

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

Cognitive Control as Alternation of Activation and Suppression in a Stroop Task

#### **Permalink**

<https://escholarship.org/uc/item/6950q24k>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

#### **ISSN**

1069-7977

#### **Authors**

Juvina, Ion  
Taatgen, Niels A.  
Dickison, Daniel

#### **Publication Date**

2007

Peer reviewed

# Cognitive Control as Alternation of Activation and Suppression in the Stroop Task

Ion Juvina (juvina@andrew.cmu.edu)  
Niels A. Taatgen (taatgen@cmu.edu)  
Daniel Dickison (danieldickison@cmu.edu)

Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue  
Pittsburgh, PA 15213 USA

## Abstract

Execution of tasks in which information rapidly changes its relevance employs both excitatory and inhibitory top-down control. This assertion is derived from empirical data and computational cognitive modeling of a screen-based Stroop task. Empirical data showing both within-trial (interference and facilitation) and between-trial (negative priming) effects are presented. A computational cognitive model developed within the ACT-R architecture is proposed. This model uses a spreading activation mechanism that activates information pertaining to the current trial and suppresses information pertaining to the preceding trial. A good fit of this model to human data is reported and the neuro-cognitive plausibility of top-down suppression is discussed.

**Keywords:** Cognitive control; Cognitive modeling; Stroop.

## Introduction

Cognitive control has been conceptualized in several ways. In the ACT-R theory cognitive control is localized in the anterior cingulate cortex (ACC) (Anderson, 2005) and is viewed as sequencing of steps needed to accomplish a task, resistance to distraction via abstract control states, and allocation of cognitive resources in a parsimonious way (Taatgen, 2007). In the parallel distributed processing (PDP) framework cognitive control is conceived of as top-down excitatory biasing (TEB) from the prefrontal cortex (PFC) to other brain structures involved in performing a task (Miller & Cohen, 2001). The PFC holds task representations that provide biasing input to the processing areas in the posterior cortex, thus selectively activating task relevant information. Task irrelevant information is not exposed to TEB thus losing the competition in favor of task relevant information (Herd, Banich, & O'Reilly, 2006).

Two distinct dimensions of cognitive control seem to be employed in tasks where information rapidly changes its relevance. A symbolic dimension of cognitive control ensures the proper sequencing of steps necessary to perform the task. A sub-symbolic dimension of control ensures that relevant information is activated and irrelevant information is suppressed. This paper seeks to demonstrate that the sub-symbolic dimension of control is achieved not only through a TEB mechanism but also with the aid of a top-down inhibitory biasing (TIB) mechanism.

The two dimensions of control seem to have complementary functional roles. The symbolic dimension enforces and maintains a course of action in face of

distraction and/or interruption; it results in consistent behavior. The sub-symbolic dimension intervenes by activating or suppressing various contents or processes depending on their relevance to the task at hand; this dimension ensures that the system is able to react timely and flexibly to changes in environment.

## Background

The Stroop task is one of the most investigated tasks in Cognitive Psychology (MacLeod, 1991). Since it is so well documented we used it as a reference task in conjunction with other tasks in an individual differences study aiming to understand the role of cognitive control in multitasking.

## The Stroop Task

The standard Stroop task, which will be discussed here (Stroop, 1935), consists of a serial presentation of various words written with various ink colors. Words can be color names such as “red” or neutral such as “bag”. Participants have to name the ink color of the presented words. One of the most known findings is that it takes longer to name the color of incongruent color words (e.g. the word “red” in *green* ink) than to name the color of neutral words. This effect is known as *Stroop interference* and it is usually computed as the difference between the median reaction time on incongruent and neutral trials. It also takes shorter to name the color of congruent color words (e.g. the word “red” in *red* ink) than to name the color of neutral words. This effect is known as *Stroop facilitation* and it is usually computed as the difference between the median reaction time on congruent and neutral trials. Stroop interference and facilitation are referred to here as *within-trial effects* since they are computed in a way that ignores any dependencies between trials.

Performance on a particular Stroop trial seems to be related not only to features of the current stimulus but also to features of the preceding stimulus. For example, when the to-be-ignored word on the previous trial turns out to be the to-be-named color on the current trial (e.g., the word “green” in *red* ink precedes the word “blue” in *green* ink) reaction time increases. The phenomenon, called “negative priming”, is very robust and fairly general (MacLeod, 1991). Theoretical accounts of negative priming (NP) refer to either processes of selective inhibition or episodic memory retrieval (Egner & Hirsch, 2005).

The selective inhibition account (Houghton, Tipper, Weaver, & Shore, 1996) posits an initial activation of both features (word and color) followed by an active inhibition of the to-be-ignored feature (word) of the current stimulus. If the inhibited feature returns as the to-be-named feature of the next stimulus its inhibition has to be overridden by re-activation. This account predicts longer reaction times when the previous word feature re-occurs as the current color feature. Since only the to-be-ignored feature is inhibited (i.e., inhibition is selective), this account predicts that reaction time will not increase when the previous color re-occurs as the current color. In fact, reaction times should decrease, since the to-be-named feature (color) has just been activated in the previous trial.

