The Index of Cognitive Activity: Measuring Cognitive Workload

Sandra P. Marshall

Abstract—**The Index of Cognitive Activity is an innovative technique that provides an objective psychophysiological measurement of cognitive workload. As users operate in increasingly complex environments, it is essential that the designers of these environments understand the cognitive demands placed on the users. The Index of Cognitive Activity (ICA) provides an important estimate of the levels of cognitive effort of the user. The ICA is based on changes in pupil dilation that occur as a user interacts with a visual display. The Index is described, and several applications are presented.**

Index Terms—**cognitive science, user interfaces, visual system, workstation human factors**

I. Introduction

Eye movements and changes in pupil dilation provide important information about how users interact with complex visual displays. Both types of data can be obtained by using eye-tracking apparatus that captures eye data in a nearly continuous signal, providing precise information about what the user looks at, how long he looks at it, and how much his pupils dilate as he looks.

New eye-tracking systems yield an extraordinary amout of data. With binocular tracking and sampling rates up to 500 Hz, it is not unusual to obtain over a million data points for one user during a single interactive session. Thus, the primary challenge for today's eye-tracking researcher is to reduce the data to a meaningful synthesis. An equally important challenge is to produce the synthesis in a useful amount of time—and preferably in real time.

The value of eye position alone is evident: the point of gaze reveals whether a user looks at a specific feature, whether he ignores it, or whether he notices it at an inappropriate time. The value of pupillary change is equally great although not as widely known. Psychologists have argued for more than 20 years that changes in pupil dilation accompany effortful cognitive processing [1]. Many studies have validated this argument across a variety of tasks, including reading, problem solving, and visual search [2]-[4]. However, hardware and software limitations made it difficult to measure pupil changes on complex tasks over extended periods of time.

Only recently has a technique been introduced that allows reliable and rapid estimation of cognitive workload from changes in pupil dilation. The technique is the Index of Cognitive Activity (ICA) [5]. The ICA is described in the following sections, with examples that illustrate its use. A concluding section discusses the range of potential applications for the Index and outlines current research and development efforts.

II. Overview of ICA

The Index of Cognitive Activity measures abrupt discontinuities in the signal created from continuous recording of pupil diameter. In the presence of effortful cognitive processing, the pupil responds rapidly with a reflex reaction. At the same time, the pupil also makes a reflex reaction to changes in light. The ICA successfully separates the light reflex from the dilation reflex. Details about the techniques used to calculate ICA may be found elsewhere [6].

The ICA has several advantages over other techniques that measure changes in pupil dilation. First, it does not require averaging over trials or over individuals. Second, it can be applied to a signal of any length. And, third, it can be computed in nearly real time (i.e., within a few seconds). The index is computed as the number of times per second that an abrupt discontinuity in the pupil signal is detected.

A simple example will illustrate the ICA. Consider two factors: the presence or absence of light and the presence or absence of cognitive effort. When fully crossed, these two factors yield four testable conditions: light plus cognitive effort, dark plus cognitive effort, light plus no cognitive effort, dark plus no cognitive effort.

The ICA should detect significant difference between cognitive effort and no cognitive effort in both

The research described here was supported by the U.S. Department of Defense under the Defense Augmented Research Projects Agency (DARPA Grant No. N00014-02-1-0349), the Office of Naval Research (ONR Grant No. N00014-00-1-0353), and the Air Force Office of Scientific Research (AFOSR Grant No. F49620-97-1-0353).

Sandra Marshall is Professor of Psychology and Director of the Cognitive Ergonomics Research Facility, San Diego State University. She is also President and CEO of EyeTracking, Inc. [www.eyetracking.com] (e-mail: smarshall@sciences.sdsu.edu)

IEEE 7th Human Factors Meeting
Scottsdale Arizona 2002
0-7803-7450-9/$17.00©2002 IEEE

lighting conditions—light and dark. Conversely, it should show small or no difference between light and dark conditions for both levels of cognitive effort—no task and task. In this example, the cognitive task was mental computation of arithmetic items that were presented orally.

In the Dark Conditions, the screen was dark and the room lights were turned off. In the Light Conditions, the screen was on but empty and the room lights were at normal workplace settings. For the No Task Conditions, the subject was asked to sit quietly before a computer screen and do nothing. For the Math Task Conditions, the subject listened to a series of items and responded verbally to each one. The items were presented at a rate of one every ten seconds.

The results of all four conditions are shown in Figure 1, which contain the data from a single subject for both eyes.

