths: Portable, noninvasive, and suitable for real-time team

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
(fNIRS)	(changes in oxygenated and deoxygenated hemoglobin concentrations) during task engagement. It enables real-time monitoring of brain activation related to team interaction and shared mental states (Algumaei et al., 2023)	peak for task-related segments (Algumaei et al., 2023) Frequency domain: Coherence and phase-locking value (PLV) analyses (Algumaei et al., 2023) Non-linear: Not typically applied; most analyses focus on statistical and signal-based synchronization (Algumaei et al., 2023) Other: Region-of-interest (ROI) mapping and hyperscanning across dyads and triads (Algumaei et al., 2023)	group decision-making accuracy (Algumaei et al., 2023)	interaction studies (Algumaei et al., 2023). Weaknesses: Limited spatial resolution compared to fMRI; signal quality can be affected by scalp and skin variations (Algumaei et al., 2023)
Respiration	Respiratory synchronization refers to the alignment of breathing patterns between team members during cooperative tasks. It is	Time domain: Breath-by-breath signal comparison and waveform alignment (Hemakom et al., 2017) Frequency domain:	Team coordination, physiological alignment, and task performance in group settings such as surgical teams and choir ensembles (Hemakom et	Strengths: Captures real-time physiological alignment; noninvasive and applicable in naturalistic settings (Hemakom et al., 2017)

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
	typically measured using respiratory belts or sensors and analyzed alongside HRV (Hemakom et al., 2017; Kazi et al., 2019).	Synchrony assessed through coherence and spectral overlap (Hemakom et al., 2017)	al., 2017; Kazi et al., 2019)	Weaknesses: Requires precise sensor placement and complex signal processing; may require specialized hardware and analytical software (Hemakom et al., 2017)
Eye-Tracking/ Pupillometry	Eye-tracking or pupillometry measures gaze behavior and pupil dilation to assess visual attention, emotional arousal, and coordination in teams. Metrics include fixation duration, gaze synchrony, and pupil size variability during joint tasks (Chanel et al., 2012; He et al., 2021; Järvelä et al., 2014; Kazi et al., 2019).	Time domain: Fixation sequences, and time spent on areas of interest (AOIs) to assess attention and visual processing (Punde et al., 2017) Spatial domain: Gaze point distribution, scanpaths, and AOI transitions to map visual navigation across stimuli (Punde et al., 2017)	Workload, visual attention, and gaze synchrony during collaboration (Montague et al., 2014, as cited in Kazi et al., 2021)	Strengths: High temporal precision for monitoring gaze behavior and visual attention; noninvasive and suitable for real-time team interaction studies (Montague et al., 2014) Weaknesses: Sensitive to confounding factors (e.g., pupil size changes in response to lighting, fatigue, workload, gaze angle), making isolation of specific cognitive processes more difficult; sensitive to signal noise (e.g., motion, occlusion, and head movement) and often requires recalibration

Method	Measures (i.e., description of the method)	How is it analyzed	Related Teamwork Variable Studied	Strengths/Weakness
				between participants (Harezlak & Kasprowski, 2018; Holmqvist, Nyström, & Mulvey, 2012; Ryan, Duchowski, & Birchfield, 2008; Tolvanen et al., 2022)
Endocrine/ Hormonal (Cortisol via Saliva or Blood)	Hormones serve as biochemical signals throughout the body. Cortisol is a steroid hormone released in response to stress, typically measured via saliva or blood sampling. It reflects activation of the hypothalamic–pituitary–adrenal (HPA) axis and is used to assess chronic stress in team contexts (Kazi et al., 2019; Hellhammer et al., 2009; Yılmazer, 2024).	Time domain: Time-sensitive; highest salivary cortisol levels typically occur in the early morning and decline throughout the day, necessitating time-controlled sampling (Papacosta & Nassis, 2011)	Salivary cortisol and alpha-amylase may be jointly used to assess stress response, team coping, and recovery (Papacosta & Nassis, 2011)	Strengths: Noninvasive and well-tolerated sampling method, especially suited for children and older adults (Papacosta & Nassis, 2011) Weaknesses: Cotton swabs may alter saliva acidity and biomarker concentrations, reducing data reliability (Papacosta & Nassis, 2011)

Note. The “related teamwork studied” column was initially populated using the “significant findings – unequivocal” column from Kazi et al. (2019), which includes seminal works on physiological measures in team contexts. To expand this list, additional papers were selected from the fields of medicine, engineering, and psychology.

WHEN: When is it optimal to measure physiological indicators?

Physiological measurement can offer many strengths, such as capturing real-time and dynamic changes in response to events as they unfold. However, the interpretation of these real-time data often depends on baseline measurements that provide a meaningful point of comparison (Fishel et al., 2007). These baselines can serve as a reference, anchoring the data and allowing deviations from typical functioning to be identified, thereby improving the sensitivity and interpretability of the data (Fishel et al., 2007). Therefore, baseline data should be collected under controlled and standardized conditions, given the sensitivity of physiological signals to numerous biological, environmental, and behavioral factors (Fishel et al., 2007). For instance, heart rate variability (HRV) is sensitive to a range of confounding variables such as age, disease, or lifestyle habits (e.g., caffeine or alcohol consumption) (Sammito et al., 2024). Depending on the analytic approach, event-related HRV may rely on baseline windows on the order of 30–60 seconds to capture the beat-to-beat fluctuations that define this measure (Sammito et al., 2024). In contrast, event-related neural measures may require baseline periods on the order of milliseconds (Stark & Squire, 2001).

Baseline requirements similarly vary across other physiological modalities and should align with the temporal characteristics of the signal being measured. For example, electrodermal and respiratory measures often require a brief resting period to establish a stable tonic reference before task onset, as both signals can fluctuate with spontaneous autonomic and respiratory activity (Lie et al., 2022; Oostveen et al., 2013). Neuroimaging modalities such as fMRI and fNIRS often use a brief resting or fixation period before task onset to establish a comparison point for interpreting changes in brain activity (Stark & Squire, 2001). Eye-tracking and pupillometry also require an initial calibration process so that changes in gaze behavior or pupil size can be interpreted relative to an individual's normal visual state (Krafka et al., 2016). Table 3 provides a summary of recommended experimental conditions, contraindicated contexts, and temporal characteristics for various physiological measures.

Additionally, the type of physiological measure being implemented can influence the timing of data collection (i.e., sample rate). Continuous metrics, such as heart rate, require ongoing, high-frequency sampling to accurately track changes over time (Burma et al., 2021). In contrast, cortisol exhibits delayed responses (Becker & Rohleder, 2019). Research shows that cortisol levels typically peak 20–30 minutes after the onset of a stressor, making immediate sampling ineffective (Becker & Rohleder, 2019). Additionally, cortisol levels are affected by other factors such as time of day(Liening et al., 2010). Thus, understanding the temporal characteristics of the specific physiological indicators being collected is essential for determining the appropriate timing of data collection windows.

There are additional contingencies when it comes to the timing of collecting continuous data, particularly when collecting multiple physiological measures together. This includes timestamping to allow for accurate synchronization since sensors for different physiological measures may operate at different sampling rates (Sharma et al., 2019; Vick & Ikehara, 2003). For instance, eye-tracking devices may sample at 60 Hz, while skin conductance sensors may sample at 13 Hz (Khazaei & Faghih, 2024; Vick & Ikehara, 2003). Without alignment, aliasing can occur, where events that did not occur at the same time can become falsely aligned. A method to address this issue of aliasing is anchoring, which involves using the device with the fastest sampling rate as a reference clock by aligning all other data streams to its timeline using the faster sampling rate device (e.g., an eye-tracker) as the master clock and aligning slower sampling devices to the nearest corresponding timestamp (Vick & Ikehara, 2003).

Additionally, different modalities require varying levels of event-related time synchronization to extract meaning. Neural measures, such as EEG, can be measured as time and phase-locked responses to an event or the presentation of a stimulus (Stevens et al., 2012). These so-called event-related potentials can require second to millisecond resolution (Stevens et al., 2012). Similar time resolution is required when measuring neural functional connectivity metrics between two or more members of a team. In contrast, measures of heart rate or heart rate variability recorded in response to an event or between team members require a time resolution on the order of seconds or tens of seconds (Dias et al., 2019).

