# COGNITION AND PHYSIOLOGICAL RESPONSE: TOWARDS A MODEL OF VALIDATED PHYSIOLOGICAL MEASUREMENT

Ashley M. Hughes, William G. Volante, Kimberly Stowers, Kevin Leyva, James M. Oglesby, Tiffany Bisbey, Eduardo Salas
University of Central Florida and Institute for Simulation and Training

Benjamin A. Knott, Michael A. Vidulich
U.S. Air Force Research Laboratory

Complex tasks in large and error-prone environments require unobtrusive, unbiased and real-time measurement of cognitive variables to promote safety and to achieve optimal performance. Despite the prevalence of physiological measurement of cognitive constructs and cognitive performance, such as workload, little has been done to justify the inference of cognitive states from physiological measures. We develop a framework based on the extant literature to provide the groundwork for further validation of physiological measurement. Specifically, we leverage theoretically-grounded conditions of measurement to aid in investigating the logical sampling and construct validity for use of such metrics. Further meta-analytic investigation is warranted to validate the model and justify use of physiological measures.

## INTRODUCTION

The transition to tasks from involving multiple human operators to humans interacting with automatic systems and virtual tools requires specific skills and abilities on behalf of the personnel operating them (Nemeth, 2004). With the constant advancement of technology in these automated systems, more complex tasks in dynamic and complex environments are able to be undertaken by workers and employees; thusly, the cognitive demands of the operator become increasingly important to understand. Similarly, as these tasks become exceedingly complex, they require a higher degree of cognitive focus and vigilance to ensure safety of the human within the system (Proctor & Zandt, 2008). New methods in science and technology must be implemented to improve measurement to allow for real-time measurement. These methods should enable an improved understanding of these cognitive demands, and in doing so, allow for the development of countermeasures to address the risks of potential performance decrement.

The physiological process the body undergoes while performing a task is a crucial aspect to this understanding. Neglecting the physiological context of behavior, computational models of human information processing and cognition are incomplete. Specifically, without accounting for physiological response, these models are incomplete and fail to provide adequate information to understand processes and outputs of cognitive performance (Robert & Hockey, 1997).

Physiological measurement of cognitive and psychological factors is in widespread use for both applied and basic research purposes. Reviews over the last decade have sought to evaluate the extant literature on physiological measurement of psychological constructs, such as vigilance. Physiological measurement covers a wide span of voluntary and involuntary responses (e.g., EKG, saccades, EEG, EMG, speech analysis) to infer a wide variety of cognitive performance variables (e.g., cognitive workload, fatigue, situation awareness).



Despite these extensive reviews of physiological measurement of key psychological constructs such as workload, situation awareness, vigilance, and fatigue (e.g., Kramer, 1990; Oken, Salinsky, & Elsas, 2006), cause for the inference of psychological attributes from physiological states (Cacioppo & Tassinary, 1990; Stevens et al., 2010) is yet to be quantified and standardized for many forms of physiological measurement. Given the wide range of interpretation of physiological responses (e.g., EKGs used to measure stress as well as workload) and lack of available items for traditional psychometric assessment (Allen & Yen, 2002), we seek to provide theoretical guidance to inform validated inference of physiological metrics of cognition. Specifically, we seek to provide theoretical guidance for construct and logical validity (see Allen & Yen, 2002) of use of such measures. Through synthesis of the extant literature, we present a comprehensive framework to provide insight to critical aspects of the task, environment, and individual involved in physiological measurement to better assess its validity. Thus, robust and valid measurement can be further leveraged to provide insights into human cognition, performance, and improved safety in complex and hazardous environments. Below, we lay the rationale for each condition. We also present the practical outcomes and implications of leveraging this framework.

