Applying Real Time Physiological Measures of Cognitive Load to Improve Training

Joseph T. Coyne[1], Carryl Baldwin[2], Anna Cole[3],
Ciara Sibley[2,3], and Daniel M. Roberts[2]

[1] Naval Research Laboratory, Washington, DC 20375
[2] George Mason University, Department of Psychology, Fairfax, VA 22030
[3] Strategic Analysis Incorporated, Arlington, VA 22203
cbaldwi4@gmu.edu

Abstract. This paper discusses how the fields of augmented cognition and neuroergonomics can be expanded into training. Several classification algorithms based upon EEG data and occular data are discussed in terms of their ability to classify operator state in real time. These indices have been shown to enhance operator performance within adaptive automation paradigms. Learning is different from performing a task that one is familiar with. According to cognitive load theory (CLT), learning is essentially the act of organizing information from working memory into long term memory. However, our working memory system has a bottleneck in this process, such that when training exceeds working memory capacity, learning is hindered. This paper discusses how CLT can be combined with multiple resource theory to create a model of adaptive training. This new paradigm hypothesizes that a system that can monitor working memory capacity in real time and adjust training difficulty can improve learning.

1 Introduction

Whether it is called the field of Augmented Cognition or Neuroergonomics [1] there has been a recent push to apply findings from the field of neuroscience and the "Decade of the Brain" to improve human performance. While Neuroergonomics focuses on how neuroscience can be applied to work and everyday environments, Augmented Cognition places an emphasis on the design of closed loop systems based upon real time physiological assessment. Until recently there has not been a focus on how advances within these two areas could be applied to learning or training. As the use of computers and simulation become an increasingly important component of the learning process, it seems that applying these two fields and developing a closed loop adaptive training system would be a natural extension of these two areas. Such a system could adjust the content, presentation format, and pace of training to match the specific skills and abilities of the trainee. It is proposed that such a system would reduce the amount of time required to train an individual by reducing the amount of time the trainee is under or overloaded. The present paper reviews previous research in neuroscience and learning which serve as a theoretical basis for the adaptive training system. The paper will discuss the research done with respect to real time

D.D. Schmorrow et al. (Eds.): Augmented Cognition, HCII 2009, LNAI 5638, pp. 469–478, 2009.

physiological assessment, describe the cognitive load theory and explain how these two separate research areas can be merged into a theory of adaptive training.

2 Real Time Physiological Assessment

2.1 Sensors

At the heart of neuroscience are the tools available to measure activity within the brain. There are a number of different sensor technologies available which provide either direct or indirect indices of the brain's activity. These include measures taken from the brain such as electroencephalogram (EEG), functional near-infrared imaging (fNIR), magnetoencephalography (MEG), functional magnetic resonance (fMRI), as well as indirect physiological measures of brain activity such as cardiorespiratory activity, for example heart rate (HR) and heart rate variability (HVR), as well as measures of electrodermal activity such as skin conductance and galvanic skin response (GSR) and pupilometry. The advantages of each of these methods can be assessed along three criteria including spatial resolution, temporal resolution, and ease of use [2]. Of these sensor technologies EEG has been the most widely used for real time assessment due to its temporal resolution and ease of use. Although eye tracking data only provides an indirect measure of brain activity; it is widely used, unobtrusive (particularly with new off the head systems) and easy to collect. For those reasons the present paper will focus on advances made with respect to these two different sensors.

2.2 Real Time Cognitive State

For a physiological sensor to be useful it must be sensitive to an aspect of operator state that has cognitive or performance implications (i.e., stress, arousal, mental workload). Ultimately, much of the research has been looking to find a gauge of mental workload: how hard the brain is working at a given point in time. One of the key aspects of mental workload is the relationship between the task being performed and an individual's limited pool of resources available. Wickens' multiple resource theory [3] distinguishes between three orthogonal resource dimensions including perceptual modality (e.g., visual, auditory), information code (e.g., verbal, spatial) and processing stage (e.g., encoding, central processing, and responding). Multiple resource theory parallels that of Baddeley's model of working memory [4] which also includes a spatial and verbal component. In fact there is some suggestion that resource theory is essentially synonymous with working memory [5]. Mental resource capacity can be conceptualized as essentially how much information can be maintained and manipulated in working memory. Theoretical conceptions of both constructs make the assumption that capacity is in some way limited and that a given task or set of tasks can exceed that capacity. Individual differences in working memory capacity are consistently found and these differences are strongly correlated with performance on a number of different cognitive tasks [6-8]. Because working memory capacity affects both the difficulty of and the strategies used for learning complex tasks as well as the susceptibility to different forms of distraction [9-11], its assessment may provide a powerful tool for improving existing training protocols.