Finally, the measurement needs to be practical for the environment in which it is being conducted (Rosen et al., 2013; Salas et al., 2017). For example, during a long surgical procedure, overly complex, or intrusive sensor setups may interfere with the surgical team's performance or become unsustainable over long periods of time. As such, the optimal timing for physiological measurement requires a balance between theoretical rigor (i.e., accurate and high-resolution data) and practical feasibility, including setup time, baseline acquisition, and the physical and cognitive burden placed on participants.

Emerging Principle 10: Collect baseline physiological data prior to task performance using modality-appropriate baseline procedures, while accounting for individual differences and other factors that may influence physiological signals.

Emerging Principle 11: When using multiple physiological measures, use synchronized, high-resolution timestamps to collect data at different rates, helping construct a comprehensive and temporally accurate performance profile.

436 **Table 3**

437 *Recommended Experimental Conditions for Physiological Measures in Team Research*

Measure	Primary Construct(s)	Example Experimental Conditions	Contraindicated Conditions	Temporal Resolution
HRV	Autonomic regulation, sustained stress, inter-member physiological synchrony (Curtin et al., 2022; Kaur et al., 2014; Von Borell et al., 2007)	Extended cooperative tasks, workload transitions, team coordination over minutes to hours; ≥5-min recordings for frequency-domain HRV (Delliaux et al., 2019; Sammito & Böckelmann, 2016; Sariñana-González, Romero-Martinez, & Moya-Albiol, 2018)	Very short tasks (<60 s); use as the sole measure for emotional valence discrimination (Töyry et al., 1995)	Milliseconds to seconds (derived HRV metrics); beat-to-beat (R-R interval detection) (Ruha, Sallinen, & Nissilä, 1997; Zimmermann et al., 1991)
EEG	Cognitive workload, attentional allocation, inter-brain synchrony, and event-related potentials (ERPs) (Beres, 2017; Chikhi,	Stationary tasks: command-and-control, co-located team decision-making, and hyperscanning paradigms for shared	High-mobility field tasks; environments with gross motor movement or excessive electrical noise (Cutmore & James, 1999)	Sub-milliseconds to millisecond resolution; seconds for spectral power analyses (Davidson, Jones, &

Measure	Primary Construct(s)	Example Experimental Conditions	Contraindicated Conditions	Temporal Resolution
	Matton, & Blanchet, 2022; Deng et al., 2023;Ladouce et al., 2019)	attention (Curtin et al., 2022; Hang et al., 2026; Micek et al., 2026; Shedeed, Issa, & El-Sayed, 2013)		Peiris, 2007; Shamas et al., 2015)
EDA/GSR	Sympathetic arousal, emotional reactivity, and acute stress (Amiri et al., 2016; Greco et al., 2021; Posada-Quintero et al., 2016)	Discrete event-driven tasks: alarms, critical handoffs, threat cues, and high-stakes team decisions (Studer, Scheibehenne, & Clark, 2016; Zhou et al., 2019)	Sustained low-arousal vigilance or high ambient temperature or fluctuationing humidity; high physical exertion (Posada-Quintero et al., 2016)	Seconds (phasic SCR); minutes (tonic SCL; Posada-Quintero & Chon, 2019; Subramanian, Barbieri, & Brown, 2020)
fMRI	Hemodynamic alterations in relation to neural activity; localization of brain activation; measurement of non-conscious processing; blood flow mapping (Cohen, Noll, & Schneider, 1993; Matthews et al., 1999; Reimann et al., 2011)	Controlled laboratory tasks involving cooperative problem-solving, joint decision-making, or shared attention paradigms (Glover, 2012; Martin et al., 2015; Reimann et al., 2011)	Tasks requiring significant movement; tasks requiring significant face-to-face communication; naturalistic face-to-face teamwork; designs requiring sub-second temporal resolution (Glover, 2012; Havsteen et al., 2017)	Seconds (for hemodynamic response in fMRI) (Kim, Richter, & Uğurbil, 2005; Matthews & Jezzard, 2004)
fNIRS	Prefrontal executive function, cognitive workload, hemodynamic coupling (Aghajani, Garvey, & Omurtag, 2017; Byun et al., 2014;	Mobile tasks requiring cortical monitoring: flight simulation, walking protocols, face-to-face team interactions; field	Designs requiring deep subcortical or whole-brain coverage; rapid event-related responses (<1s; Klein, 2024)	Seconds (hemodynamic response) (Wilcox & Biondi, 2015)

Measure	Primary Construct(s)	Example Experimental Conditions	Contraindicated Conditions	Temporal Resolution
	Keles, Barbour, & Omurtag, 2016; Xie et al., 2021)	settings unsuitable for fMRI (Cui et al., 2011; Gateau et al., 2015; Hirsch et al., 2021; Holtzer et al., 2011)		
Respiration	Cognitive load, stress, physical effort, autonomic co-regulation (Grassmann et al., 2016; Hemakom et al., 2017)	Operational environments with chest-band sensors; complementary measure for HRV interpretation (e.g., respiratory sinus arrhythmia [RSA]); extended monitoring periods (Hemakom et al., 2017)	Use as a standalone modality for discriminating cognitive vs. emotional states; tasks involving frequent speech or vocalization (may confound respiratory rate) (Boyadzhieva & Kayhan, 2021; Grassmann et al., 2016)	Seconds (breath-by-breath respiratory cycles) (Grassmann et al., 2016)
Eye-tracking/ Pupillometry	Attention allocation, shared situation awareness, visual search strategies, and cognitive workload (Atweh & Riggs, 2025; Chita-Tegmark, 2016; Kita et al., 2010; Wetzel et al., 1997; Willbanks, Aroke, & Dudding, 2021)	Simulation-based training, display design, team visual coordination, and performance monitoring during dynamic task environments (Lai et al., 2013; Ma et al., 2024; Peißl, Wickens, & Baruah, 2018; Weiss et al., 2023)	Tasks requiring significant movement, tasks under uncontrolled lighting conditions; studies requiring unambiguous mapping of gaze behavior to specific cognitive processes (Deng & Gao, 2022; Evans et al., 2012)	Approximately 31-1000 Hz (Gibaldi et al., 2016; Prasse et al., 2020)
Salivary Cortisol	Cumulative stress, HPA-axis activation, pre/post	Pre- and post-task assessments, extended	Studies requiring moment-to-moment	Minutes to hours (peak salivary responses

Measure	Primary Construct(s)	Example Experimental Conditions	Contraindicated Conditions	Temporal Resolution
	stressor assessment, team readiness (Hellhammer et al., 2009; Papacosta & Nassis, 2011; Moreira et al., 2009)	operations, longitudinal team stress tracking; controlled diurnal sampling windows (Papacosta & Nassis, 2011; Moreira et al., 2009)	dynamics; tasks requiring continuous data streams; uncontrolled diurnal or dietary confounds (Papacosta & Nassis, 2011)	approximately 20–30 minutes post-stressor) (Moreira et al., 2009)

Note. Contraindicated conditions refer to research contexts in which a physiological measure may be less suitable for use.

Discussion

Physiological measurements may offer an unparalleled advantage: the ability to capture subtle, moment-to-moment changes that reflect underlying cognitive and affective states (Kazi et al., 2021). Although the use of physiological measures has become increasingly accessible and theoretically relevant across organizational research (Ganster, Crain, & Brossoit, 2018), their application to understanding the dynamic processes that unfold during teamwork remains relatively underutilized. This discussion examines the key limitations and challenges associated with physiological measurement in team research, as well as the ethical implications and future potential of these measures in both research and applied contexts. The emerging principles discussed in this work are summarized in Table 4, which provides a consolidated set of practical recommendations for selecting, implementing, and interpreting physiological measures in team performance contexts.