## FRAMEWORK OVERVIEW

This framework conceptualizes conditions critical to understanding the theoretical basis for physiological metrics of cognitive performance through identifying links of study conditions to the cognitive variables of interest. Specifically, these conditions of interest are conceptualized as input, adaptation and output variables. By including these variables which affect cognitive performance, the construct and logical sampling validity for physiological metrics as measures of cognitive performance can be bolstered. Justification for use of these key variables is determined through careful consideration of theoretical and practical conditions which influence individual adaptation and, consequently, response (e.g., task load increasing cognitive workload). In examining influencing variables of cognition, the framework outlines input and adaptation variables which serve to improve or weaken variables of cognition (e.g., usability and situation awareness), or conditions which may reduce physiological measure effectiveness (e.g., extreme temperatures). Additionally this can be leveraged to ensure measurement strategies correlate and diverge appropriately. This can and has been established for particular physiological approaches to measurement of cognitive variables through correlation (e.g., Camili, Terenzi, & Di Nocera, 2007). However, with theoretical conditions, we can control and better determine that physiological measurement measures the intended constructs and does not measure constructs for which it was not intended, which may confound the measurement strategy (i.e., construct and discriminant validity; see Bobko, 2001). More specifically, the construct validity of physiological measures can be tested through determining whether they correlate with traditional measures of the intended construct and whether they do not correlate with constructs unrelated to the construct of interest.

The principle components of our model involve the input, including stimulus, adaptation of the individual (also known as the individual's response or preconditions of the operator) and output or behavioral response to stressors (i.e., "trinity of stress"; Hancock & Warm, 1989), including individual differences and physiological response. We further describe specifics within each input, adaptation, and output category in terms of the task, the environment, and the individual as conditions (Hancock & Warm, 1989). Conditions of this nature have been identified as significantly impacting cognitive performance (Hancock, Ross, & Szalma, 2007). Conversely, we seek to explain these theoretical elements within the framework as critical components to physiological measurement and the assessment of vigilance, situation awareness, cognitive workload, as well as fatigue. Ultimately, we aim to provide guidance on providing psychometrically sound methods for assessing and interpreting cognitive states from physiological measures.

## Input Variables/Stressors

The input variables play an important role in the framework by acting as preconditions to the situation or environmental context. Each of the input variables outlined must be recorded and accounted for in order to maintain construct validity. Specifically, we account for said variables to ensure that physiological measurement are sensitive to these external changes and truly measure psychological state and to determine conditions in which unobtrusive physiological measurement is best. These variables can be broken down into two sub-categories which are: (1) Variables within the task itself, and (2) variables within the environment. Below we explore each of these sub-categories.

*Task Characteristics.* There are several variables related to the task itself that we will be looking at, these include: (1) system transparency, (2) modality, (3) taskload (event rate, number of tasks, duration, and breaks), and (4) feedback. First, transparency of the system indicates what information is given to the user in regards to the state of system processes (Parasuraman & Riley, 1997) and can directly translate to an individuals' ability to effectively complete taskwork and mitigate workload. Previous research has also indicated that system transparency—a significant component in the conceptualization of the levels of automation—can affect the vigilance and situation awareness of an operator (Parasuraman & Riley, 2000). For example, a system that informs the operator about the tasks and processes it conducts, or allows for input by them can help in maintaining vigilance of the operator. Modality, or the sensory and cognitive paradigms (See, Howe, Warm & Dember, 1995) which demonstrate differences in performance in vigilance based on whether the task is visual, auditory, or haptic, relates to the Multiple Resource Model (Wickens, 2002), where personnel have certain capacities to how they can interact with various modalities based on the task conditions.

There are several types of feedback that can be received for a task such as continuous output feedback or post-feedback. Feedback on a task can enhance performance in subsequent performance sessions by as much as 20% (Tannenbaum & Cerasoli, 2013). Additionally, feedback received during a performance session can significantly alter strategies for situation awareness and performance on vigilance tasks (Kluger & DeNisi, 1996). Therefore, it is critical to know how, when and if participants received feedback in addition to their tasking conditions. See table 1 for specific forms of taskload.