The agreement between the two theories is that an individual's mental capacity is in some way limited and that a given task or set of tasks can exceed that capacity. A cognitive state gauge that provides a real time indicator of an individual's available resources could provide great insight into performance, task design and training. However, moving from a physiological signal filled with noise to a real time gauge requires a significant amount of signal processing.

2.3 EEG Algorithms

A cognitive state gauge that provides a real time indicator of an individual's current level of engagement and the availability of different types of resources could provide great insight into performance, task design and training. However, improving methods of signal amplification, filtering, and the analog to digital conversions required to extract physiological signals from the background of noise is an on-going challenge to the successful implementation of real time cognitive state gauges.

Linear Classification. Early real time metrics derived from EEG analyzed the changes of spectral power in the five frequency bands (alpha, beta, theta, gamma, and delta). These changes were used to provide an indication of the operator's engagement, attention and mental workload. Pope, Bogart, and Bartolome [12] developed an index based on a ratio of EEG power, defined as (beta/(alpha + theta)) which can be computed in real-time by calculating a running average over a 20 second window. This index was said to determine the level of engagement/alertness of an individual while performing a task. The researchers were able to demonstrate the index could be used in real time to improve performance on a vigilance task [13] and a complex tracking task [14].

A second linear algorithm for processing cognitive state from EEG is the eXecutive Load Index (XLI) [15], which was designed to monitor changes in cognitive load related to processing messages in real-time. This was done by computing the ratio of ((delta+theta)/alpha) over a moving 2 second window, with the change determined by comparing the value to the previous 20 second running average.

Researchers at Advanced Brain Monitoring (ABM) developed several gauges of cognitive state based upon linear and quadratic discriminant function analysis (DFA) [16]. The gauges for mental workload and engagement are of particular interest. The index for engagement tracks the demands for sensory processing and attentional resources, whereas the index for mental workload tracks the level of cognitive function and is considered to be a correlate of executive function. The algorithms for both indexes are derived for each individual based upon his or her EEG signals on a series of baseline vigilance tasks. The measures have both been validated in a series of basic cognitive tasks. The mental workload metric has been shown to track task demand in mental arithmetic and digit span tasks as well as show a significant correlation with subjective measures of workload and task performance. The gauge for engagement has been shown to decrease as a function of time during a vigilance task whereas workload did not. The algorithms for both engagement and workload output data every second.

Non-linear classification. Research at the Air Force Research Lab has investigated the ability of an Artificial Neural Network (ANN) to classify operator mental workload in a complex laboratory task and during a UAV simulation [17, 18]. The ANN derives its classification from EEG, EOG, and heart rate data and was successfully able to classify high versus low workload with a 85-90% accuracy rate when the ANN was trained for each individual. The ANN was also successfully implemented in an adaptive automation UAV paradigm where vehicle speed was reduced during periods of high workload. The adaptive automation system was able to significantly improve performance over both a non-adaptive system and a system with random changes to the speed. Although ANN has been shown to be highly successful, it requires a large amount of data to "train" the model. It is also unclear how stable an ANN would be for a particular individual over time.

2.4 Eye Tracking

Visual scanning strategies may provide an indication of mental workload. Di Nocera, et al. [19] implemented the Nearest Neighbor Index (NNI) to investigate whether a statistical index that provides information on dispersion of points, or fixations, would have differential patterns for high workload and low workload conditions. The index is based on the Complete Spatial Randomness (CSR) method, which is the spatial analysis equivalent of uniformly and independently distributed random variables. The index is computationally straightforward and is feasible to compute in near real-time, which lends potential to be used as a metric or trigger for adaptive training. Essentially, higher values in the NNI show higher levels of entropy in scanning. Preliminary analysis from a case study showed that higher NNI values were correlated with higher workload, and that the NNI was sensitive to varying workload conditions. However it was suggested that more studies be performed to fully understand the correlation between the randomness of fixations and mental workload.