Table 4

Summary of Emerging Principles for Using Physiological Measures to Assess Teamwork

Why use physiological measurement?
Emerging Principle 1: Conduct a needs assessment to determine whether physiological data are necessary, feasible, and relevant. ***Emerging Principle 2:*** Clearly define the purpose of collecting moment-to-moment physiological measurements and identify the specific team constructs the tool(s) are intended to capture.
What content needs to be captured by the measurement system?
Emerging Principle 3*:* The content captured by the measurement system should be supplemented with additional metrics of team performance (i.e., triangulation). ***Emerging Principle 4****:* Ensure that your subject matter experts measure only three to five constructs during a performance window, as measuring more than five constructs can reduce

each construct's potential to provide distinct and actionable information due to their potential correlation.
Who collects the data from whom, and who "makes sense" of it?
Emerging Principle 5*:* Ensure that subject matter experts possess a thorough understanding of relevant teamwork behaviors, taskwork, and the operational context before measuring and assessing physiological measures. ***Emerging Principle 6****:* When triangulating, minimize the use of self-assessments and peer evaluations to reduce the risk of bias in team performance assessments when feasible.
Where are the measurements made?
Emerging Principle 7*:* Select the environment for physiological measurement based on the specific task participants are performing to support the generalizability of the results. ***Emerging Principle 8****:* When possible, employ unobtrusive measures, keeping in mind that what is considered obtrusive can vary depending on individual perception.
How can physiological measures be used?
Emerging Principle 9: Select analysis techniques that account for the complex, nonlinear nature of physiological signals rather than relying solely on traditional linear methods.
When is it optimal to measure physiological indicators?
Emerging Principle 10: Collect baseline physiological data prior to task performance using modality-appropriate baseline procedures, while accounting for individual differences and other factors that may influence physiological signals. ***Emerging Principle 11:*** When using multiple physiological measures, use synchronized, high-resolution timestamps to collect data at different rates, helping construct a comprehensive and temporally accurate performance profile.

Limitations

All measurement approaches have limitations, and physiological measurement is no exception. For instance, ensuring that physiological measures are conceptually linked to true team-level constructs and not just individual experiences is a challenging endeavor. Teams are composed of members with different roles, responsibilities, and task demands, which means they may not be exposed to identical stimuli or react to similar situations in similar ways. Even when physiological responses across team members appear similar (e.g., elevated heart rate), this does not necessarily provide insight into how effectively the team is coordinating, communicating, or performing. Therefore, caution is warranted when interpreting physiological indicators or aggregating individual physiological responses into team-level insights.

Additionally, while often perceived as objective, physiological signals are influenced by individual characteristics and contextual variables (Cosoli et al., 2023). Many physiological findings are inconsistently reported and standard test protocols are often lacking (Cosoli et al., 2023). The absence of methodological rigor makes it challenging to compare results across devices or applications, especially those that are not established as research-grade devices. Moreover, physiological measurements are affected by natural biological variability that cannot be eliminated. For instance, many physiological measures are sensitive to chronic and acute health conditions, lifestyle factors, age, weight, and environmental stressors (Sammito et al., 2024). Without controlling for these confounding variables, there is a risk of attributing physiological changes to incorrect causes. To improve construct validity, physiological data should be triangulated with psychometrically-validated self-report measures, behavioral observations, objective mission metrics, or contextual information. The challenge then becomes: Which data do we trust more, and which do we use as the “gold standard”?

Ethical and privacy concerns are also important to consider. Cosoli and colleagues (2023) highlight the importance of informed consent and ethical review in all experimental protocols to ensure transparency and accountability in the collection and use of physiological data. Such concerns are particularly important in applied settings, where the distinction between supportive monitoring and surveillance must be clearly defined and upheld. Without boundaries and safeguards, the use of physiological data intended to promote team performance and well-being may be misused in ways that blur the line between support and surveillance.

Research and Practical Implications

The integration of physiological measurement into team science presents an opportunity for theoretical advancement by improving our understanding of team performance from a physiological perspective, which has been largely underdeveloped to date. This shift also promotes a move toward multimodal team assessment, aligning with best practices advocated by team science researchers (e.g., Cooke, 2015; Kazi et al., 2021; Rosen et al., 2013; Salas et al., 2017). Additionally, this work takes an interdisciplinary approach, blending insights from psychology and physiology, to create a more holistic and nuanced understanding of team performance.

In terms of practical implications, the use of physiological data in applied settings holds promise for enhancing diagnostic precision and supporting more adaptive, personalized training and interventions. By identifying physiological markers that correlate with specific states or patterns of team interaction, practitioners may hypothetically be able to tailor interventions in real-time. When physiological signals are combined with validated behavioral and self-report measures, they may help reduce subjective bias in performance evaluations and increase the accuracy of assessments. However, the value of such tools is not in replacing traditional methods, but in complementing them, ensuring a more comprehensive approach to evaluating and developing team performance.

Conclusion

Team performance is complex, and physiological signals are not a panacea for measuring teamwork. Physiological measures do not offer a precise equation that yields definitive answers;

rather, they serve as guideposts that provide a perspective on team performance and complement the rich body of existing work in team measurement. This perspective is especially valuable in high-risk settings, where safety hinges not only on procedures, but on the team members working together effectively. The use of physiological measures in team research is a modern continuation of Wundt’s (1904) call for capturing dynamic physiological processes and connecting them to meaningful psychological constructs. Understanding team performance is not about relying on standalone indicators, but about integrating those elements into a theoretically-grounded and already scientifically-rich understanding of teamwork.

References

Aghajani, H., Garbey, M., & Omurtag, A. (2017). Measuring mental workload with EEG+ fNIRS. *Frontiers in Human Neuroscience*, *11*, 359. https://doi.org/10.3389/fnhum.2017.00359

Al-Ayyad, M., Owida, H. A., De Fazio, R., Al-Naami, B., & Visconti, P. (2023). Electromyography monitoring systems in rehabilitation: A review of clinical applications, wearable devices and signal acquisition methodologies. *Electronics*, *12*(7), 1520.

Algumaei, M., Hettiarachchi, I. T., Farghaly, M., & Bhatti, A. (2023). The neuroscience of team dynamics: Exploring neurophysiological measures for assessing team performance. *IEEE Access*, *11*, 129173–129194.

Amiri, A. M., Abtahi, M., Rabasco, A., Armey, M., & Mankodiya, K. (2016, March). Emotional reactivity monitoring using electrodermal activity analysis in individuals with suicidal behaviors. In *2016 10th International Symposium on Medical Information and Communication Technology (ISMICT)* (pp. 1-5). IEEE. https://doi.org/10.1109/ISMICT.2016.7498896

American Psychological Association. (n.d.). *Physiological measure.* In APA Dictionary of Psychology. Retrieved April 16, 2025, from https://dictionary.apa.org/physiological-measure

Atweh, J., & Riggs, S. (2025, September). Can We See Team Situation Awareness? Predicting and Classifying Team SA Using Eye Tracking and Machine Learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 69, No. 1, pp. 198-206). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/10711813251357922

Barnes, C. M., & Hollenbeck, J. R. (2009). Sleep deprivation and decision-making teams: Burning the midnight oil or playing with fire? *Academy of Management Review*, *34*(1), 56–66.