Table 1
*Characteristics of Taskload*

| *Form of Taskload* | *Theory* | *References* |
|---|---|---|
| *Event Rate* | Increase event rate can impact cognitive variables of performance | Rose, Murphy, Byard, &Nikzad, 2002 |
| *Number of Tasks* | Can be multiple tasks, or a battery of tasks; Impact cognitive performance by divided attention | Hart & Staveland, 1988;Yagoubi & Slimani, 2007 |
| *Duration* | Increase duration, increases workload, fatigue, decreases SA and vigilance | Block, Hancock & Zakay, 2010 |
| *Number of Breaks* | Regains mental resources, increases performance, decreases fatigue, avert vigilance decrement | Phipps-Nelson & Redman et al., 2011; Ariga, Lleras et al., 2011 |
| *Signal to Noise Ratio* | Response sensitivity increases as signal strength decreases and response bias and sensitivity affect performance and vigilance decrement | Mackworth & Taylor, 1963; Teichner, 1974 |

*Environmental Characteristics*. In order to accurately measure physiological measures, environmental factors must also be controlled or accounted for accurate measurement. Environmental characteristics critical to interpreting physiological measurement include temperature, humidity level, as well as the location of the study. Temperature, namely extreme cold and extreme heat can greatly impact physiological performance, but can also decrease variables of cognitive performance (Hancock, Ross, & Szalma, 2007). Therefore, extreme temperature could confound the physiological measures' ability to accurately capture cognitive states. Humidity plays a role in this as well, more specifically when attempting to accurately capture specific forms of physiological measurement (e.g., galvanic skin response). Features of the environment, such as lighting, air quality, and sound can also have an impact on physiological measures and self –report measures (Thayer, et al., 2010). Therefore, in accordance with Hancock & Warm's (1989) dynamic model, conditions of the environment should be a priority for determining measurement confounds and the validity of inference of psychological states and performance.

**Throughput**

*Individual differences.* Preconditions of the operator characterize individual differences of the person interacting with the task/environment to accomplish a goal. While there are anatomical differences which can increase physiological variability (e.g., sex of participant; Hochwarter, Perrewe, & Dawkins, 1995), we argue that psychological preconditions of the operator are critical for physiological assessment. Both the state and trait characteristics of the operator affect how this interaction takes place. Expertise of the operator on the subject, interface or environment can play a key role in examining cognition and physiological response. Arousal, as an individual state, is strongly linked to the availability of mental resources an individual can access, and can indicate stress (Morgan, Matthews, & Winton,, 1995). Arousal, for instance may provide important information to distinguishing the measurement of stress from workload measurement in using physiological criteria. Additionally, more stable personality traits of the individual can influence cognitive appraisal and performance. This is especially true in vigilance tasks such that personality can impact the timing and severity of a vigilance decrement (Szalma, 2009), and ability to maintain situation awareness (Doherty & Walker, 1966). Finally, expertise on a given subject, interface, or environment can also impact performance on a cognitive task and influence cognitive states, such as subjective workload. More specifically, individual expertise can act as a "buffer" against psychological stressors within an environment. See Table 2 for additional information on Individual differences.

*Physiology and Physiological Measurement.* As individuals respond to the characteristics of the task and environmental context, they elicit a physiological response. While there are numerous forms of physiological measurement (e.g., Biomarkers, on-body, off-body), the quality of a given physiological measure can be captured in similar ways. Kramer (1991) advanced a framework for evaluating physiological measures for workload based on the sensitivity/diagnosticity, reliability, and intrusiveness which he framed as the generality of application of the measure. The guidelines stipulates that while individuals will exhibit cognitive workload. While his guidelines provide direction for capturing cognitive workload, these features can be generalized to other variables of cognitive performance including fatigue, situation awareness, and vigilance. Additionally, this framework can be applied as an aid in diagnosing, analyzing, and differentiating between physiological measures of brain activity, eye movements, voluntary muscular activity, cardiovascular activity, electro dermal activity, chemical activity and other bodily processes such as respiration and temperature. Adding this framework as a theoretical basis for diagnosing the appropriateness of measurement can aid in determining which is most appropriate based on their generality of application.

Table 2.
*Non-clinical Individual Differences and Cognition*

| *Individual Difference* | *Theory* | *References* |
|---|---|---|
| *Personality* | Can influence situation Awareness, fatigue, cognitive workload, and vigilance | Doherty & Walker, 1966; Szalma, 2009 |
| *Arousal* | Can indicate stress when high, fatigue when low; affective states of energetic arousal, tense arousal, and hedonic tone | Wickens & Holland, 2000; Matthews et al., 1995 |