Cognitive workload has also been evaluated using measures of eye movement and pupil dilation to detect cognitive strategy shifts [20]. A psycho-physiological index of workload based on pupil dilation, the Index of Cognitive Activity (ICA), was used in a case study by Marshall et al. to detect shifts in strategy based upon large changes of ICA. ICA does not require averaging over trials or individuals, it can be applied to a signal of any given length, and it can be computed in near real-time. ICA is calculated as the frequency of a detection of an abrupt discontinuity in the pupil signal [21]. Marshall, et al.'s study demonstrated that cognitive strategy shifts can be identified from eye tracking data, and observed fluctuations of ICA can identify the time and location of those strategy shifts. Identification of cognitive strategy shifts may be beneficial not only for instructional design based on cognitive load theory [22], but also for adaptive training.

A recent review identified and evaluated the ability of seven eye tracking metrics to classify an operator's cognitive state, while taking into account the sensitivity and specificity of the classification [23]. The metrics under evaluation included the Index of Cognitive Activity (ICA), blinks, movement, and divergence between eyes, where separate right and left eye values were calculated for the ICA, blink, and movement metrics. Each of the seven metrics can be computed in near real time, making them attractive candidates to apply and incorporate into adaptive training applications.

For statistical analysis, all metrics were transformed to a common scale ranging from 0 to 1. Two classification models, linear discriminate function analysis and non-linear neural network analysis were employed, and the sensitivity and specificity were evaluated to determine classification adequacy. Two-state classifications were calculated for three separate studies (problem solving, driving simulation, and visual search) to differentiate between an engaged or relaxed state, focused or distracted driving, and a fatigued versus alert state. For all three studies, both classification models were successful in differentiating cognitive states (69% to 92%) based solely upon the aforementioned eye metrics. Discriminant analyses with systematic elimination of each metric were conducted to confirm that all metrics were needed to obtain the same accuracy of results. In addition, it was determined that all metrics were needed to obtain the level of demonstrated classification, and that no particular metric was salient across all subjects within any study.

3 Cognitive Load Theory

Cognitive load theory (CLT) is model of learning based around components of human information processing, particularly working memory and long term memory. A core principle of CLT theory is that learning places demands on a limited capacity working memory system [24, 25]. Since working memory capacity (WMC) is limited [26, 27], learning is integrally tied to both the working memory capacity (WMC) of the learner and the working memory demand of the instruction and instructional material. While the WMC of an individual is limited, his or her long term memory is almost unlimited. Thus the theory is concerned with how information from working memory is organized and grouped together (into schemata) and stored in long term memory. Once information is stored in long term memory, it enables the individual to access it later and reduces the burden placed upon the working memory system. Much of the research on CLT has focused on working memory since it serves as the bottleneck to learning.

CLT proposes three specific types of cognitive load with additive effects; the sum total of which must not exceed a learner's working memory resources if optimal learning is to be achieved. The first, termed *intrinsic cognitive load* is the difficulty imposed by the material or task to be learned. It is heavily influenced by the elemental interactivity of the material – how many interacting elements must be maintained in working memory at any given time. Complex material may have high elemental interactivity. The more inherent elemental interactivity, the higher the cognitive load. Often there is little if anything that instructional design can do to change the intrinsic cognitive load of the material or task to be learned. As expertise develops, schemas are formed and elements become grouped together; enabling the individual to deal with more elements simultaneously and allowing them to overcome the working memory bottleneck. This process reflects learning and the dynamic nature of intrinsic cognitive load within CLT. The number of elements which make up intrinsic load are based upon the individual's ability to group them together.

The second form of cognitive load is *extraneous cognitive load* and this is where instructional design has the potential to make vast improvements in training. Extraneous load typically refers to how the information is presented, e.g., graphically versus verbally. Ineffective instructional designs impose an additional level of extraneous cognitive load, which is particularly problematic when the intrinsic load is high. Much of CLT has emphasized reducing extraneous load as a method of reducing overall load and enabling learning.

Germane cognitive load, the third type, is the process of creating and organizing information into schema. Germane load is the result of the instructional design [28]. It promotes the development of accurate mental models of the task and relevant schemas as well as facilitating the transition from controlled to automatic processing that accompanies expertise. Germane load is influenced by the manner, modality and sequence in which the material is presented and the learning activities involved. Differential sensitivity has been observed between various measures for each type of load [29].