Becker, L., & Rohleder, N. (2019). Time course of the physiological stress response to an acute stressor and its associations with the primacy and recency effect of the serial position curve. *PloS one*, *14*(5), e0213883. https://doi.org/10.1371/journal.pone.0213883

Beres, A. M. (2017). Time is of the essence: A review of electroencephalography (EEG) and event-related brain potentials (ERPs) in language research. *Applied Psychophysiology and Biofeedback*, *42*(4), 247-255. https://doi.org/10.1007/s10484-017-9371-3

Buller, M. J., Welles, A. P., & Friedl, K. E. (2018). Wearable physiological monitoring for human thermal-work strain optimization. *Journal of applied physiology (Bethesda, Md. : 1985)*, *124*(2), 432–441. https://doi.org/10.1152/japplphysiol.00353.2017

Bullock, M., Jackson, G. D., & Abbott, D. F. (2021). Artifact Reduction in Simultaneous EEG-fMRI: A Systematic Review of Methods and Contemporary Usage. *Frontiers in neurology*, *12*, 622719. https://doi.org/10.3389/fneur.2021.622719

Burma, J. S., Lapointe, A. P., Soroush, A., Oni, I. K., Smirl, J. D., & Dunn, J. F. (2021). Insufficient sampling frequencies skew heart rate variability estimates: Implications for extracting heart rate metrics from neuroimaging and physiological data. *Journal of biomedical informatics*, *123*, 103934. https://doi.org/10.1016/j.jbi.2021.103934

Boyadzhieva, A., & Kayhan, E. (2021). Keeping the Breath in Mind: Respiration, Neural Oscillations, and the Free Energy Principle. *Frontiers in neuroscience*, *15*, 647579. https://doi.org/10.3389/fnins.2021.647579

Byun, K., Hyodo, K., Suwabe, K., Ochi, G., Sakairi, Y., Kato, M., ... & Soya, H. (2014). Positive effect of acute mild exercise on executive function via arousal-related prefrontal activations: an fNIRS study. *Neuroimage*, *98*, 336-345. https://doi.org/10.1016/j.neuroimage.2014.04.067

Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2nd ed., pp. 687–732). Consulting Psychologists Press.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81–105.

Cascio, W. F., & Aguinis, H. (2008). Staffing twenty-first-century organizations. *Academy of Management Annals*, *2*(1), 133-165. https://doi.org/10.1080/19416520802211461

Chanel, G., Kivikangas, J. M., & Ravaja, N. (2012). Physiological compliance for social gaming analysis: Cooperative versus competitive play. *Interacting with Computers*, *24*, 306–316. https://doi.org/10.1016/j.intcom.2012.04.012

Chikhi, S., Matton, N., & Blanchet, S. (2022). EEG power spectral measures of cognitive workload: A meta-analysis. *Psychophysiology*, *59*(6), e14009. https://doi.org/10.1111/psyp.14009

Chita-Tegmark, M. (2016). Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders*, *3*(3), 209-223. https://doi.org/10.1007/s40489-016-0077-x

Christopoulos, G. I., Uy, M. A., & Yap, W. J. (2019). The body and the brain: Measuring skin conductance responses to understand the emotional experience. *Organizational Research Methods, 22*(1), 394–420. https://doi.org/10.1177/1094428116681073

Cohen, J. D., Noll, D. C., & Schneider, W. (1993). Functional magnetic resonance imaging: Overview and methods for psychological research. *Behavior Research Methods, Instruments, & Computers*, *25*(2), 101-113. https://doi.org/10.3758/BF03204483

Cooke, N. J. (2015). Team cognition as interaction. *Current directions in psychological science*, *24*(6), 415-419. https://doi.org/10.1177/0963721415602474

Cosoli, G., Antognoli, L., & Scalise, L. (2023). Methods for the metrological characterization of wearable devices for the measurement of physiological signals: state of the art and future challenges. *MethodsX*, *10*, 102038. https://doi.org/10.1016/j.mex.2023.102038

Cui, X., Bray, S., Bryant, D. M., Glover, G. H., & Reiss, A. L. (2011). A quantitative comparison of NIRS and fMRI across multiple cognitive tasks. *Neuroimage*, *54*(4), 2808-2821. https://doi.org/10.1016/j.neuroimage.2010.10.069

Curtin, A. B., Watson, J., Topoglu, Y., DeFilippis, N., Ye, H., Suri, R., & Ayaz, H. (2022, January). Brains in sync: Team coordination and interpersonal prefrontal neural synchrony during cooperative e-gaming. In *NATO Human Factors & Medicine (HFM) Symposium*.

Cutmore, T. R., & James, D. A. (1999). Identifying and reducing noise in psychophysiological recordings. *International Journal of Psychophysiology*, *32*(2), 129-150. https://doi.org/10.1016/S0167-8760(99)00014-8

Dabbabi, T., Bouafif, L., & Cherif, A. (2023, April). A review of non invasive methods of brain activity measurements via EEG signals analysis. In *2023 IEEE International Conference* *on Advanced Systems and Emergent Technologies (IC_ASET)*(pp. 01-06). IEEE. https://doi.org/10.1109/IC_ASET58101.2023.10150607

Davidson, P. R., Jones, R. D., & Peiris, M. T. (2007). EEG-based lapse detection with high temporal resolution. *IEEE Transactions on Biomedical Engineering*, *54*(5), 832-839. https://doi.org/10.1109/TBME.2007.893452

Delliaux, S., Delaforge, A., Deharo, J. C., & Chaumet, G. (2019). Mental workload alters heart rate variability, lowering non-linear dynamics. *Frontiers in Physiology*, *10*, 565. https://doi.org/10.3389/fphys.2019.00565

Deng, R., & Gao, Y. (2023). A review of eye tracking research on video-based learning. *Education and Information Technologies, 28*(6), 7671–7702. https://doi.org/10.1007/s10639-022-11486-7

Deng, X., Yang, M., Chen, X., & Zhan, Y. (2023). The role of mindfulness on theta inter-brain synchrony during cooperation feedback processing: An EEG-based hyperscanning study. *International Journal of Clinical and Health Psychology*, *23*(4), 100396. https://doi.org/10.1016/j.ijchp.2023.100396

Di Lascio, E., Gashi, S., & Santini, S. (2018). Unobtrusive assessment of students' emotional engagement during lectures using electrodermal activity sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(3), 1-21. https://doi.org/10.1145/3264913

Dias, R. D., Zenati, M. A., Stevens, R., Gabany, J. M., & Yule, S. J. (2019). Physiological synchronization and entropy as measures of team cognitive load. *Journal of biomedical informatics*, *96*, 103250. https://doi.org/10.1016/j.jbi.2019.103250

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological science in the public interest*, *5*(3), 69-106.

Edmondson, A. C., & McManus, S. E. (2007). Methodological fit in management field research. *Academy of Management Review, 32*, 1155–1179. https://doi.org/10.2307/20159361

Elkins, A. N., Muth, E. R., Hoover, A. W., Walker, A. D., Carpenter, T. L., & Switzer, F. S. (2009). Physiological compliance and team performance. *Applied ergonomics*, *40*(6), 997-1003. https://doi.org/10.1016/j.apergo.2009.02.002

Evans, K. M., Tarduno, J. A., Jacobs, R. A., & Pelz, J. B. (2012). Collecting and analyzing eye-tracking data in outdoor environments. *Journal of Eye Movement Research*, *5*(2), 11. https://doi.org/10.16910/jemr.5.2.6

Farina, D., Merletti, R., & Enoka, R. M. (2004). The extraction of neural strategies from the surface EMG. *Journal of applied physiology (Bethesda, Md. : 1985)*, *96*(4), 1486–1495. https://doi.org/10.1152/japplphysiol.01070.2003

Feitosa, J., Grossman, R., Kramer, W. S., & Salas, E. (2020). Measuring team trust: A critical and meta-analytical review. *Journal of Organizational Behavior, 41*(5), 479–501. https://doi.org/10.1002/job.2436

Fernandez, G. C., Khalid, M., & Salas, E. (2023). Team Performance. *Oxford Bibliographies in Management.* Ricky Griffin (Ed.). New York: Oxford University Press.