The episodic retrieval account (Neill, 1997) holds that the to-be-named feature of the current stimulus triggers a retrieval of the most recent episode in which the concept corresponding to that feature has been used and the associated reaction. For example, assuming the word “red” re-occurred as the color *red*, it would trigger the retrieval of an episode composed of the concept “red” and the reaction “no-response”. Since the reaction derived from the retrieved episode is not adequate for the current stimulus, an additional retrieval or a strategy change is required to generate the proper reaction, which explains the time delay. Similar to the selective inhibition account, this account predicts longer reaction times when the previous word feature re-occurs as the current color feature. In the case of the previous color feature re-occurring as the current color feature, this account would also predict no increase in reaction times; the most recent episode involving the current color contains exactly the reaction needed for the current stimulus.

Both the selective inhibition and episodic retrieval accounts explain between-trial effects based on the processes (activation and suppression, or retrieval of previous solutions) responsible for the within-trial effects; in other words, between-trial effects are *by-products* of within-trial effects.

### Computational Cognitive Models of Stroop Effects

In an ACT-R model of the Stroop task developed by Lovett (2005), a representation of the current stimulus is part of the goal representation and spreads activation toward color-association representations in declarative memory. Representations are positively linked when they represent matching colors and negatively linked when they represent mismatching colors. When the word and color features of the stimulus are congruent, the association representation receives positive activation from both features causing the facilitation effect; when the word and color are incongruent the to-be-retrieved representation receives positive activation from the color feature and negative activation from the word feature causing the interference effect. This model seems to support the selective inhibition account: activation and suppression (i.e., activation spreading through negatively weighted links) mechanisms are

employed to model the Stroop interference and facilitation effects. Although it does not address between-trial effects, it could account for negative priming cases where the previous word re-occurs as the current color.

Herd, Banich, and O’Reilly (2006) argue that a TIB mechanism is not necessary to account for interference and facilitation. Using the Leabra modeling environment and only a TEB approach, they were able to not only account for interference and facilitation but also explain fMRI data showing increased activity in brain regions processing to-be-ignored information. A general concept of color, which subsumes linguistic and perceptual features, is represented in the PFC and spreads activation toward processing areas for both word and color. This is how the to-be-ignored word feature gets activated as shown by fMRI data (Herd, Banich, & O’Reilly, 2006). However, the color feature gets extra-activation from another PFC representation responsible for maintaining the task set, that is, the requirement to respond to the color feature. This model demonstrates that there is no need for active top-down inhibition of the word feature because the color feature gets more top-down activation and wins the competition. Moreover, the authors argue that top-down inhibition would be biologically implausible, as inhibitory projections in the human brain are strictly local. It is not clear though how this model would account for between-trial effects using only a TEB mechanism.

## Empirical Study

### Participants, Method, and Procedure

A screen-based Stroop task was administered to 53 participants. Age ranged from 18 to 59 with an average of 24. There were 16 women and 37 men. Participants received a fixed amount of monetary compensation for their participation.

The standard Stroop task was adapted for screen-based administration. Stimuli were color words (red, blue, yellow and green) and neutral words colored with one of the four colors denoted by the mentioned color words. They were presented one at a time and remained on the screen until the participant responded. Two response options were also displayed flanking the stimulus on its left and right sides. Response options were non-colored (i.e., colored with black ink) color words. One response option would contain the right answer and the other one a wrong answer. Every subject received 150 trials, 50 trials for each condition – conflicting, congruent and neutral. The location of stimuli on the screen was kept constant. All the other factors were randomized.

Participants were instructed to select the response option that matched the color of the stimulus and started with a short computer-guided tutorial that emphasized the correct response. During the actual task no feedback was provided.

### Results of the Empirical Study

The data of one participant were excluded from analyses due to failure to follow task instructions. Accuracy and

reaction time (RT) were recorded for each trial. Accuracy data are consistent with previous studies, showing less than 1% errors for the congruent and neutral conditions and less than 10% errors for the incongruent condition. For the following analyses only RT from correct trials was considered. Significant interference and facilitation effects were found. Since within-trial effects were very consistent with those found in previous studies they will not be treated in more detail here. However, the actual mean values can be seen in the next section where model simulations will be compared to the empirical data.

Two significant between-trial effects were found (see Table 1). They will be referred here to as Word-Color (W-C) and Color-Color (C-C) effects, respectively. The W-C effect is observed when the word feature of the preceding stimulus re-occurs as the color feature of the current stimulus. The average RT for the W-C prime trials is significantly higher than for no-prime trials ( $t(51)=6.4$ ,  $p<0.001$ ).

The C-C effect is observed when the color feature of the preceding stimulus re-occurs as the color feature of the current stimulus. The average RT for the C-C prime trials is significantly higher than for no-prime trials ( $t(51)=2.86$ ,  $p<0.01$ ).

Table 1: Mean RT (ms) and t-tests showing between-trial effects (\* significant at alpha = 0.01 two tailed).

