The Development and Application of a Novel Physiological Metric of Cognitive Workload

Jeremy C. Rietschel[1] and Matthew W. Miller[2]

[1] Maryland Exercise and Robotics Center of Excellence
VA Rehabilitation Research & Development
Baltimore, MD 21201, USA
jcrietschel@yahoo.com
[2] Department of Kinesiology
Auburn University
Auburn, AL 36849, USA
mwm0024@auburn.edu

Abstract. An objective assessment of the cognitive burden imposed by a task (cognitive workload) is of fundamental interest in that it would provide a "window" into one's current allocation of cognitive resources. Such insight would have tremendous implications in maximizing human performance through a multitude of applications including human-computer interaction. The authors propose a novel, electroencephalographic (EEG)-derived metric, which relies on the event-related potential (ERP) component, novelty-P3. A theoretical rationale and experimental evidence supporting the metric's utility are provided, followed by future directions.

Keywords: Cognitive workload, human performance, EEG, novelty-P3.

1 Introduction

This paper will present a novel method to assess the cognitive burden imposed when one performs a task (i.e., cognitive workload). First, the importance of such a metric will be discussed followed by how this metric was conceptualized and developed. Next, three experiments aimed at validating the capability of this measure regarding the assessment of cognitive workload will be presented. Finally, the paper will conclude with recommendations for future research regarding this metric.

1.1 Why Measure Cognitive Workload?

An accurate measure of cognitive workload would be useful in a multitude of ways. For instance, one would be able to determine how different task conditions impact the mental state. This information would be useful in designing a task so as to reduce excessive cognitive workload and limit mental fatigue. Additionally, a cognitive workload assessment could serve as a forecast of future behavior. For example, two

D.D. Schmorrow and C.M. Fidopiastis (Eds.): AC/HCII 2013, LNAI 8027, pp. 582–590, 2013.

individuals could be executing the same task at comparable levels of performance and thus would be indistinguishable from each other using a behavioral level of analysis. However, it could be that one individual is performing the task at a considerable cognitive 'cost' whereas the other individual is able to perform similarly with little strain placed on his/her cognitive resources. Knowing this, one could predict which individual could maintain his/her level of performance longer and who would be better able to cope with unexpected increases in task demands (the latter individual). Similarly, task mastery has been robustly associated with automaticity (the ability to perform a task with little mental effort), and as such, measuring cognitive workload would inform skill level beyond that of looking at the performance alone. In addition, continual monitoring of cognitive workload would reveal the dynamic mental state of an individual. This information could be used to maximize user/task interaction by adjusting task demands to match the user's current cognitive state. For example, if a cognitive workload metric detects that an aircraft pilot is experiencing excessive cognitive workload while flying the aircraft, the machine (aircraft) could assume task demands by engaging an autopilot feature. Similarly, in a team environment task, demands could be dynamically allocated among team members based on their respective cognitive workloads such that each member maintains a manageable load. For pictorial examples illustrating the utility of a metric assessing cognitive workload, see Figure 1.

1.2 Background and Development

Cognitive resources are limited in regards to quantity [1]. As one engages in a task, the cognitive workload imposed by the task draws upon these finite cognitive resources. The spare resources not currently being utilized by the task are referred to as attentional reserve and are available to allocate to additional task demands (e.g., unexpected events). In this regard, cognitive workload and attentional reserve are inversely related such that when cognitive workload increases, attentional reserve decreases. Conversely, when cognitive workload is reduced, attentional reserve grows [2]; see Figure 2A. Therefore, assessing attentional reserve provides insight into the current state of cognitive workload.

Thus, in order to develop a technique to measure cognitive workload, we sought to objectively quantify attentional reserve using a neurobiological approach. The electroencephalographic (EEG) technique measures the electrical activity of the brain. Brain activity associated with the processing of stimuli can be assessed by extracting a portion of the EEG signal time-locked to the onset of the stimuli—these EEG segments are known as event-related potentials (ERPs). ERPs are comprised of different components, each of which reflects a distinct cognitive process. The component known as the novelty-P3 reflects the automatic orienting of attention to novel stimuli, and the amplitude of the novelty-P3 component is positively related to the degree of this cognitive process [3]. The degree to which attention can be oriented to novel stimuli depends on the availability of cognitive resources for such orienting (i.e., the magnitude of attentional reserve). Thus, we reasoned that novelty-P3 component amplitude would reflect the quantity of attentional reserve. Specifically, when attentional

