

Attentional Blink: An Internal Traffic Jam?

Niels A. Taatgen (taatgen@cmu.edu)

Ion Juvina (ijuvina@cmu.edu)

Department of Psychology, Carnegie Mellon University
5000 Forbes Av., Pittsburgh PA 15213

Seth Herd (sethherd@grey.colorado.edu)

Department of Psychology, University of Colorado
Boulder CO 80309

David Jilk (djilk@jilk.com)

eCortex, Inc., Boulder CO

Sander Martens (s.martens@med.umcg.nl)

NeuroImaging Center, University Medical Center Groningen
Antonius Deusinglaan 2, 9713 AW, Groningen, Netherlands

Abstract

In cognitive models, cognitive control can be measured in terms of the number of control states that are used to do the task. In most cases more control leads to better performance. Attentional Blink is an example in which the opposite is true: more control leads to poorer performance. A hybrid ACT-R/Leabra model is used to model both high- and low-control participants using two and one control states, respectively.

Keywords: Cognitive Control, Multi-tasking, ACT-R, Leabra

Introduction

The term *cognitive control* is used to refer to cognitive processes that help us focus on our goals and plans, and prevent external stimuli and events from interfering with them. Most of the time a higher level of control improves performance on tasks. However, too much control can make behavior brittle and less flexible. It is therefore likely that cognition strives for a level of control that is just enough for proper performance (the minimal control principle, Taatgen, 2005, 2007). In this paper we will discuss a task that demonstrates that too much control can hurt performance. The task is a Rapid Serial Visual Presentation (RSVP) task (Raymond, Shapiro, & Arnell, 1992).

In RSVP tasks, participants are presented with rapid streams of visual stimuli. Each of these streams contains 0, 1 or 2 targets that the participants have to identify. In the version that we will discuss in this paper, streams consist of 20 characters that are presented at a rate of 100ms/character. The targets are letters, while distracters are digits. The streams of interest are the ones with two targets. In these streams the time between the two targets is of importance, usually referred to by the *lag*. A lag of 1 means the two targets appear in sequence, and are 100 ms apart in time (given our presentation rate), lag 2 means that there is one distracter in between the targets, etc. Example sequences are:

Lag 1: 49204039GF3434329237

Lag 3: 0230349023Y94D324294

Lag 9: 9430R32305129K235209

The phenomenon of interest is that on lags 2-5, but mainly on lag 2 and 3, accuracy on the second target is much lower than on the first target. On other lags, including lag 1, the accuracies are the same. This phenomenon is referred to as *Attentional Blink*. The interesting aspect from the viewpoint of cognitive control is that there are indications that *less* control improves performance. Certain experimental manipulations can decrease the amount of blink people exhibit, for example if the stimuli are presented in a star field (Arend, Johnston & Shapiro, 2006), when music is played in the background (Olivers & Nieuwenhuis, 2005), or when participants receive instructions to focus less on the task (Olivers & Nieuwenhuis, 2006). In addition, there are strong individual differences in attention blink: some individuals do not exhibit attention blink at all (Martens, Munneke, Smid & Johnson, 2006).

Nieuwenhuis, Gilzenrat, Holmes and Cohen (2005) have modeled the attentional blink using a neural network. Their model consists of three layers, input, decision and detection, and a cell representing the Locus Coeruleus (LC), which is connected to the decision and detection layers. The LC provides extra activation to these layers, making them more sensitive to targets. Once the decision layer has decided that an input is a target, it is stored in the detected layer, but it also sends a signal to the LC. The result of the signal to the LC is that its contribution to activating the decision layer temporarily diminishes, decreasing the detection rate of targets that appear within 200-300 ms. This decrease is relatively slow, so it has no effect on the lag 1 trials.

Although the Nieuwenhuis et al. model is successful in predicting the outcome of several new experiments, the impact of control is outside the scope of that model. An additional finding in RSVP experiments is that in lag 1 trials, the order in which the targets are reported is often

reversed, while this almost never happens in any of the other lags (Hommel & Akyürek, 2005).