Fishel, S. R., Muth, E. R., & Hoover, A. W. (2007). Establishing appropriate physiological baseline procedures for real-time physiological measurement. *Journal of Cognitive Engineering and Decision Making*, *1*(3), 286-308. https://doi.org/10.1518/155534307X255636

Fowlkes, J., Dwyer, D. J., Oser, R. L., & Salas, E. (1998). Event-based approach to training (EBAT). The international journal of aviation psychology, 8(3), 209-221. https://doi.org/10.1207/s15327108ijap0803_3

Friesen, C. L., Lawrence, M., Ingram, T. G., Smith, M. M., Hamilton, E. A., Holland, C. W., ... & Boe, S. G. (2022). Portable wireless and fibreless fNIRS headband compares favorably to a stationary headcap-based system. *PLoS One*, *17*(7), e0269654. https://doi.org/10.1371/journal.pone.0269654

Fusaroli, R., Bjørndahl, J. S., Roepstorff, A., & Tylén, K. (2016). A heart for interaction: Shared physiological dynamics and behavioral coordination in a collective, creative construction task. *Journal of Experimental Psychology: Human Perception and Performance, 42,* 1297–1310. doi:10.1037/xhp0000207

Ganster, D. C., Crain, T. L., & Brossoit, R. M. (2018). Physiological measurement in the organizational sciences: A review and recommendations for future use. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*(1), 267-293. https://doi.org/10.1146/annurev-orgpsych-032117-104613

Gateau, T., Durantin, G., Lancelot, F., Scannella, S., & Dehais, F. (2015). Real-time state estimation in a flight simulator using fNIRS. *PloS One*, *10*(3), e0121279. https://doi.org/10.1371/journal.pone.0121279

Gerin, W., Christenfeld, N., Pieper, C., Derafael, D. A., Su, O., Stroessner, S. J., ... & Pickering, T. G. (1998). The generalizability of cardiovascular responses across settings. *Journal of Psychosomatic Research, 44*(2), 209-218.

Gibaldi, A., Vanegas, M., Bex, P. J., & Maiello, G. (2017). Evaluation of the Tobii EyeX Eye tracking controller and Matlab toolkit for research. *Behavior Research Methods*, *49*(3), 923-946. https://doi.org/10.3758/s13428-016-0762-9

Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, *22*(2), 133. https://doi.org/10.1016/j.nec.2010.11.001

Gorman, J. C., Martin, M. J., Dunbar, T. A., Stevens, R. H., Galloway, T. L., Amazeen, P. G., & Likens, A. D. (2016). Cross-level effects between neurophysiology and communication during team training. *Human Factors, 58,* 181–199.doi:10.1177/0018720815602575

Grassmann, M., Vlemincx, E., von Leupoldt, A., Mittelstädt, J. M., & Van den Bergh, O. (2016). Respiratory Changes in Response to Cognitive Load: A Systematic Review. *Neural plasticity*, *2016*, 8146809. https://doi.org/10.1155/2016/8146809

Greco, A., Valenza, G., Lazaro, J., Garzon-Rey, J. M., Aguilo, J., de la Cámara, C., ... & Scilingo, E. P. (2021). Acute stress state classification based on electrodermal activity modeling. *IEEE Transactions on Affective Computing*, *14*(1), 788-799. https://doi.org/10.1109/TAFFC.2021.3055294

Guastello, S. J., Marra, D. E., Perna, C., Castro, J., Gomez, M., & Peressini, A. F. (2016). Physiological synchronization in emergency response teams: Subjective workload, drivers and empaths. *Nonlinear Dynamics, Psychology, and Life Sciences*, *20*, 223–270.

Hang, Y., Wu, W., Li, R., Huang, L., Shioiri, S., & He, X. (2026). Trust Emerges From Shared Attention: Behavioural and Neural Evidence From Dual EEG Hyperscanning. *Human Brain Mapping*, *47*(4), e70504. https://doi.org/10.1002/hbm.70504

Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel psychology*, *41*(1), 43-62. https://doi.org/10.1111/j.1744-6570.1988.tb00631.x

Havsteen, I., Ohlhues, A., Madsen, K. H., Nybing, J. D., Christensen, H., & Christensen, A. (2017). Are movement artifacts in magnetic resonance imaging a real problem?—a narrative review. *Frontiers in Neurology*, *8*, 232. https://doi.org/10.3389/fneur.2017.00232

Hays, R. T., Jacobs, J. W., Prince, C., & Salas, E. (1992). Flight simulator training effectiveness: A meta-analysis. *Military Psychology*, *4*(2), 63–74. https://doi.org/10.1207/s15327876mp0402_1

Harezlak, K., & Kasprowski, P. (2018). Application of eye tracking in medicine: A survey, research issues and challenges. *Computerized Medical Imaging and Graphics*, *65*, 176-190. https://doi.org/10.1016/j.compmedimag.2017.04.006

He, W., Jiang, X., & Zheng, B. (2021). Synchronization of pupil dilations correlates with team performance in a simulated laparoscopic team coordination task. *Simulation in Healthcare*, *16*(6), e206-e213.

Hellhammer, D. H., Wüst, S., & Kudielka, B. M. (2009). Salivary cortisol as a biomarker in stress research. *Psychoneuroendocrinology*, *34*(2), 163-171. https://doi.org/10.1016/j.psyneuen.2008.10.026

Hemakom, A., Powezka, K., Goverdovsky, V., Jaffer, U., & Mandic, D. P. (2017). Quantifying team cooperation through intrinsic multi-scale measures: respiratory and cardiac synchronization in choir singers and surgical teams. *Royal Society open science, 4*(12), 170853.

Henning, R. A., Armstead, A. G., & Ferris, J. K. (2009). Social psychophysiological compliance in a four-person research team. *Applied Ergonomics, 40*, 1004–1010. doi:10.1016/j .apergo.2009.04.009

Henning, R. A., Boucsein, W., & Claudia Gil, M. (2001). Social- physiological compliance as a determinant of team performance. *International Journal of Psychophysiology, 40*, 221–232. doi:10.1016/S0167-8760(00)00190-2

Hirsch, J., Tiede, M., Zhang, X., Noah, J. A., Salama-Manteau, A., & Biriotti, M. (2021). Interpersonal agreement and disagreement during face-to-face dialogue: an fNIRS investigation. *Frontiers in Human Neuroscience*, *14*, 606397. https://doi.org/10.3389/fnhum.2020.606397

Holmqvist, K., Nyström, M., & Mulvey, F. (2012, March). Eye tracker data quality: What it is and how to measure it. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (pp. 45-52). https://doi.org/10.1145/2168556.2168563

Holtzer, R., Mahoney, J. R., Izzetoglu, M., Izzetoglu, K., Onaral, B., & Verghese, J. (2011). fNIRS study of walking and walking while talking in young and old individuals. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, *66*(8), 879-887. https://doi.org/10.1093/gerona/glr068

Holzbach, R. L. (1978). Rater bias in performance ratings: superior, self-, and peer ratings. *Journal of applied psychology*, *63*(5), 579. https://psycnet.apa.org/doi/10.1037/0021-9010.63.5.579

Hughes, A. M., Gregory, M. E., Joseph, D. L., Sonesh, S. C., Marlow, S. L., Lacerenza, C. N., Benishek, L. E., King, H. B., & Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare. *The Journal of applied psychology*, *101*(9), 1266–1304. https://doi.org/10.1037/apl0000120

Jain, A. K., & Kumar, A. (2012). Biometric recognition: an overview. *Second generation biometrics: The ethical, legal and social context*, 49-79.

Järvelä, S., Kivikangas, J. M., Kätsyri, J., & Ravaja, N. (2014). Physiological linkage of dyadic gaming experience. *Simulation & Gaming*, *45*, 24–40. doi:10.1177/1046878113513080

Johannessen, E., Szulewski, A., Radulovic, N., White, M., Braund, H., Howes, D., ... & Davies, C. (2020). Psychophysiologic measures of cognitive load in physician team leaders during trauma resuscitation. *Computers in Human Behavior*, *111*, 106393.