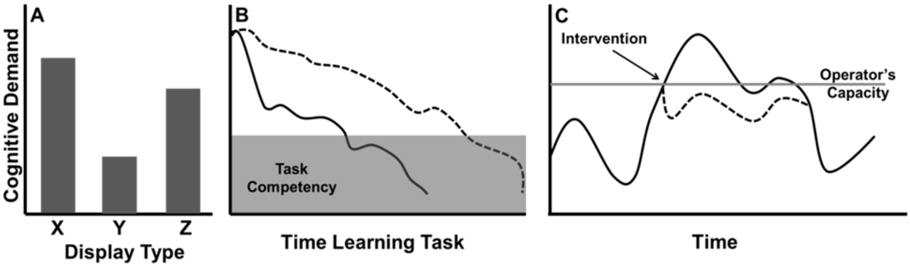

Fig. 1. Illustrative scenarios demonstrating the utility of an assessment of cognitive workload. A) In order to determine the most efficient means to convey information, three different visual displays representing the same information are presented to individuals. Their cognitive workload is assessed during these presentations and it is revealed that display Y conveys the information with the least cognitive demand imposed, thus it is the most efficient. B) In order to determine when a flight controller has had enough training to begin real-world operation their cognitive workload is assessed as they learn how to perform their task. When they can perform the task with minimal cognitive workload (i.e., perform the task below a specified threshold of cognitive demand—"task competency"), then they are considered adequately trained. In the current example the two tracings correspond to two trainees with the solid line representing a trainee who reached the competency threshold quicker than the trainee represented by the dotted line. C) During the dynamic production of a task, if the cognitive demand associated with the task exceeds the operator's capacity then the probability of failure greatly increases. The ability to monitor the cognitive workload during task production (solid line) would inform when demand is exceeding capacity, which could trigger an intervention aimed at reducing the demand (dotted line) thus averting the increased risk of failure. For example, if a pilot became overloaded during a flight, then the co-pilot could begin to take over some of the responsibilities, effectively reducing the pilot's load and, thus, the probability of an accident.

reserve is high, many cognitive resources are available to be oriented to novel stimuli, which should then be reflected by large novelty-P3 component amplitudes. Conversely, when attentional reserve is lower, fewer cognitive resources are available to be oriented to novel stimuli, which should result in reduced amplitude (see Figure 2B). Given the inverse relationship between attentional reserve and cognitive workload, we predicted high cognitive workload should result in small novelty-P3 amplitude, whereas lower cognitive workload should result in larger novelty-P3 amplitude. In this regard, we predicted the novelty-P3 component should be effective in assessing cognitive workload.

In line with this rationale, our approach in assessing cognitive workload involves probing individuals with stimuli known to elicit the novelty-P3 component while they engage in a primary task (a task for which cognitive workload measurement is of interest). Specifically, we present individuals with novel, task-irrelevant, ecologically-valid auditory stimuli (e.g., a woman coughing, a dog barking, a glass breaking). Concurrently, EEG is recorded and time-locked to the stimuli. Next, ERPs to the stimuli are extracted and the average amplitude of the novelty-P3 is computed.

There are three distinct advantages to this approach. First, the EEG signal is an objective assessment and thus not influenced by the subjectivity typically introduced when employing self-report methods of cognitive workload assessment. Second, the most commonly employed method used to measure cognitive workload (i.e., the dual-task paradigm; [e.g., 4-7]) may risk inherently confounding the assessment [8-9]. In dual-task paradigms, participants are probed with stimuli to which they are asked to attend (secondary task) while performing the primary task. For example, participants may be asked to count auditory stimuli (secondary task) while performing a simulated aircraft flight (primary task). The major limitation of such paradigms is that the addition of a having to attend to secondary task stimuli may fundamentally interact with the primary task, thus compromising the magnitude of cognitive workload imposed by the primary task alone. As our method probes individuals with task-irrelevant stimuli (i.e., stimuli to which individuals are not instructed to attend), it avoids this limitation altogether. Third, we probe individuals with ecologically-valid, novel stimuli. The salience of such stimuli is believed to induce a compulsory orienting of spare cognitive resources [3]. Therefore, this method is likely to provide a robust assessment of attentional reserve and thus cognitive workload.