Both phenomena, reduced control leads to less blink, and the reversed report of lag 1 targets, are outside the scope of the Nieuwenhuis et al. model, because it neither incorporates higher-level aspects of control nor the fine-level details of perception. In this paper we will present a model that encompasses both. To incorporate both the fine details of perception and the higher-level control aspects, we used the hybrid model that combines the Leabra neural network architecture (O'Reilly & Munakata, 2000) and the ACT-R architecture (Anderson, 2007). More specifically, visual perception is handled by a Leabra model, which passes on the information to the visual input buffer of ACT-R. ACT-R takes care of the classification of the symbol, and storing it, if it is a target.

Experiment

Method

Forty-one volunteers from the Carnegie Mellon student population participated in this experiment, which was part of a larger experiment on individual differences in cognitive control. The larger experiment included three other tasks that we selected to assess levels of cognitive control: the N-Back memory task (McElree, 2001), the abstract decision making task (ADM task, Joslyn & Hunt, 1998), and a dual-tasking task (DTT task, Taatgen, van Rijn & Anderson, in press). In the N-Back task, participants were shown sequences of letters, and they had to detect repetitions of letters and judge how many letters back that repetition was. In the ADM task, participants had to ask questions about properties of objects, and sort the objects into bins once they had obtained enough information. In the DTT task, participants had to do two visual tasks and a time estimation task in parallel. Four participants had to be excluded from the dataset due to a problem in the experimental software.

The stimuli consisted of sequences of 20 characters. Distracters were digits from 2 to 9, and targets were C, D, F, G, H, I, J, K, L, M, N, P, R, T, V, W and X. No consecutive characters were identical, and if there were two targets, they were different from each other. Targets were never in the first four positions in the sequence, nor in the last four.

The experiment consisted of 6 practice trials, and 4 blocks of 37 experimental trials. Each block of experimental trials consisted of 5 zero-target trials, 5 one-target trials, and 27 two-target trials. The two-target trials consisted of 3 trials of each 9 different lag lengths (1 through 9).

Each trial was preceded by a 500 ms fixation point, after which symbols in the sequence were presented one at a time for 100 ms each. At the end of the sequence, participants were asked to type all the targets (if any) they had seen, and press enter. The next trial started immediately after the participant had pressed enter.

Results

Figure 1 shows the correctness on the second target by lag. There is a clear blink effect in lags 2-4, consistent with many earlier findings. On 16% of the Lag 1 trials, both targets were reported correctly, but in the wrong order.

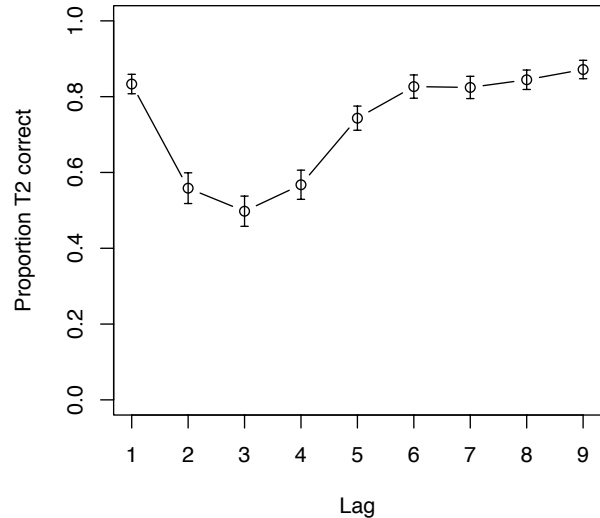


Figure 1: Blink by Lag. Bars are standard errors.