Within the framework of cognitive load theory, intrinsic cognitive load is set by the task, and extraneous cognitive load is typically manipulated through instructional design. This ensures an individual's cognitive resources are not being exceeded, and that learning is promoted. Instructional design that reduces extraneous cognitive load frees more resources for germane cognitive load thus facilitating the development of schema acquisition and a shift toward automatic processing and expertise.

4 Adaptive Training

Although traditionally cognitive load theory has focused on adjusting extraneous load simulation based training allows for an adjustment in the amount of intrinsic cognitive load presented at a given time. This ability to manipulate intrinsic load combined with the capability to measure working memory capacity in real time provides the basis for the development of an adaptive training approach.

Combining elements from Wickens' Multiple Resource Theory [3] which is designed to describe workload and ultimately help predict performance and Sweller's CLT [24] we have developed an initial throughput model of learning that can be used in a closed loop system. Within the adaptive training paradigm the intrinsic load acts as the input into the system. How that information is presented in terms of modality and processing code produces the extraneous load. As with the traditional model of CLT presenting information in a spatial code versus a verbal code can yield different amounts of extraneous load depending on the task/information. Following multiple resource theory and CLT it would be possible to reduce extraneous load by using different modalities and processing codes.

As with CLT the germane load is still the organization of information into schema. The overall load is still the combination of intrinsic, extraneous, and germane load. However, as with multiple resource theory there are now potentially multiple capacities. Figure 1 represents a simple throughput model of how adaptive training would work. Operator capacity would be assessed in real time based upon a real-time physiological metric described above. Presently their may not be a separate metric for each potential resource pool. However an overall gauge of spare capacity could still serve to trigger an adaptive training screen. Based upon whether the physiological metrics indicate spare capacity the screen could add or remove certain elements of

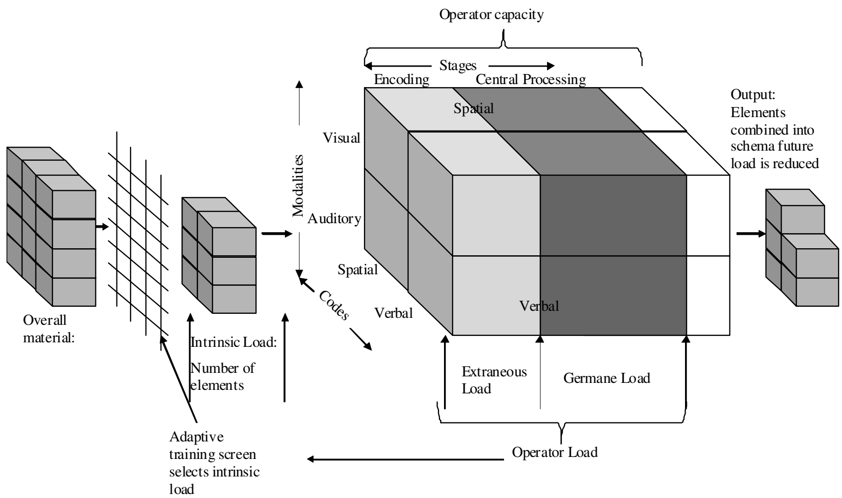

Fig. 1. Model of Adaptive Training based upon Wickens' Multiple Resource Theory and Sweller's Cognitive Load Theory

intrinsic load or add/reduce the size of a single element (e.g., driving at high speed versus low) from the overall material to be learned.

For example, in a computer simulation designed to teach target identification, the cognitive load of the task can be manipulated by adjusting the number of targets presented at a given time, target speed or salience. These manipulations do not change the intrinsic load of the task, per se. However, the load imposed on the novice learner is adjusted. Extraneous load may be reduced by making the targets more visible, by using an auditory modality to supplement identification of salient aspects of the visual targets, etc.

Germaine cognitive load, which supports skill development, can be increased by presenting various targets to be identified in random order rather than for instance all enemy tanks and then all enemy helicopters. The proposed adaptive training system would manipulate the amount of overall cognitive load in the scenario based on assessment of the trainee's current expenditure of mental resources. When physiological metrics indicate that mental workload is high, for example enemy units could slow down or decrease in number. Alternatively, when a trainee displays a low level of mental resource utilization, the scenario can be made more difficult. Changes to the cognitive load in the adaptive training system must be based not only on an individual's available resources but also their level of expertise with the system. There is a dynamic relationship between level of expertise and cognitive load required by the task.