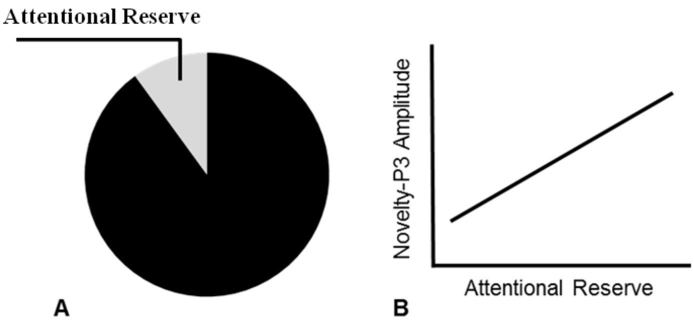

Fig. 2. A) The conceptual model indicating (1) that cognitive resources are fixed with regard to total capacity, and (2) when a cognitive workload is imposed, the resources that are spared are referred to as attentional reserve. Accordingly, this relationship reveals that measuring attentional reserve will, in turn, reveal the magnitude of cognitive workload. B) Hypothesized relationship between attentional reserve and novelty-P3 amplitude. As attentional reserve increases, this is reflected in increased novelty-P3 amplitude. Conversely, as attentional reserve decreases, novelty-P3 amplitude becomes reduced.

2 Experimental Assessment of the Metric

2.1 Experiment 1

The first experiment aimed at testing the validity of our cognitive workload metric involved incrementally varying the difficulty of a primary task [10]. We reasoned that increasing task difficulty would elicit a corresponding cognitive burden, thus raising cognitive workload. Therefore, we predicted that incremental modulations in task difficulty would induce dose-dependent changes in cognitive workload and, as such, our metric should be sensitive to these changes.

Twenty participants performed the videogame Tetris at three levels of difficulty presented in random order: View, Easy, and Hard. Tetris requires individuals to use a keyboard to manipulate different-shaped game pieces presented on a video screen in order to place them in an optimal location (limiting the space between the current piece's placement and previously played pieces). During the View level of difficulty, participants watched Tetris but did not manipulate the game pieces. This level was expected to impose the least cognitive burden as individuals did not directly interact with the game. During the Easy level, participants maneuvered game pieces moving down the video screen at a velocity of 1.67 cm/s, whereas during the Hard level participants manipulated pieces moving at 3.56 cm/s. This difference in speed was believed to elicit greater cognitive workload in the Hard condition as compared to the Easy, as participants had to more quickly decide where to place the current game piece, execute the placement, and update their planning for successive pieces. During each level, we employed our cognitive workload assessment. Specifically, participants were probed with novel, task-irrelevant, ecologically-valid auditory stimuli. Concurrently, EEG was recorded and time-locked to the stimuli. Next, ERPs to the stimuli were extracted and the average amplitude of the novelty-P3 was computed.

Behavioral results revealed poorer task performance in the Hard level than the Easy level, suggesting a successful manipulation of task difficulty. As predicted, novelty-P3 amplitude incrementally changed as a function of task difficulty level. Specifically, novelty-P3 amplitude was largest in the View level, second-largest in the Easy level, and smallest in the Hard level (see Figure 3). These results suggest that our metric is able to provide an effective assessment of cognitive workload. Specifically, a negative relationship between cognitive workload and novelty-P3 amplitude was observed, which is consistent with our conceptual model.

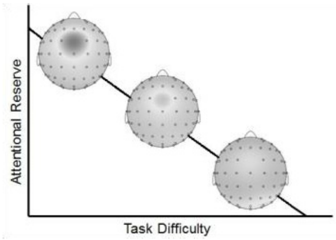

Fig. 3. Support of our conceptual model. Specifically, as a task becomes more difficult, cognitive workload increases, resulting in reduced attentional reserve. The scalp maps of the novelty-P3 are actual data from the three difficulty levels: View, Easy, and Hard (darker grey indicates higher novelty-P3 amplitude). As expected, novelty-P3 amplitude was inversely related to task difficulty, suggesting that our metric is able to provide an effective assessment of cognitive workload.

2.2 Experiment 2

In our second experiment task difficulty was held constant while participants' skill level improved [11]. It is generally accepted that as individuals learn a new task, the cognitive workload required to perform that task becomes reduced [12]. Accordingly, we sought to examine if our metric was sensitive to changes in cognitive workload related to individuals' current skill level.