In order to assess individual differences, we calculated the AB magnitude (based on Martens et al., 2006) according to the following equation:

$$\left(\frac{T1_{lag2} - T2_{lag2} | T1_{lag2}}{T1_{lag2}} + \frac{T1_{lag3} - T2_{lag3} | T1_{lag3}}{T1_{lag3}} \right) / 2$$

In this equation, T1 means proportion correct identification of the first target (given the indicated lag), and T2|T1 means correct identification of the second target given a correct identification of the first target. An AB magnitude of 0 or less means there is no blink at all, while higher values imply more blink. We correlated the AB magnitude with outcomes of the other tasks that are associated with control. For the DTT task, the outcome most associated with control is how often participants make a response on the time estimation task, the one of three tasks that is not cued by visual input. Low levels of response are an indication of low-level control. The response rate on the DTT task and the AB magnitude correlated with $r=0.46$ ($p=0.005$), indicated that more control in the DTT task implied a higher blink magnitude. Similarly, the score on the N-Back task and the AB magnitude had a correlation of $r=0.42$ ($p=0.009$). The correlation with the ADM task was only weak: $r=0.29$ ($p=0.08$). All of these correlations suggest that a higher level of control correlates with a higher AB magnitude.

For the purposes of identifying non-blinkers in the dataset, we classified participants with a AB magnitude of 0.2 or lower as non-blinker, and participants with 0.2 or

higher as blinker (consistent with Martens et al., in press), amounting to 8 non-blinkers and 29 blinkers.

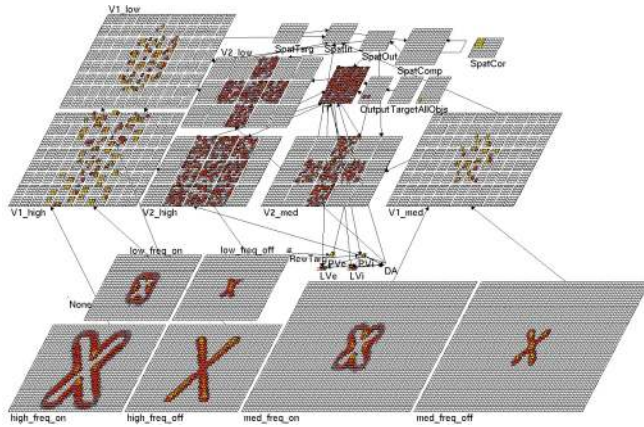


Figure 2: Visual input module

Model

Overview

Cognitive models allow a more precise characterization of what “more control” means. To keep outside events from completely controlling behavior, we maintain internal goals. The current goal can be in a certain state to keep track of progress of that goal. Any action or progress on the goal can change the state, or keep it as it is. Taken together, states and possible actions create a state space. The more states there are in the state space, the more it is associated with a higher level of control (Taatgen, 2007).

We designed two possible control structures for a model of attentional blink, one with two states, which models the blinkers, and one with one state, which models the non-blinkers. The first, blinker, model assumes two control states. One state is used to signify the model is searching for a target in the input stream. Once a target has been found, the model switches to a second state that is used to consolidate the target in memory. When the target is

consolidated, the state switches back to the first state. When the model is in the consolidation state, it no longer fully processes the input stream. When the model switches back to the search state it has too much to do at the same time, creating an internal “traffic jam” that causes the model to sometimes miss targets in the 200-500ms range. The second, non-blinker, model uses only a single state. In other words, there is no state to protect the consolidation process, but it also misses targets less often.

Vision

The visual input is projected on the input layer of a Leabra (O’Reilly & Munakata, 2000) neural network of the ventral visual stream (Figure 2). This model processes the input, and arrives at a classification in the output layer, in which it has a single cell for each of the possible symbols in the input. Because of the speed in the visual presentation, the network is not always able to fully settle, and there is often still residual activation of the previous symbol. Figure 3 shows an example of how activation in the output layer changes over time based on the “2829P” sequence. The consequence of the rising and falling of activations is that if the visual input is sampled at a particular moment, it is possible two output cells are active, in which case it is impossible to determine in which order the two have been presented. This explains why in lag 1 the two targets are often reported in the wrong order.

Central Cognition and Control

The ACT-R architecture is structured as a set of interacting modules. Modules communicate through buffers, but otherwise operate asynchronously. Each module can only work on one thing at a time, but because all the modules work in parallel, the cognitive system as a whole can work on several tasks at the same time. At this level of abstraction, cognitive control is a matter of optimally engaging all modules in doing the task or tasks.