Figure 1 shows the second-by-second history of cognitive effort required for each of the four conditions over one minute. The cognitive effort is estimated by the Index of Cognitive Activity, which is shown here as three levels of intensity: low levels are white, medium levels are light gray, and high levels are dark gray.

In the top two conditions, the observed levels of ICA were very low. Only twice in the first condition and once in the second condition did the Index exceed the low level, as shown by the spikes in the graph. In contrast, the values for ICA were often in the medium or high range for the bottom two conditions.

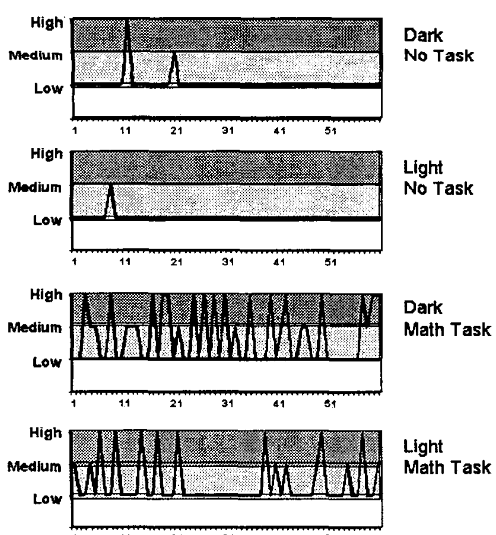

Figure 1. Illustration of ICA

Figure 1 shows clearly that the ICA is not a function of light. When no cognitive task was given, the light condition and the dark condition are virtually identical. Similarly, in the presence of the cognitive task, the light and dark conditions are very similar.[1] High levels of ICA were recorded during the cognitive tasks. Low levels were recorded when no task was given.

III. Validation Through AugCog Collaborations

As part of DARPA's Augmented Cognition (AugCog) Program, the applied neuroscience and cognitive science research communities have begun a number of interdisciplinary collaborations. These collaborations are intended to investigate multiple measures of cognitive workload such as EEG, cardiovascular/physiological arousal, and pupil dilation.

A. ICA and Arousal

A recent AugCog collaboration between the author of this paper and E. Muth of Clemson University undertook simultaneous estimation of the Index of Cognitive Activity with several measurements of physiological arousal. The tasks used for this study were several cognitive tasks involving problem solving and visual search as well as a resting period and a physical exertion task.

The overall results of this collaboration with respect to the ICA are shown in Figure 2. This figure gives the mean numerical values of the index rather than applying the categorical labels of low, medium, and high used in Figure 1. (For this particular study, the participants maintained an unusually high level of stress, presumably because of the nature of the study and the number of sensors that were attached to them for the physiological measurements. Additional investigations will need to be carried out to determine the appropriate category boundaries under these circumstances.)

The six tasks shown in Figure 2 are of two types: four cognitive tasks on the left and two non-cognitive tasks on the right. The cognitive tasks elicited significantly higher ICA for both eyes than did the non-cognitive tasks ($F=41.83$; $df=1,15$; $p<.001$). Follow-up comparisons showed that all possible pairwise comparisons between any cognitive and any non-cognitive task for both eyes were statistically significant.

[1] The Dark Math Task condition preceded the Light Math Task condition. Two different sets of items were used. The full study is described in detail in [6].

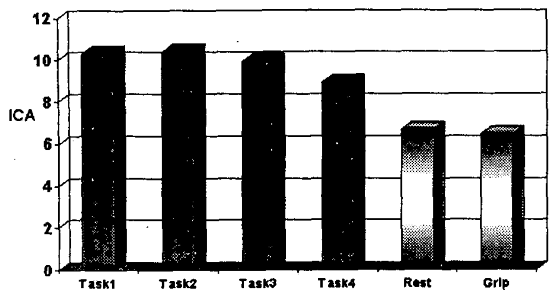

Figure 2. Cognitive and Non-Cognitive Tasks

This result provides statistical support in two ways to the argument that pupil dilation is not a measure of arousal. First, significantly lower levels of ICA were observed for the non-cognitive tasks than for the cognitive ones. Second, no statistical difference was detected between the two non-cognitive tasks. This latter finding is quite interesting, because the grip task was designed especially to show higher physiological rates.