The primary difference between a novice and an expert in any given task domain hinges on two things. First, the expert has an extensive knowledge base of well developed relevant schemas held in long term memory. Schemas allow a person to treat multiple elements as one item. For example, an expert chess player has literally thousands of schemas for movement patterns stored in long term memory. Secondly, for the expert many of the relevant tasks and skills as well as access to the stored schemas are automatic, no longer requiring resource demanding controlled processing [30]. Novices have neither extensive or well developed schemas, nor can they initiate many of the task components automatically [31]. Problem solving routines and access to schemas can become automated as when one automatically knows to solve the algebraic equations within the brackets before moving on to the relationships between the bracketed and non-bracketed items. Individual features of letters, nor even individual letters need to be processed once a reader has developed sufficient skill. As schemas develop and tasks become automated, working memory load is reduced and learning is accelerated. As learning is accelerated the amount and or rate of information presentation can *and should* be accelerated. Monitoring the transition from novice to expert is essential for efficient learning and is a key element of the proposed adaptive training strategy.

As with CLT the adaptive training paradigm will recognize the task-learner interaction, or expertise reversal effect, meaning that as a learner develops expertise, the methods of instruction that are effective should change [28, 32]. Optimal training protocols must continuously monitor in real-time both the working memory resources being utilized or mental workload of the learner and changes in skill level associated with developing expertise. Under or over utilization of working memory processes or a mismatch with the learner's current skill level will result in less efficient learning.

5 Conclusions

It is believed that the proposed adaptive training model will be able to significantly improve learning by eliminating the time in which the learner is not in an optimal state as determined by their working memory capacity. An adaptive training system will be capable of reducing the intrinsic load when working memory capacity is exceeded, or adding to the intrinsic load when there is sufficient reserve working memory capacity. Additionally the new model allows for a diagnostic approach to implementing the adaptive training screen. Advances in the real time sensors may eventually be capable of assessing capacity within the different pools (i.e., spatial versus verbal) and therefore allow for more specific changes to the material being presented. Eventually such a system will be capable of moving between information codes and processing modalities of the information being presented to capitalize on an individual's multiple resources.

References

1. Parasuraman, R.: Neuroergonomics: Research and Practice. Theoretical Issues in Ergonomics Science 4, 5–20 (2003)
2. Parasuraman, R., Wilson, G.F.: Putting the Brain to Work: Neuroergonomics Past, Present, and Future. Human Factors 50, 468–474 (2008)
3. Wickens, C.D.: Processing resources in attention. In: Parasuraman, R., Daives, D.R. (eds.) Varieties of attention, pp. 63–102. Academic Press, New York (1984)
4. Baddley, A.D.: Working Memory. Oxford University Press, Oxford (1986)
5. Parasuraman, R., Caggiano, D.: Neural and genetic assays of mental workload. In: McBride, D., Schmorrow, D.D. (eds.) Quantifying Human Information Processing Rowman and Littlefield, Lanham, MD (2005)