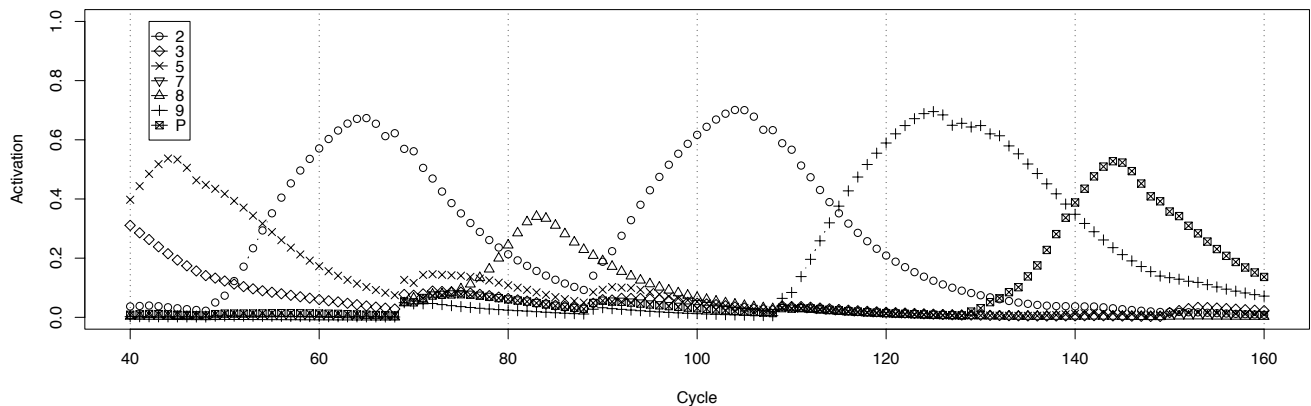


Figure 3: Example of visual module output for the “2829P” sequence. Vertical lines indicate where a new stimulus is presented, e.g. at cycle 40 the symbol “2” is presented. Each cycle corresponds to 5ms real time.

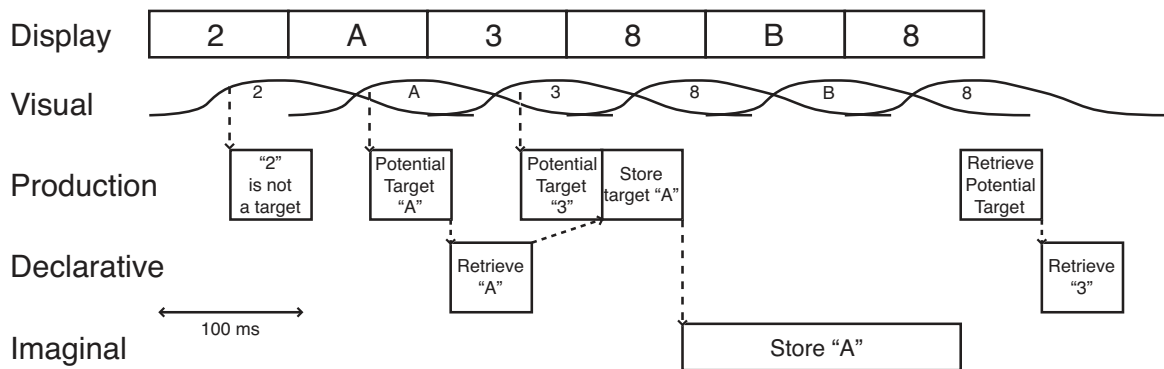


Figure 4: Time diagram of a Lag 3 presentation: blinker model.