**Output**

*Cognitive Outcomes.* Although there are many measures of output of individual adaptation, and particularly cognitive performance variables; previous papers have distinguished cognitive performance as memory tasks, accuracy and response time measurements, tracking and detection tasks (Hancock, Ross, & Szalma, 2007). Specifically, we conceptualize situation awareness as "the perception of the elements in the environment …, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1988). In terms of fatigue, we aim to capture individuals' level of cognitive exhaustion, weakness, or "loss of maximal force-generating capacity during muscular activity" (Chadler et al., 1993, p. 147). Vigilance can also be captured as the sustained attention, or tonic alertness, of an individual maintaining focused attention over time (Oken, Salinsky, & Elsas, 2006). Finally, we are exploring cognitive workload as the amount of attentional and working memory demands experienced during a specified time period (Block, Hancock, & Zakey, 2010).

## DISCUSSION

### Practical Implications

Applications of this framework can be utilized for the purposes of improving safety, training and the efficiency of operations through early detection of critical psychological states. Based on identifying the critical cognitive constructs for personnel working in complex environments, we plan to gain insight into what physiological measurements are the most effective for practical purposes to these factors. As each physiological measurement currently available have their own advantages and limitations related to practicality (e.g., self-report issues; Mabe & West, 1982) and construct validity, it is theorized that a combination of multiple measurements will be the most effective in order to provide a comprehensive summary of personnel conducting specific tasks at a given point in time.

It is expected that this research effort will aid in the development and implementation of countermeasures that can help maintain performance in tasks requiring high levels of cognitive performance. Potential practical implementations can involve detection of physiological changes that can lead to performance decrement, such as hypoxia or temperature changes. Additionally, this effort also aims to guide the development of systems and protocols to ensure appropriate implementation for performance maintenance related to the cognitive factors of interest, such as changes in shift work, role and task assignment, selection for field operations, and insights into task training effectiveness.

Current limitations associated with this effort involve verifying the validity of the proposed framework in this paper. Research is required to expand upon previous efforts for validating physiological measures (e.g., Berka et al., 2007; Haapalainen, Kim, Forlizzi, & Dey, 2010) to be implemented in real world environments. Comparison of these measurements across a variety of tasks and domains can help in selection purposes, as different measurements may act differently from one another based on these conditions.

### Conclusion

This paper introduces a theoretically-grounded framework to provide guidance for identifying and measuring cognitive constructs utilizing real-time physiological measurements to monitor individuals in their work environment. In order to effectively evaluate the validity of real-time physiological measurements in complex environments, it is important to consider the effects of potential confounding effects of the person, task, and environment. Measurements implemented in real world settings must be valid taking into account these confounding and moderating effects. The overall goal of the framework is to set the ground in developing specifications for what physiological measurements will be most effective for monitoring personnel in real time in work environments. Similarly, future research should seek to validate this framework. Specifically, the field would benefit from determination of a meta-analytic effect size to help determine a true correlation between metrics of cognition and physiology (see Mabe & West, 1982).

## REFERENCES

Allen, W. J. & Yen, W.M. (2002). *Introduction to Measurement Theory.* Long Grove, IL: Waveland Press.

Ariga, A., & Lleras, A. (2011). Brief and rare mental “breaks” keep you focused: Deactivation and reactivation of task goals preempt vigilance decrements. *Cognition, 118*(3), 439-443.

Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., & Craven, P. L. (2007). EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, space, and environmental medicine*, *78*(Supplement 1), B231-B244.

Block, R. A., Hancock, P. A., & Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta psychologica*, *134*(3), 330-343.

Bobko, P. (Ed.). (2001). *Correlation and regression: Applications for industrial organizational psychology and management.* Sage.

Camilli, M., Terenzi, M., & Di Nocera, F. (2007). Concurrent validity of an ocular measure of mental workload. *Human factors issues in complex system performance*, 117-129.

Cacioppo, J. T., & Tassinary, L. G. (1990). Inferring psychological significance from physiological signals. *American Psychologist, 45*(1), 16.

Chalder, T., Berelowitz,G., Pawlikowska, T.,Watts, L., Wessely, S., Wright, D. and Wallace, E.P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research, 37*(2), 147-153

Doherty, M. A., & Walker, R. E. (1966). The relationship of personality characteristics, awareness, and attitude in a verbal conditioning situation. *Journal of Personality, 34*(4), 504-516.