Kaur, B., Durek, J. J., O'Kane, B. L., Tran, N., Moses, S., Luthra, M., & Ikonomidou, V. N. (2014, May). Heart rate variability (HRV): an indicator of stress. In *Independent Component Analyses, Compressive Sampling, Wavelets, Neural Net, Biosystems, and Nanoengineering XII* (Vol. 9118, pp. 177-184). SPIE. https://doi.org/10.1117/12.2051148

Kazdin, A. E. (1979). Unobtrusive measures in behavioral assessment. Journal of Applied *Behavior Analysis, 12*(4), 713-724. https://doi.org/10.1901/jaba.1979.12-713

Kazi, S., Khaleghzadegan, S., Dinh, J. V., Shelhamer, M. J., Sapirstein, A., Goeddel, L. A., Chime, N. O., Salas, E., & Rosen, M. A. (2021). Team Physiological Dynamics: A Critical Review. *Human factors*, *63*(1), 32–65. https://doi.org/10.1177/0018720819874160

Keles, H. O., Barbour, R. L., & Omurtag, A. (2016). Hemodynamic correlates of spontaneous neural activity measured by human whole-head resting state EEG+ fNIRS. *Neuroimage*, *138*, 76-87. https://doi.org/10.1016/j.neuroimage.2016.05.058

Kerr, K. L., Cosgrove, K. T., Ratliff, E. L., Burrows, K., Misaki, M., Moore, A. J., ... & Morris, A. S. (2020). TEAMwork: Testing emotional attunement and mutuality during parent-adolescent fMRI. *Frontiers in Human Neuroscience, 14*, 24.

Khazaei, S., & Faghih, R. T. (2024). Eye tracking is more sensitive than skin conductance response in detecting mild environmental stimuli. *PNAS nexus*, *3*(9), pgae370. https://doi.org/10.1093/pnasnexus/pgae370

Kim, S. G., Richter, W., & Uğurbil, K. (1997). Limitations of temporal resolution in functional MRI. *Magnetic Resonance in Medicine*, *37*(4), 631-636. https://doi.org/10.1002/mrm.1910370427

Kita, Y., Gunji, A., Sakihara, K., Inagaki, M., Kaga, M., Nakagawa, E., & Hosokawa, T. (2010). Scanning strategies do not modulate face identification: Eye-tracking and near-infrared spectroscopy study. *PloS One*, *5*(6), e11050. https://doi.org/10.1371/journal.pone.0011050

Klein, F. (2024). Optimizing spatial specificity and signal quality in fNIRS: an overview of potential challenges and possible options for improving the reliability of real-time applications. *Frontiers in Neuroergonomics*, *5*, 1286586. https://doi.org/10.3389/fnrgo.2024.1286586

Knight, A. P. (2018). Innovations in unobtrusive methods. In A. Bryman and D. A. Buchanan (Eds.). *Unconventional methodology in organization and management research*, *1*, 64-83.

Kozlowski, S. W. (2015). Advancing research on team process dynamics: Theoretical, methodological, and measurement considerations. *Organizational Psychology Review*, *5*(4), 270-299.

Kozlowski, S. W., & Chao, G. T. (2018). Unpacking team process dynamics and emergent phenomena: Challenges, conceptual advances, and innovative methods. *American Psychologist*, *73*(4), 576.

Krafka, K., Khosla, A., Kellnhofer, P., Kannan, H., Bhandarkar, S., Matusik, W., & Torralba, A. (2016). Eye tracking for everyone. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2176-2184).

Kusserow, M., Amft, O., & Tröster, G. (2012). Monitoring stress arousal in the wild. *IEEE Pervasive Computing*, *12*(2), 28-37.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., & Ietswaart, M. (2019). Mobile EEG identifies the re-allocation of attention during real-world activity. *Scientific Reports*, *9*(1), 15851. https://doi.org/10.1038/s41598-019-51996-y

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., ... & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90-115. https://doi.org/10.1016/j.edurev.2013.10.001

Leonard, M. W., & Frankel, A. S. (2011). Role of Effective Teamwork and Communication in Delivering Safe, High-Quality Care. *Mount Sinai Journal of Medicine, 78*(6), 820–826. https://doi.org/10.1002/msj.20295

Li, M., Rojas, E., Bracken, B., Cooke, N., Desrochers, P., Gorman, J., Huang, L., Kilcullen, M., Rosen, M., Winder, A., Yadav, A., Yin, X., Zahmat Doost, E., & Zhou, S. (2025). Physiological Synchrony as a Predictor of Team Performance: A Machine Learning Approach. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *69*(1), 333-337.

Li, S., Sung, B., Lin, Y., & Mitas, O. (2022). Electrodermal activity measure: A methodological review. *Annals of Tourism Research, 96*, 103460. https://doi.org/10.1016/j.annals.2022.103460

Liening, S. H., Stanton, S. J., Saini, E. K., & Schultheiss, O. C. (2010). Salivary testosterone, cortisol, and progesterone: two-week stability, interhormone correlations, and effects of time of day, menstrual cycle, and oral contraceptive use on steroid hormone levels. *Physiology & behavior, 99*(1), 8-16. https://doi.org/10.1016/j.physbeh.2009.10.001

Ma, S., Zhang, W., Noland, R. B., & Andrews, C. J. (2024). Eye tracking measures of bicyclists' behavior and perception: A systematic review. *Transportation Research Part F: Traffic Psychology and Behaviour, 107*, 52-68. https://doi.org/10.1016/j.trf.2024.08.026

Martin, A., Schurz, M., Kronbichler, M., & Richlan, F. (2015). Reading in the brain of children and adults: A meta-analysis of 40 functional magnetic resonance imaging studies. *Human Brain Mapping, 36*(5), 1963-1981. https://doi.org/10.1002/hbm.22749

Matthews, P. M., Clare, S., & Adcock, J. (1999). Functional magnetic resonance imaging: Clinical applications and potential. *Journal of inherited metabolic disease, 22*(4), 337-352. https://doi.org/10.1023/A:1005535102667

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry, 75*(1), 6-12.

McCraty, R. (2017). New frontiers in heart rate variability and social coherence research: Techniques, technologies, and implications for improving group dynamics and outcomes. *Frontiers in Public Health, 5*, 267.

Micek, C., Warnke, L., Rodrigues, L. A., Putze, F., & Solovey, E. (2026). Capturing Team Cognition: A Multimodal Dataset for Adaptive Collaborative Interfaces. In Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI ’26), April 13–17, 2026, Barcelona, Spain. ACM, New York, NY, USA, 24 pages. https://doi.org/10.1145/3772318.3791607

Mitkidis, P., McGraw, J. J., Roepstorff, A., & Wallot, S. (2015). Building trust: Heart rate synchrony and arousal during joint action increased by public goods game. *Physiology and Behavior*, *149*, 101–106. doi:10.1016/j.physbeh.2015.05.033

Montague, E., Xu, J., & Chiou, E. (2014). Shared experiences of technology and trust: An experimental study of physiological compliance between active and passive users in technology-mediated collaborative encounters. *IEEE Transactions on Human-Machine Systems, 44,* 614–624. doi:10.1109/ THMS.2014.2325859

Moreira, A., Arsati, F., de Oliveira Lima Arsati, Y. B., Da Silva, D. A., & de Araújo, V. C. (2009). Salivary cortisol in top-level professional soccer players. *European journal of applied physiology*, *106*(1), 25-30.

Müller, W., Jung, A., & Ahammer, H. (2017). Advantages and problems of nonlinear methods applied to analyze physiological time signals: human balance control as an example. *Scientific reports*, *7*(1), 1-11.

Neubauer, C., Woolley, J., Khooshabeh, P., & Scherer, S. (2016, October). Getting to know you: a multimodal investigation of team behavior and resilience to stress. *In Proceedings of the 18th ACM International Conference on Multimodal Interaction* (pp. 193-200).

Nijstad, B. A., Homan, A. C., Heerdink, M. W., & van Kleef, G. A. (2025). Data aggregation in team research: Theoretical considerations and practical recommendations. *Organizational Psychology Review*, 20413866251333058.

Oostveen, E., Boda, K., van der Grinten, C. P., James, A. L., Young, S., Nieland, H., & Hantos,

Z. (2013). Respiratory impedance in healthy subjects: baseline values and bronchodilator response. *European respiratory journal*, *42*(6), 1513-1523.

Orasanu, J., Tada, Y., Kraft, N., & Fischer, U. (2005, May). Physiological monitoring of team and task stressors. In *Biomonitoring for Physiological and Cognitive Performance during Military Operations* (Vol. 5797, pp. 182-192). SPIE

Papacosta, E., & Nassis, G. P. (2011). Saliva as a tool for monitoring steroid, peptide and immune markers in sport and exercise science. *Journal of science and medicine in sport*, *14*(5), 424–434. https://doi.org/10.1016/j.jsams.2011.03.004

Peabody, J. E., Ryznar, R., Ziesmann, M. T., Gillman, L., Ryznar, R. J., & Gillman, L. M. (2023). A systematic review of heart rate variability as a measure of stress in medical professionals. *Cureus*, *15*(1).