To date, several AugCog collaborations involve data collection using eye tracking and EEG sensors. These efforts have demonstrated on several occasions that it is possible to collect data simultaneously without deleterious interference, although data collection techniques continue to be refined to reduce noise in the recordings. Two collaborations are described below to illustrate how the ICA may be used effectively in combination with other sensors.

B. ICA and EEG #1

In a recent joint study, the present author teamed with two researchers from England, B. Dickson of QinetiQ, Ltd. And C. Pleydell-Pearce of Bristol University. Testing in a research facility in England, the group studied closely the responses of a single subject to a number of tasks. This case study is reported in detail elsewhere [7].

Several important results emerged from this collaboration. As with the earlier collaborative study using arousal measures, this research effort also found high values for the Index of Cognitive Activity for known cognitive tasks and low values for resting state.

One of the tasks used in the study required the subject to monitor a set of gauges and to correct the gauges when they moved outside a given range. The subject responds by toggling the arrow keys on a standard keyboard. The task was created by Professor Pleydell-Pearce and has been used extensively by him in several EEG studies. The task was designed with five levels of difficulty, with 8 trials at each level. The 40 trials were presented in random order with each trial lasting 45 seconds. Each trial was followed by a 15 second rest period.

Two important results emerge from the analysis of this task. First, the subject clearly learned to perform the task, and his mental effort lessens over time. Second, the difficult trials (as defined by the task developer) elicited more cognitive effort and did so in the predicted direction.

Figures 3 and 4 show these results. The ICA values reported in the figures are the mean numbers of abrupt changes in pupil dilation per second for each trial. Both eyes recorded similar values, so their combined ICA results were used for this analysis. The same data are shown in both figures. To facilitate comparisons, the ICA values were normalized.

Figure 3 gives the overall ICA results for all trials in the order they were encountered by the subject. The most striking feature of this graph is the abrupt change that occurred about midway through the task. The ICA changes significantly across Trials 19-21. These data suggest that the subject modified his performance in some way so that the task was suddenly easier for him. (As a parenthetical note, the subject mentioned after the study that he was aware of making an abrupt strategy shift, but he did not recall on which trial this occurred.)

Figure 4 shows the same results grouped by trial level. The trials have been sorted by level, beginning with the easiest trials (Level1) and ending with the most difficult trials (Level5). The means of the five levels are significantly different ($F=2.78$, $df=4,35$, $p<.05$), and the linear trend is significant ($F=11.03$, $df=1,35$, $p<.01$), with Level 1 showing the lowest values of ICA and Level 5 having the highest.

An overall pattern of learning or strategy shift is evident within each level. That is, each level begins with a higher ICA value and moves to a lower value by the final trial within the level. At the same time, the difficulty pattern is also clear, with lower values of the ICA found for the easiest trials and higher ICA values observed for the more difficult trials.