Twenty-one participants all performed a center-out reaching task that required moving as quickly and accurately as possible to targets. However, they were randomly assigned to either a group that learned a novel visuomotor distortion (i.e., requires learning) or to a control group that performed the same task with no distortion element (i.e., no learning). For the duration of the task, our metric was employed to assess cognitive workload. We predicted novelty-P3 amplitude would initially be low in the Learning group relative to the Control group, but that there would be a progressive increase in amplitude in the Learning group as a function of learning. Additionally, as the Control group was not required to learn the distortion, we predicted that novelty-P3 amplitude would remain relatively stable.

Behavioral evidence supported that the Learning group experienced learning whereas the control group did not (i.e., the Learning group significantly improved task performance, whereas the Control group's performance remained stable). As expected, the Learning group exhibited a progressive increase in novelty-P3 amplitude over the course of learning, whereas the Control group did not exhibit significant changes in amplitude (Figure 4). In other words, across the time period where individuals learned a new skill, our metric revealed a progressive decrease in their cognitive workload. Moreover, our metric revealed no change in the cognitive workload of individuals assessed across the same time period, performing the same task but without the learning component. Thus, our metric was sensitive to predictable changes in cognitive workload associated with skill learning.

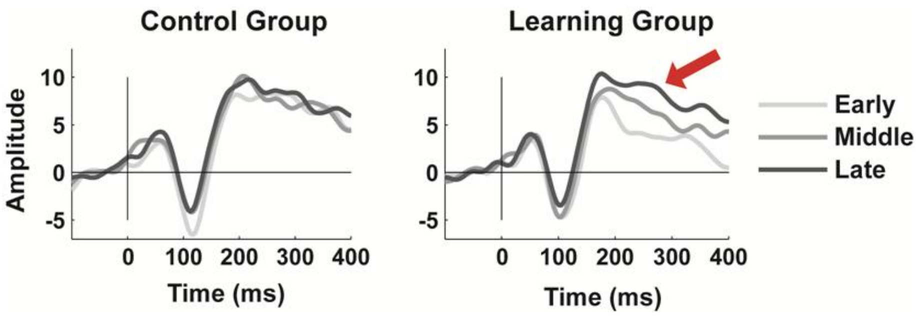

Fig. 4. The change in novelty-P3 amplitude as a function of skill level. On the right panel, the Learning group's ERPs to the auditory stimuli are presented. The light grey, darker grey, and black lines correspond to early, middle, and late learning, respectively. Consistent with predictions, novelty-P3 amplitude (indicated by the arrow) become larger as a function of learning, supporting our method as a valid assessment of cognitive workload. On the left panel, the Control group's data are presented. As expected, there was no change in novelty-P3 amplitude over the course of task performance.

2.3 Experiment 3

In our third experiment, task difficulty and participants' skill levels were held constant while participants' environments were manipulated [13]. Specifically, 12 participants performed Tetris at a difficulty level yoked to his/her respective skill level in two social environments: a high quality team environment and a low quality team environment. In the High Quality Team Environment, participants performed Tetris with a teammate who they perceived as being competent. Conversely, in the Low Quality Team Environment, participants performed with a teammate who they perceived as being incompetent. Prior research has indicated that individuals performing in high quality team environments experience significantly reduced cognitive workload relative to performing in lower quality environments [14]. Accordingly we sought to examine if our metric was sensitive to changes in cognitive workload related to this aspect of the social environment.

Participants reported, via a questionnaire, that the High Quality Team Environment was the preferred social environment. As expected, participants exhibited higher novelty-P3 amplitudes in the High Quality Team Environment relative to the Low Quality Team Environment, suggesting that cognitive workload was lower in the former. Thus, our metric detected predictable changes in cognitive workload as a function of social environment.

2.4 Summary of Experiments

Collectively, these three studies support our novel metric's ability to assess cognitive workload. Specifically, novelty-P3 amplitude was demonstrated to be sensitive to multiple factors known to influence cognitive workload: changes in task difficulty while holding skill level constant, changes in skill level while holding task difficulty constant, and changes in environmental factors in which both task difficulty and skill level were held constant. Further, in the case of the first two experiments, the metric behaved in a dose-dependent, predictable fashion. Specifically, the metric revealed graded increases in cognitive workload concomitant with incremental increases in task difficulty, and progressive decreases in cognitive workload as a function of skill learning. These results underscore the fidelity and sensitivity of the measure as well as its utility in application.