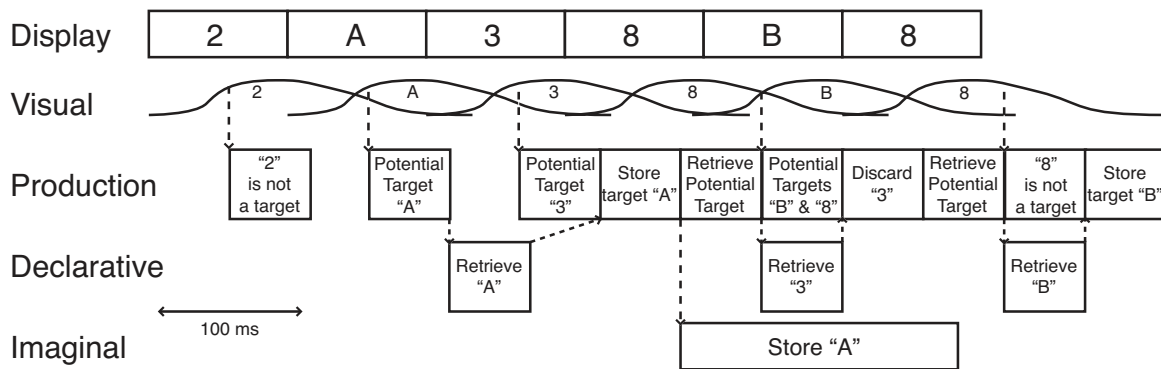


Figure 5: Time diagram of a Lag 3 presentation: non-blinker model.

In the attentional blink tasks, several modules have to operate in parallel. The visual module has to scan the incoming stream of characters. These characters have to then go through a decision process, which determines whether a character is a target, which involves the procedural and declarative modules. Once a target is found, it has to be consolidated in ACT-R's imaginal buffer, a place to temporarily store problem-related information.

Figure 4 shows a diagram of how the model operates in the case of a Lag 3 sequence. Each row in the diagram represents one of ACT-R's modules, with the exception of the top line, which shows the state of the display. Boxes in each row show activity in a module at a particular time, and the width of the box indicates the duration. The visual module follows the input, and outputs activations corresponding to the classification of the input. This input triggers production rules that try to determine whether the character is a target. The assumption of the model is that some characters can be recognized as non-targets straight away (with a probability of 60%). Other characters have to be classified by a declarative memory retrieval. In the example, the first "2" is immediately recognized as non-target. The second character ("A") is a target. Targets are always retrieved from memory to verify that they are targets. During this retrieval, the next character ("3") is

identified as a potential target. However, the memory retrieval that can verify this has to be postponed until the retrieval of the "A" finishes. However, once the retrieval of "A" finishes, the model decides "A" is indeed a target, and stores it in the imaginal buffer. In this version of the model, once a target has been found the production rule that initiates storing the target changes the control state to consolidate, blocking further processing of the input. Only when the imaginal buffer is done storing the target can the target detection process resume. As a consequence, the "8" and the "B" will not be considered as targets, leading to a blink trial in this particular example. Also note that the "3" that directly followed the first target is considered a potential target by the model, although delayed, which means that if this had been a Lag 1 trial and the "3" would have been a target, both targets would have been detected.

The non-blinker model

Although the model described above changes state to protect its memory consolidation, this is an unnecessary exertion of control. If the model does not change state when a target has been detected, detection of the second target can proceed while the first target is consolidated. Figure 5 shows an example trace of that variation of the model. Once the target "A" has been detected and transferred to the

imaginal buffer, the next candidate, “3” is requested from declarative memory. The next production rule samples this visual input again, now detecting two potential targets (“B” and “3”), because both output cells of the neural network are active.

Reversal of target on Lag 1

Figure 6 demonstrates how reversals can occur. Delayed by the possibility that “2” is a target, the model samples the visual buffer at a moment when both the “A” and “B” cells are active. Having no means to determine the order of the two, the model decides to retrieve the “B” first.

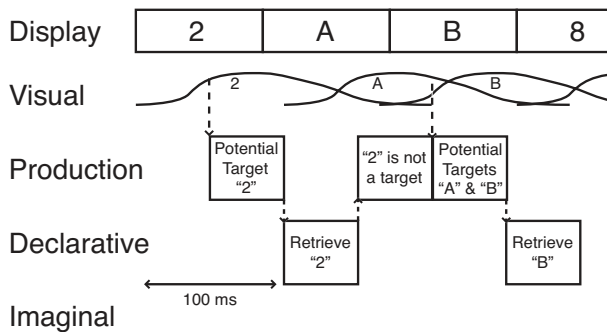


Figure 6: Example of a reversal of targets

Parameters

The model has a number of parameters that influence its outcome. The main parameter that was used to fit the model was the probability that a character could be recognized as a foil straight away (60%). Other parameters that will probably influence the outcome when changed are the latency factor that determines how long declarative retrievals take (set so that they take about 50ms), and the time to store an item in the imaginal buffer (left at the default of 200ms). In addition, the Leabra model has several parameters, but those were left untouched for the purpose of fitting the data.