Peißl, S., Wickens, C. D., & Baruah, R. (2018). Eye-tracking measures in aviation: A selective literature review. *The International Journal of Aerospace Psychology*, *28*(3-4), 98-112. https://doi.org/10.1080/24721840.2018.1514978

Posada-Quintero, H. F., & Chon, K. H. (2019, May). Phasic component of electrodermal activity is more correlated to brain activity than tonic component. In *2019 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)* (pp. 1-4). IEEE. https://doi.org/10.1109/BHI.2019.8834567

Posada-Quintero, H. F., Florian, J. P., Orjuela-Cañón, A. D., Aljama-Corrales, T., Charleston-Villalobos, S., & Chon, K. H. (2016). Power spectral density analysis of electrodermal activity for sympathetic function assessment. *Annals of Biomedical Engineering*, *44*(10), 3124-3135. https://doi.org/10.1007/s10439-016-1606-6

Prasse, P., Jäger, L. A., Makowski, S., Feuerpfeil, M., & Scheffer, T. (2020). On the relationship between eye tracking resolution and performance of oculomotoric biometric identification. *Procedia Computer Science, 176*, 2088-2097. https://doi.org/10.1016/j.procs.2020.09.245

Punde, P. A., Jadhav, M. E., & Manza, R. R. (2017, October). A study of eye tracking technology and its applications. In *2017 1st International Conference on Intelligent Systems and Information Management (ICISIM)* (pp. 86-90). IEEE.

Rampichini, S., Vieira, T. M., Castiglioni, P., & Merati, G. (2020). Complexity Analysis of Surface Electromyography for Assessing the Myoelectric Manifestation of Muscle Fatigue: A Review. *Entropy (Basel, Switzerland)*, *22*(5), 529. https://doi.org/10.3390/e22050529

Ryan, W. J., Duchowski, A. T., & Birchfield, S. T. (2008, March). Limbus/pupil switching for wearable eye tracking under variable lighting conditions. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 61-64).

Reimann, M., Schilke, O., Weber, B., Neuhaus, C., & Zaichkowsky, J. (2011). Functional magnetic resonance imaging in consumer research: A review and application. *Psychology & Marketing*, *28*(6), 608-637. https://doi.org/10.1002/mar.20403

Richer, R., Groh, B. H., Blank, P., Dorschky, E., Martindale, C., Klucken, J., & Eskofier, B. M. (2016, June). Unobtrusive real-time heart rate variability analysis for the detection of orthostatic dysregulation. In *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)* (pp. 189-193). IEEE. https://doi.org/10.1109/BSN.2016.7516257

Rosen, M. A., Schiebel, N., Salas, E., Wu, T. S., Silvestri, S., & King, H. B. (2013). How can team performance be measured, assessed, and diagnosed? In E. Salas, K. Frush, D. P. Baker, J. B. Battles, H. B. King, & R. L. Wears (Eds.), *Improving patient safety through teamwork and team training* (pp. 59–79). Oxford University Press.

Rosen, M. A., Salas, E., Silvestri, S., Wu, T. S., & Lazzara, E. H. (2008). A measurement tool for simulation-based training in emergency medicine: the simulation module for assessment of resident targeted event responses (SMARTER) approach. *Simulation in Healthcare*, *3*(3), 170-179. doi: 10.1097/SIH.0b013e318173038d

Rosen, M. A., Schiebel, N., Salas, E., Wu, T. S., Silvestri, S., & King, H. B. (2012). How can team performance be measured, assessed, and diagnosed. *Improving patient safety through teamwork and team training*, 59-79.

Roy, R. N., Verdière, K. J., & Dehais, F. (2020). EEG covariance-based estimation of cooperative states in teammates. In *Adaptive Instructional Systems: Second International Conference, AIS 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22* (pp. 383-393). Springer International Publishing.

Ruha, A., Sallinen, S., & Nissila, S. (2002). A real-time microprocessor QRS detector system with a 1-ms timing accuracy for the measurement of ambulatory HRV. *IEEE Transactions on Biomedical Engineering*, *44*(3), 159-167. https://doi.org/10.1109/10.554762

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

Salas, E., Stevens, R., Gorman, J., Cooke, N. J., Guastello, S., & Von Davier, A. A. (2015, September). What will quantitative measures of teamwork look like in 10 years?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 59, No. 1, pp. 235-239). Sage CA: Los Angeles, CA: SAGE Publications.

Salas, E., Bisbey, T. M., Traylor, A. M., & Rosen, M. A. (2020). Can teamwork promote safety in organizations? *Annual Review of Organizational Psychology and Organizational Behavior*, *7*, 283–313. https://doi.org/10.1146/annurev-orgpsych-012119-045411

Sammito, S., & Böckelmann, I. (2016). Reference values for time-and frequency-domain heart rate variability measures. *Heart Rhythm*, *13*(6), 1309-1316. https://doi.org/10.1016/j.hrthm.2016.02.006

Salas, E., Reyes, D. L., & Woods, A. L. (2017). The assessment of team performance: Observations and needs. In A. A. von Davier, M. Zhu, & P. C. Kyllonen (Eds.), *Innovative assessment of collaboration* (pp. 21–36). Springer International Publishing/Springer Nature. https://doi.org/10.1007/978-3-319-33261-1_2

Sammito, S., Thielmann, B., & Böckelmann, I. (2024). Update: factors influencing heart rate variability-a narrative review. *Frontiers in physiology*, *15*, 1430458. https://doi.org/10.3389/fphys.2024.1430458

Sariñana-González, P., Romero-Martínez, Á., & Moya-Albiol, L. (2018). Cooperation between strangers in face-to-face dyads produces more cardiovascular activation than competition or working alone. *Journal of Psychophysiology*. https://doi.org/10.1027/0269-8803/a000210

Shaffer, F., & Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, *5*, 258.

Shamas, M., Wendling, F., El Falou, W., & Hassan, M. (2015, April). EEGNET: A novel tool for processing and mapping EEG functional networks. In *2015 7th International IEEE/EMBS Conference on Neural Engineering (NER)* (pp. 1064-1067). IEEE. https://doi.org/10.1109/NER.2015.7146811

Sharma, K., Castellini, C., Van Den Broek, E. L., Albu-Schaeffer, A., & Schwenker, F. (2019). A dataset of continuous affect annotations and physiological signals for emotion analysis. *Scientific data*, *6*(1), 196.

Shedeed, H. A., Issa, M. F., & El-Sayed, S. M. (2013, November). Brain EEG signal processing for controlling a robotic arm. In *2013 8th International Conference on Computer Engineering & Systems (ICCES)* (pp. 152-157). IEEE. https://doi.org/10.1109/ICCES.2013.6707191

Smets, E., De Raedt, W., & Van Hoof, C. (2018). Into the wild: the challenges of physiological stress detection in laboratory and ambulatory settings. *IEEE journal of biomedical and health informatics*, *23*(2), 463-473.

Smith-Jentsch, K. A., Johnston, J. H., & Payne, S. C. (1998). Measuring team-related expertise in complex environments. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making Decisions Under Stress: Implications for Individual and Team Training* (1st ed., pp. 61-87). American Psychological Association.

Stark, C. E., & Squire, L. R. (2001). When zero is not zero: the problem of ambiguous baseline conditions in fMRI. *Proceedings of the National Academy of Sciences*, *98*(22), 12760-12766. https://doi.org/10.1073/pnas.221462998

Stevens, R. H., & Galloway, T. L. (2014). Toward a quantitative description of the neurodynamic organizations of teams. *Social Neuroscience, 9,* 160–173. doi:10.1080/17470919.2014.883 324

Stevens, R. H., Galloway, T., Berka, C., & Sprang, M. (2009). Can neurophysiologic synchronies provide a platform for adapting team performance?. In Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience: 5th International Conference, *FAC 2009 Held as Part of HCI International 2009 San Diego, CA, USA, July 19-24, 2009 Proceedings 5* (pp. 658-667). Springer Berlin Heidelberg.