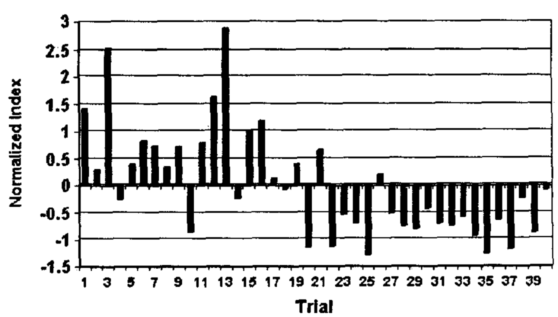

Figure 3. The normalized ICA values for 40 trials

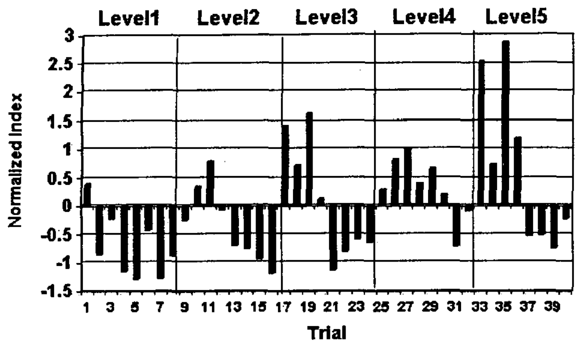

Figure 4. The same trial results, grouped by trial difficulty

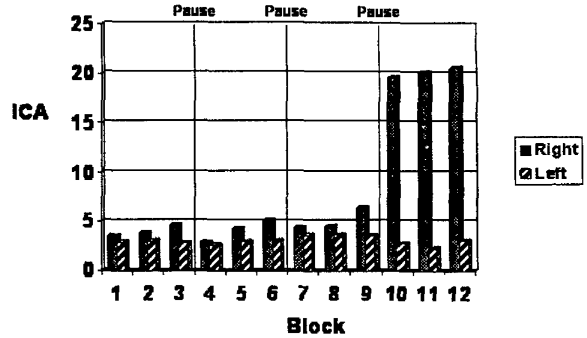

Figure 5. Flanker Task Results

C. ICA and EEG #2

Performance on tasks does not always improve over time. In a second collaborative effort undertaken with L. Parra of Sarnoff Corporation and J. Gelfand of Princeton University, the present author investigated the ICA levels for a very simple repetitive task called the 'flanker task.' The task consisted of a series of 12 blocks with 68 trials in each block.

For each trial, the subject was shown a stimulus consisting of 5 symbols: >>>>> or <<<<<. The center symbol may point either left or right, so that it is flanked on either side by two symbols that are identical to each other. The center symbol may be congruent to these symbols, or it may be incongruent.

The participant uses two simple buttons to indicate whether the center symbol points left or points right, and responses are made as quickly as possible. The task is not cognitively challenging. The participant need only maintain vigilance to make a response, and most individuals make very few errors.

In this collaborative study, the normal experimental conditions were modified. The participant wore EEG sensors as well as an eye-tracking headset, and both preparation and testing time were longer than usual. Moreover, the flanker task was the last task to be performed. Thus, the task was potentially more stressful and required more effort than would typically be the case.

The ICA results for a single subject are shown in Figure 5. The values in the figure are the mean number of instances per second recorded for each block for each eye. The most striking result is the enormous increase in ICA that occurred in the right eye during the final three blocks. It is important to note that nothing unusual happened during the transition from Block 9 to Block 10 and that the left eye ICA did not yield a similar increase. Thus, the increase is presumably not the result of an equipment error.

A second noticeable result is the slight but steady increase seen in the right ICA across each of the 3-block sets, i.e., from Block 1 to Block 3, from Block 4 to Block 6, from Block 7 to Block 9, and finally from Block 10 to Block 12. After the first two rest periods, the ICA values dropped and then began to rise again.

The steady rise in ICA across the 3-block sets may have occurred because the participant became fatigued and needed to use more cognitive resources to performing the task. The drop in ICA immediately following the rest periods supports this explanation as does the repeated pattern across all four sets.

The sharp increase at Block 9 cannot be satisfactorily explained without further testing. As mentioned above, equipment failure is possible but unlikely. Two plausible hypotheses to be considered are: the subject was tired and was able to refocus with greater intensity knowing that only three blocks remained; or the subject was tired and also physically bothered by the sensors and headset during the final three blocks.

IV. SUMMARY AND CONCLUSIONS

The Index of Cognitive Activity has been used in a wide variety of applications ranging from simple laboratory tasks to complex user interfaces. Several important issues are now under investigation, such as differences in modalities (i.e., visual, spatial, auditory) during cognitive processing, hemispheric differences as reflected in left/right eye dilations, and reliable detection of overload.

Collaborative research efforts to link EEG measurements and ICA estimates are continuing, and new experiments are underway. The objective now is to relate specific EEG activity to measured changes in ICA on tasks that are well understood. These efforts should yield new insights about the nature of cognitive processing that is required of users who operate in both simple and/or complex computer environments.

V. References

[1] D. Kahneman, *Attention and Effort.* Englewood Cliffs, NJ: Prentice Hall, 1973, pp. 17-24.

[2] Ahern, S. & Beatty, J. (1979). Pupillary responses during information processing vary with Scholastic Aptitude Test scores. Science, 205, 1289-1292.

[3] Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. Psychophysiology, 33, 457-461.

[4] J ust, M. A. & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. Canadian Journal of Experimental Psychology, 47, 310-339.

[5] S. Marshall, "Method and Apparatus for Eye Tracking and Monitoring Pupil Dilation to Evaluate Cognitive Activity," U.S. Patent 6,090,051, Jul. 18, 2000.

[6] S. Marshall, S. Knust, & C. Davis, *The Index of Cognitive Activity: Estimating Cognitive Effort from Pupil Dilation.* Unpublished manuscript, 2002.

[7] S. Marshall, B. Dickson, & C. Pleydell-Pearce, *Combining EEG and Pupil Dilation to Measure Cognitive Workload: A Case Study.* Paper submitted to the 36th Annual Hawaii International Conference on Systems Sciences, to be held January 2003.