Model Results

To assess the model we divided the dataset in blinkers and non-blinkers according to the criteria discussed earlier. Figure 7 shows the model/data comparison for the blinker group ($r=0.97$, $MSE=0.03$), and Figure 8 for the non-blinker group ($r=0$, $MSE=0.04$). The blinker model has a good correlation between model and data and a low MSE. The non-blinker model has no correlation with the data because there is no meaningful variability in the data to model. However, the MSE is comparable to that of the blinker model, confirming that there is a decent fit.

Figure 9 shows the proportion of trials in which the two targets were reported in the wrong order for both groups together (there was no difference between blinkers and non-blinkers).

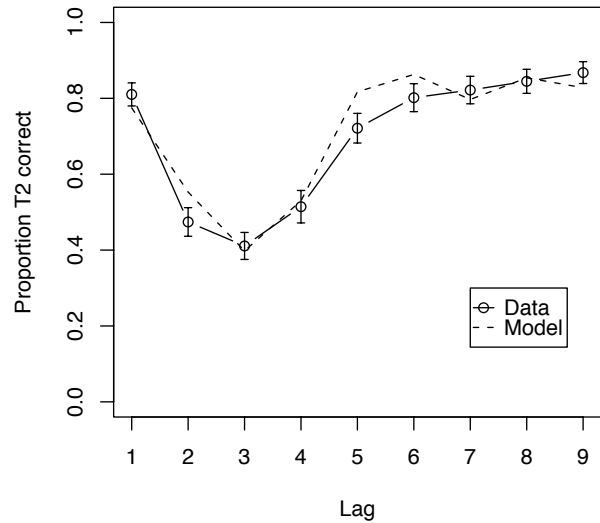


Figure 7: T2 correctness for blinkers (n=29), model/data comparison.

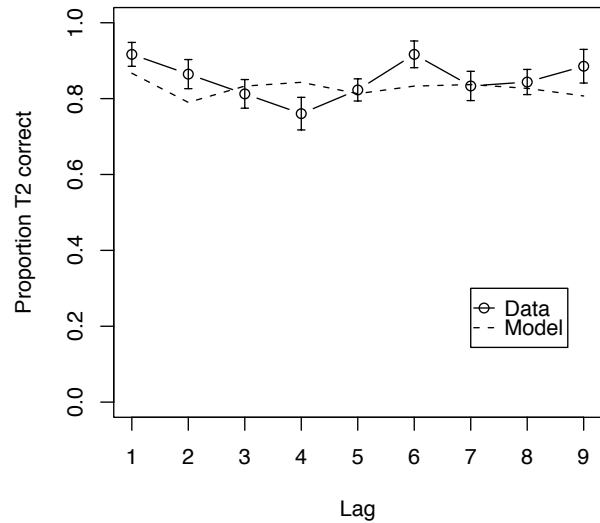


Figure 8: T2 correctness for non-blinkers (n=8), model/data comparison.

Discussion

The central question that we tried to address in this paper is how cognitive control can be made more concrete in terms of a cognitive model. Our hypothesis is that more control is associated with more possible control states. More control states gives more top-down control of the task’s execution, but at the cost of flexibility. The experiment as a whole showed that the amount of blink in the RSVP task correlated with control aspects of other tasks, which is consistent with findings by Arend et al. (2007) and Olivers and Nieuwenhuis (2006) that attentional blink is related to control factors.