Stevens, R. H., Galloway, T. L., Wang, P., & Berka, C. (2012). Cognitive neurophysiologic synchronies: What can they contribute to the study of teamwork? *Human Factors, 54,* 489–502. doi:10.1177/0018720811427296

Stikic, M., Berka, C., Waldman, D., Balthazard, P., Pless, N., & Maak, T. (2013). Neurophysiological estimation of team psychological metrics. In *Foundations of Augmented Cognition: 7th International Conference, AC 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013. Proceedings 7* (pp. 209-218). Springer Berlin Heidelberg.

Strimbu, K., & Tavel, J. A. (2010). What are biomarkers?. *Current Opinion in HIV and AIDS, 5*(6), 463-466.

Studer, B., Scheibehenne, B., & Clark, L. (2016). Psychophysiological arousal and inter-and intraindividual differences in risk-sensitive decision making. *Psychophysiology*, *53*(6), 940-950. https://doi.org/10.1111/psyp.12627

Subramanian, S., Barbieri, R., & Brown, E. N. (2020). Point process temporal structure characterizes electrodermal activity. *Proceedings of the National Academy of Sciences, 117*(42), 26422-26428. https://doi.org/10.1073/pnas.2004403117

Tesluk, P., Mathieu, J. E., Zaccaro, S. J., & Marks, M. (1997). Task and aggregation issues in the analysis and assessment of team performance. In *Team performance assessment and measurement* (pp. 209-236). Psychology Press.

Thornton III, G. C. (1980). Psychometric properties of self-appraisals of job performance. *Personnel psychology, 33*(2), 263-271. https://doi.org/10.1111/j.1744-6570.1980.tb02348.x

Toppi, J., Borghini, G., Petti, M., He, E. J., De Giusti, V., He, B., & Babiloni, F. (2016). Investigating cooperative behavior in ecological settings: An EEG hyperscanning study. *PLOS ONE, 11*(4), e0154236.

Tolvanen, O., Elomaa, A. P., Itkonen, M., Vrzakova, H., Bednarik, R., & Huotarinen, A. (2022). Eye-tracking indicators of workload in surgery: A systematic review. *Journal of Investigative Surgery*, *35*(6), 1340-1349. https://doi.org/10.1080/08941939.2021.2025282

Töyry, J., Mäntysaari, M., Hartikainen, J., & Länsimies, E. (1995). Day-to-day variability of cardiac autonomic regulation parameters in normal subjects. *Clinical Physiology*, *15*(1), 39-46. https://doi.org/10.1111/j.1475-097X.1995.tb00428.x

Van Der Mee, D. J., Gevonden, M. J., Westerink, J. H., & De Geus, E. J. C. (2021). Validity of electrodermal activity-based measures of sympathetic nervous system activity from a wrist-worn device. *International Journal of Psychophysiology*, *168*, 52-64.

Vick, R. M., & Ikehara, C. S. (2003, January). Methodological issues of real time data acquisition from multiple sources of physiological data. In *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the* (pp. 7-pp). IEEE.

Volk, S., Pearsall, M. J., Christian, M. S., & Becker, W. J. (2017). Chronotype diversity in teams: Toward a theory of team energetic asynchrony. *Academy of Management Review*, *42*(4), 683-702.

Von Borell, E., Langbein, J., Després, G., Hansen, S., Leterrier, C., Marchant, J., ... & Veissier, I. (2007). Heart rate variability as a measure of autonomic regulation of cardiac activity for assessing stress and welfare in farm animals—A review. *Physiology & Behavior*, *92*(3), 293-316. https://doi.org/10.1016/j.physbeh.2007.01.007

Ward, R. D., Marsden, P. H., Cahill, B., & Johnson, C. (2002, April). Physiological responses to well-designed and poorly designed interfaces. In *Proceedings of CHI 2002 workshop on physiological computing.*

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences.* Rand Mcnally.

Weiss, K. E., Kolbe, M., Lohmeyer, Q., & Meboldt, M. (2023). Measuring teamwork for training in healthcare using eye tracking and pose estimation. *Frontiers in Psychology*, *14*, 1169940. https://doi.org/10.3389/fpsyg.2023.1169940

Wetzel, P. A., Krueger-anderson, G., Poprik, C., & Bascom, P. (1997). *An Eye Tracking System for Analysis of Pilots' Scan Paths* (No. ALHRTR19960145).

Wilcox, T., & Biondi, M. (2015). fNIRS in the developmental sciences. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(3), 263-283. https://doi.org/10.1002/wcs.1343

Wildman, J. L., Bedwell, W. L., Salas, E., & Smith-Jentsch, K. A. (2011). Performance measurement at work: A multilevel perspective. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol. 1. Building and developing the organization* (pp. 303–341). American Psychological Association. https://doi.org/10.1037/12169-010

Wilhelm, F. H., & Grossman, P. (2010). Emotions beyond the laboratory: Theoretical fundamentals, study design, and analytic strategies for advanced ambulatory assessment. *Biological psychology*, *84*(3), 552-569. http://dx.doi.org/10.1016/j.biopsycho.2010.01.017

Wilbanks, B. A., Aroke, E., & Dudding, K. M. (2021). Using eye tracking for measuring cognitive workload during clinical simulations: literature review and synthesis. *CIN: Computers, Informatics, Nursing*, *39*(9), 499-507. https://doi.org/10.1097/CIN.0000000000000704

Wilson, G. F., Caldwell, J. A., & Russell, C. A. (2007). Performance and psychophysiological measures of fatigue effects on aviation related tasks of varying difficulty. *The international journal of aviation psychology*, *17*(2), 219-247.

Woolley, A.W., Hackman, J. R., Jerde, T. E., Chabris, C. F., Bennett, S. L., & Kosslyn, S. M. (2007). Using brain-based measures to compose teams: How individual capabilities and team collaboration strategies jointly shape performance. *Social Neuroscience, 2*(2), 96-105.

Wundt, W. M. (1904). *Principles of physiological psychology* (Vol. 1; E. B. Titchener, Trans.). Sonnenschein. (Original work published 1874).

Xie, H., Karipidis, I. I., Howell, A., Schreier, M., Sheau, K. E., Manchanda, M. K., ... & Saggar, M. (2020). Finding the neural correlates of collaboration using a three-person fMRI hyperscanning paradigm. *Proceedings of the National Academy of Sciences, 117*(37), 23066-23072.

Xie, S., Wu, D., Yang, J., Luo, J., Chang, C., & Li, H. (2021). An fNIRS examination of executive function in bilingual young children. *International Journal of Bilingualism*, *25*(3), 516-530. https://doi.org/10.1177/1367006920952881

Yang, H. C., & Lee, C. C. (2019, September). Annotation matters: A comprehensive study on recognizing intended, self-reported, and observed emotion labels using physiology. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (pp. 1-7). IEEE.

Zahmat Doost, E., Gorman, J. C., Yin, X., Zhou, S., Yadav, G. A., Cooke, N., Huang, L., Bracken, B. K., Winder, A., Desrochers, P., Li, M., Rojas, E., Rosen, M., & Kilcullen, M. (2025). BioTDMS: A Multi-Sensor and Data-driven System for Real-Time Team Performance Assessment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *69*(1), 1255-1259.

Zhou, J., Jia, X., Xu, G., Jia, J., Hai, R., Gao, C., & Zhang, S. (2019). The relationship between different types of alarm sounds and children's perceived risk based on their physiological responses. *International Journal of Environmental Research and Public Health*, *16*(24), 5091. https://doi.org/10.3390/ijerph16245091

Zimmermann, M., Adamec, R., Simonin, P., & Richez, J. (1991). Beat-to-beat detection of ventricular late potentials with high-resolution electrocardiography. *American Heart Journal*, *121*(2), 576-585. https://doi.org/10.1016/0002-8703(91)90728-Z