6. Conway, A.R.A., Engle, R.W.: Individual Differences in Working Memory Capacity: More Evidence for a General Capacity Theory. Memory 4(6), 577–590 (1996)
7. Brumback, C.R., Low, K.A., Gratton, G., Fabiani, M.: Putting Things into Perspective: Individual Differences in Working-Memory Span and the Integration of Information. Experimental Psychology 52(1), 21–30 (2005)
8. Unsworth, N., Engle, R.W.: The Nature of Individual Differences in Working Memory Capacity: Active Maintenance in Primary Memory and Controlled Search From Secondary Memory. Psychological Review 114(1), 104–132 (2007)
9. Cantor, J., Engle, R.W.: Working-memory capacity as long-term memory activation: An individual-differences approach. Journal of Experimental Psychology: Learning, Memory, and Cognition 19(5), 1101–1114 (1993)
10. Sobel, K.V., Gerrie, M.P., Poole, B.J., Kane, M.J.: Individual differences in working memory capacity and visual search: The roles of top-down and bottom-up processing. Psychonomic Bulletin & Review 14(5), 840–845 (2007)
11. Graf, S., Lin, T., Kinshuk: The relationship between learning styles and cognitive traits–Getting additional information for improving student modelling. Computers in Human Behavior 24(2), 122–137 (2008)
12. Pope, A.T., Bogart, E.H., Bartolome, D.S.: Biocybernetic system evaluates indices of operator engagement in automated task. Biological Psychology 40(1), 187–195 (1995)
13. Mikulka, P.J., Scerbo, M.W., Freeman, F.G.: Effects of a biocybernetic system on vigilance performance. Human Factors 44, 654–664 (2002)
14. Freeman, F.G., Mikulka, P.J., Scerbo, M.W., Prinzel, L.J., Clouatre, K.: Evaluation of a psychophysiologically controlled adaptive automation system, using performance on a tracking task. Applied Psychophysiology and Biofeedback 25(2), 103–115 (2000)
15. DuRousseau, D.R., Mannucci, M.A.: eXecutive Load Index (XLI): Spaital-frequency EEG tracks moment-to moment changes in high-order attentional resourcesFoundations of Augmented Cognition, pp. 245–251. Lawrence Erlbaum Associates, Mahwah (2005)
16. Berka, C., Levendowski, D., Lumicao, M.N., Yau, A., Davis, G., Zivkovic, V.T., Olmstead, R.E., Tremoulet, P., Craven, P.L.: EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. Aviation, Space, and Environmental Medicine 78, B231–B244 (2007)
17. Wilson, G.F., Russell, C.A.: Real-Time Assessment of Mental Workload Using Physiological Measures and Artifiicial Neural Networks. Human Factors 45, 635–643 (2003)
18. Wilson, G.F., Russell, C.A.: Operator Functional State Classification Using Multiple Psychophysiological Features in an Air Traffic Control Task. Human Factors 45(3), 381–389 (2003)
19. Di Nocera, F., Terenzi, M., Camilli, M.: Another look at scanpath: Distance to nearest neighbour as a measure of mental workload. In: de Waard, D., Brookhuis, K.A., Toffetti, A. (eds.) Developments in human factors in transportation, design, and evaluation, pp. 295–303. Shaker, Maastricht (2006)
20. Marshall, S.P., Pleydell-Pearce, C.W., Dickson, B.T.: Integrating psychophysiological measures of cognitive workload and eye movements to detect strategy shifts. In: Proceedings of the 36th Annual Hawaii International Conference on System Sciences, vol. 6 (2003)
21. Marshall, S.P.: The Index of Cognitive Activity: measuring cognitive workload. In: Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants, pp. 7-5–7-9 (2002)
22. van Gog, T., Kester, L., Nievelstein, F., Giesbers, B., Pass, F.: Uncovering cognitive processes: Different techniques that can contribute to cognitive load research and instruction. Computers in Human Behavior (25), 325–331 (2009)
23. Marshall, S.P.: Identifying Cognitive State from Eye Metrics. Aviation, Space, and Environmental Medicine 78(5, Section II), B165–B175 (2007)

24. Sweller, J.: Cognitive load during problem solving: Effects on learning. Cognitive Science: A Multidisciplinary Journal 12(2), 257–285 (1988)
25. Sweller, J.: Discussion of Emerging Topics in Cognitive Load Research: Using Learner and Information Characteristics in the Design of Powerful Learning Environments. Applied Cognitive Psychology 20(3), 353–357 (2006)
26. Miller, G.A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review 63, 81–97 (1956)
27. Kahneman, D.: Attention and effort. Prentice Hall, Englewood Cliffs (1973)
28. Paas, F., Renkl, A., Sweller, J.: Cognitive load theory and instructional design: Recent developments. Educational Psychologist 38(1), 1–4 (2003)
29. DeLeeuw, K.E., Mayer, R.E.: A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. Journal of Educational Psychology 100(1), 223–234 (2008)
30. Shiffrin, R.M., Schneider, W.: Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review 84(2), 127–190 (1977)
31. Marcus, N., Cooper, M., Sweller, J.: Understanding instructions. Journal of Educational Psychology 88(1), 49–63 (1996)
32. Ayres, P.: Impact of Reducing Intrinsic Cognitive Load on Learning in a Mathematical Domain. Applied Cognitive Psychology 20(3), 287–298 (2006)