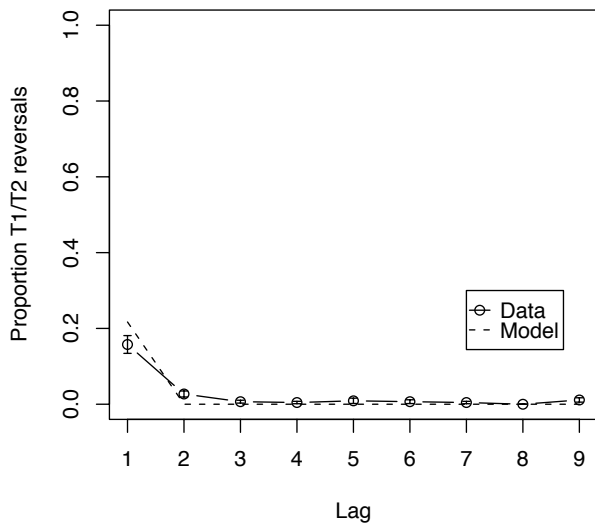


Figure 9: Proportion in which targets are reported in the wrong order, model/data comparison

The RSVP model, with two control states for high control and one control state for low control, managed to fit the data very well. Further support for the model can be found in ERP data. Martens et al. (2006) collected ERP data for both blinkers and non-blinkers. They found that the P300, which we associate with imaginal buffer activity, is present only if a target is detected. Moreover, it is later for blinkers than for non-blinkers. This is the case for both T1 and T2, but the effect on T2 is much larger. Our model currently only predicts a difference on T2. The T1 difference may be due to a different factor all together, and may be related to the proportion in which distracters can be dismissed without memory retrieval.

This model is also a demonstration of how a symbolic architecture, ACT-R, and a neural network architecture, Leabra, can work together. The perceptual part of the model was clearly outside the current capabilities of ACT-R, while the control aspects were outside of Leabra's scope.

Acknowledgments

This work was supported by Office of Naval Research grant N00014-06-1-005 and N00014-05-1-0880, and Biologically-Inspired Cognitive Architecture, DARPA.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford university press.

Arend, I., Johnston, S., & Shapiro, K. (2006). Task-irrelevant visual motion and flicker attenuate the

attentional blink. *Psychonomic Bulletin & Review*, 13(3), 600-607.

Hommel, B., & Akyürek, E. G. (2005). Lag-1 sparing in the attentional blink: benefits and costs of integrating two events into a single episode. *Quarterly Journal of Experimental Psychology*, 58A, 1415-1433.

Joslyn, S., & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. *Journal of Experimental Psychology: Applied*, 4(1), 16-43.

Martens, S., Munneke, J., Smid, H. & Johnson, A. (2006). Quick minds don't blink: electrophysiological correlates of individual differences in attentional selection. *Journal of Cognitive Neuroscience*, 18(9), 1423-1438.

McElree, B. (2001). Working Memory and Focal Attention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 27(3), 817-835.

Nieuwenhuis, S., Gilzenrat, M. S., Holmes, B. D., & Cohen, J. D. (2005). The role of the locus coeruleus in mediating the attentional blink: a neurocomputational theory. *Journal of Experimental Psychology: General*, 134(3), 291-307.

Olivers, C. N. L., & Nieuwenhuis, S. (2005). The Beneficial Effect of Concurrent Task-Irrelevant Mental Activity on Temporal Attention. *Psychological Science*, 16(4), 265-269.

Olivers, C. N. L., & Nieuwenhuis, S. (2006). The beneficial effects of additional task load, positive affect, and instruction on attentional blink. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2), 364-379.

O'Reilly, R. C., & Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience* Cambridge, MA: MIT Press.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in a RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance*, 18, 849-860.

Taatgen, N. A. (2005). Modeling parallelization and flexibility improvements in skill acquisition: from dual tasks to complex dynamic skills. *Cognitive Science*, 29(3), 421-455.

Taatgen, N. A. (2007). The minimal control principle. In W. Gray (Ed.), *Integrated models of cognitive systems*. Oxford: Oxford University Press.

Taatgen, N. A., Rijn, H. v., & Anderson, J. R. (in press). An Integrated Theory of Prospective Time Interval Estimation: The Role of Cognition, Attention and Learning. *Psychological Review*.