Endsley, M. R. (1988, October). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 32, No. 2, pp. 97-101). SAGE Publications.

Haapalainen, E., Kim, S., Forlizzi, J. F., & Dey, A. K. (2010, September). Psycho-physiological measures for assessing cognitive load. In *Proceedings of the 12th ACM international conference on Ubiquitous computing*, 301-310.

Hancock, P. A., Ross, J. M., & Szalma, J. L. (2007). A meta-analysis of performance response under thermal stressors. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 49*(5), 851-877.

Hancock, P. A., Warm, J. S. (1989). A dynamic model of stress and sustained attention. Human Factors: *The Journal of the Human Factors and Ergonomics Society, 31*(5), 519-537.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

Hochwarter, W.A., Perrewe, P.L., & Dawkins, M.C. (1995). Gender differences in perceptions of stress-related variables: Do the people make the place or does the place make the people?. *Journal of Managerial Issues, 7*(1), 62-74.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin, 119*(2), 254.Mabe, P. A. & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology, 67*(3), 280-296.

Kramer, A.F. (1991). Physiological metrics of mental workload: A review of recent progress. In D.L. Damos (Ed.), *Multiple Task Performance,* (pp.279-328). Bristol, PA: Taylor & Francis Group.

Mackworth, J. F., & Taylor, M. M. (1963). The d'measure of signal detectability in vigilancelike situations. *Canadian Journal of Psychology/Revue canadienne de psychologie, 17*(3), 302.

Morgan, I. A., Matthews, G., & Winton, M. (1995). Coping and personality as predictors of post-traumatic intrusions, numbing, avoidance and general distress: A study of victims of the Perth flood. *Behavioural and Cognitive Psychotherapy*, *23*(03), 251-264.

Nemeth, C. P. (2004). *Human factors methods for design: Making systems human-centered*. Boca Raton, FL: CRC Press.

Oken, B. S., Salinsky, M. C., & Elsas, S. M. (2006). Vigilance, alertness, or sustained attention: physiological basis and measurement. *Clinical Neurophysiology, 117*(9), 1885-1901.

Proctor, R. W., & Van Zandt, T. (2008). *Human factors in simple and complex systems*. Boca Raton, FL: CRC Press.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 39*(2), 230-253.

Phipps-Nelson, J., Redman, J. & Rajaratnam, S. (2011). Temporal profile of prolonged, night-time driving performance: breaks from driving temporarily reduce time-on-task fatigue but not sleepiness. *Journal of sleep research*, *20*(3), 404-415.

Robert, G., & Hockey, J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological psychology, 45*(1), 73-93.

Rose, C. L., Murphy, L. B., Byard, L., & Nikzad, K. (2002). The role of the Big Five personality factors in vigilance performance and workload. *European Journal of Personality, 16*(3), 185-200.

See, J. E., Howe, S. R., Warm, J. S., & Dember, W. N. (1995). Meta-analysis of the sensitivity decrement in vigilance. *Psychological Bulletin, 117*(2), 230.

Stevens, R., Galloway, T., Berka, C., & Behneman, A. (2010, January). A neurophysiologic approach for studying team cognition. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)* (Vol. 2010, No. 1). National Training Systems Association.

Szalma, J. L. (2009). Individual differences in human–technology interaction: incorporating variation in human characteristics into human factors and ergonomics research and design. *Theoretical Issues in Ergonomics Science 10*(5), 381-397.

Tannenbaum, S. I., & Cerasoli, C. P. (2013). Do team and individual debriefs enhance performance? A meta-analysis. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 55(*1), 231-245.

Teichner, W. H. (1974). The detection of a simple visual signal as a function of time of watch. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 16*(4), 339-352.

Thayer, J.F., Verkuil, B., Brosschotj, J.F., Kevin, K., West, A., Sterling, C., …. & Sternberg, E.M. (2010). Effects of the physical work environment on physiological measures of stress. *European Journal of Cardiovascular Prevention & Rehabilitation, 17*(4), 431-439.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science, 3*(2), 159-177.

Yagoubi, B., & Slimani, Y. (2007). Task Load Balancing Strategy for Grid Computing. *Journal of Computer Science*, *3*(3).