3 Future Directions

As the employment of this cognitive workload metric progresses, we recommend several future directions regarding research in this area. First, the utility and integrity of this metric need to be rigorously investigated in a myriad of ecologically valid contexts. For example, a study similar to Experiment 1 in the current paper could be conducted in a 'real-world' environment, such as having individuals drive cars during high-density versus low-density traffic. Similarly, the results of Experiment 2 need to be demonstrated to generalize to a diverse set of tasks as a function of learning and skill level. Secondly, although this metric has been shown to be sensitive to alterations in cognitive workload, a behavioral consequence associated with this index has not been demonstrated. In other words, what is the predictive ability of this metric with regard to performance? For example, one could determine that if the metric suggests an individual is under a high workload, does this correspond to a reduced ability to respond to additional challenge, such as an unexpected, 'surprise,' event. Thirdly, the metric currently requires that novelty-P3 amplitude be determined by computing its average response to multiple stimuli, thus limiting the ability to assess cognitive workload in near 'real time.' Therefore, different signal processing methods (e.g., wavelet analyses) need to be applied in order to compute the novelty-P3 after each stimulus presentation, thereby increasing the temporal resolution of this metric.

4 Conclusion

In this paper we described the utility of a metric that could reliably assess cognitive workload. We then outlined a theoretical rationale for how to assess this and conceived a corresponding novel metric. Experimental evidence was provided that suggested this metric is successful in assessing predictable changes in cognitive workload as a function of task difficulty, learning, and environment. We concluded with recommendations for future research.

References

1. Kahneman, D.: Attention and Effort. Prentice-Hall, Englewood Cliffs (1973)
2. Wickens, C., Kramer, A., Vanasse, L., Donchin, E.: Performance of concurrent tasks: a psychophysiological analysis of the reciprocity of information-processing resources. Science 221, 1080–1082 (1983)
3. Friedman, D., Cycowicz, Y.M., Gatea, H.: The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. Neurosci. Biobehav. Rev. 25, 355–373 (2001)
4. Isreal, J., Chesney, G., Wickens, C., Donchin, E.: P300 and tracking difficulty: evidence for multiple resources in dual-task performance. Psychophysiology 17, 259–273 (1980)
5. Isreal, J., Wickens, C., Chesney, G., Donchin, E.: The event-related potential as an index of display monitoring workload. Hum. Factors 22, 211–224 (1980)
6. Kramer, A.F., Sirevaag, E.J., Braune, R.: A psychophysiological assessment of operator workload during simulated flight missions. Hum. Factors 29, 145–160 (1987)
7. Sirevaag, E.J., Kramer, A.F., Coles, M.G.H., Donchin, E.: Resource reciprocity: an event-related brain potential analysis. Acta Psychol. 70, 77–97 (1989)
8. Kramer, A.F., Wickens, C.D., Donchin, E.: Processing of stimulus properties: evidence for dual task integrality. J. Exp. Psychol. Hum. 11, 393–408 (1985)
9. Papanicolaou, A., Johnstone, J.: Probe evoked potentials: theory, method and applications. Int. J. Neurosci. 24, 107–131 (1984)
10. Miller, M.W., Rietschel, J.C., McDonald, C.G., Hatfield, B.D.: A novel approach to the physiological measurement of mental workload. Int. J. Psychophysiology 80, 75–78 (2011)

11. Rietschel, J.C., Goodman, R.N., McDonald, C.G., Miller, M.W., Jones-Lush, L., Wittenberg, G.F., Hatfield, B.D.: Psychophysiologcial investigation of attentional processes during motor skill learning. In: 42nd SFN Meeting, New Orleans, LA, USA, October 13-17 (2012)
12. McGill, R.A.: Motor Learning and Control: Concepts and Applications. McGraw-Hill, New York (2011)
13. Miller, M.W., Groman, L.J., Rietschel, J.C., McDonald, C.G., Iso-Ahola, S.E., Hatfield, B.D.: The effects of team environment on attentional resource allocation and cognitive workload. Sport, Exerc., and Perform. Psychol. (2013), doi:10.1037/a0030586
14. Miller, M.W., Presacco, A., Groman, L.J., Bur, S., Rietschel, J.C., Gentili, R.J., McDonald, C.G., Iso-Ahola, S.E., Hatfield, B.D.: The effects of team environment on cerebral cortical processes and attentional reserve. Sport, Exerc., and Perform. Psychol. (in press)