Type of effect	No prime	Prime trials	t(51)	p
W-C	1036	1111	6.4	0.000*
C-W	1051	1030	-1.3	0.193
C-C	1040	1075	2.9	0.006*
W-W	1045	1055	0.7	0.481

An additional between-trial effect, which has been reported elsewhere (MacLeod, 1991), was found non-significant in this study. This effect would consist of reduced RTs for trials where the to-be-ignored word of the current stimulus names the color of the preceding stimulus. Following the same coding scheme as above this effect could be referred to as the Color-Word (C-W) effect. A reduction in RTs of about 20ms can indeed be observed in our data but it doesn't reach the threshold of statistical significance.

Table 2: Frequency of various trial types that have a potential to produce between-trial effects.

Type of effect	No prime trials	Prime trials	%
W-C	6573	1377	17.3
C-W	6660	1290	16.2
C-C	5970	1980	24.9
W-W	7048	902	11.3

Table 2 shows the frequency of occurrence of various trial types. These data are reported because the magnitude and direction of between-trial effects may be related to the low frequency of prime trials relative to the frequency of no-prime trials.

## Discussion of the Empirical Findings

The data presented here mostly confirm well-known effects in the Stroop task: interference, facilitation, and negative priming. However, the C-C effect is less known and to our knowledge has not been reported so far in the context of the Stroop task.

While it seems plausible that within-trial effects (inhibition and facilitation) can be explained only based on a TEB mechanism, it is hard to imagine how the between-trial effects can be explained without a TIB mechanism. Could this TIB mechanism be the one suggested by the selective inhibition account? This account would easily explain the W-C effect. Because the to-be-ignored word feature has been actively suppressed on the preceding trial, it takes approximately 75ms longer to reactivate the name of the color feature of the current stimulus. However, this account would fail to explain the C-W and the C-C effects. Since the color feature has been activated (and not suppressed) on the preceding trial, the word feature of the current stimulus would have an even stronger potential to interfere, that is to increase RTs (and not decrease them as in the data). When the color feature of the preceding stimulus re-occurs as the color feature of the current stimulus there would be no reason for the RTs to increase by about 35ms, as is observed in the data, because the color feature has not been inhibited according to the selective inhibition account.

Would perhaps the episodic retrieval account explain the negative priming effects better? It explains well the W-C effect but it fails to explain the other two. In the case of the C-W effect, this account would predict increase in RT (in opposition to the data) since the last episode involving the color feature of the preceding stimulus contains the reaction "response" which is not necessarily adequate for the current stimulus. In the case of the C-C effect, the last episode contains exactly the reaction that is needed for the current stimulus, thus there would be no reason for the RT to increase (as in the data).

A possible alternative account suggested by our data could be referred to as *semantic inhibition-of-return* (SIOR) by analogy with *inhibition-of-return* (IOR) – a well-known effect in the field of visual search (Klein, 2000). IOR is the delay in attending a visual stimulus at a location that was previously attended. It is thought to have the function of increasing efficiency of search and foraging behaviors by preventing attentional resources to be wasted at locations that were previously attended. By analogy, SIOR would be the suppression of the concept that has just been activated and used in the preceding trial, regardless of location. This account would predict increased reaction times not only when the preceding word but also when the preceding color is identical to the to-be-named color of the current stimulus.

Unlike the selective inhibition account, the SIOR account posits inhibition of both features (word and color) of the preceding stimulus. The functional role of SIOR would be to prevent the representation of the preceding stimulus from interfering with processing of the current stimulus. Thus, the functional role of SIOR seems identical to the role of memory decay (Altmann, 2002; Anderson, 1989). Questions arise whether an additional inhibitory mechanism is necessary and how plausible such a mechanism is from a neuro-cognitive perspective. These questions are addressed in the next sections.

### An ACT-R Model of Stroop Effects

This model is largely based on the models of Altmann and Lovett (Altmann & Davidson, 2001; Lovett, 2005) both implemented in the ACT-R architecture. An extension was necessary to account for the observed between-trial effects while preserving a good account for the within-trial effects.

The word and the color features of the current stimulus are encoded as parts of the goal representation and spread activation to related information in memory. For example, when the stimulus is the word “green” in red ink, both features spread activation and two concepts are likely to be retrieved – *redness* and *greenness*. If greenness is retrieved, the system either commits an error or performs a second retrieval to get the correct concept. If the two features of the stimulus are congruent, they spread activation toward the same concept, thus the probability of the right concept to be retrieved and the retrieval speed are very high.

In order to account for between-trial effects an inhibitory goal representation has been added. This inhibitory goal behaves exactly as the original (activating) goal of ACT-R, except it spreads negative activation in order to suppress the features (word and color) of the previous stimulus. Thus, after the current stimulus has been processed it is included in the inhibitory goal representation. Normally, this suppression of the previous stimulus reduces its chance to interfere with the current stimulus. However, in some cases, when a feature of the previous stimulus re-occurs with the current stimulus, retrieval of the corresponding concept from memory is significantly slowed down, because it has been suppressed (i.e., negative activation has been spread toward it). This way, as a result of activation and suppression the model is able to account for both within- and between-trial effects.

### Results of Model Simulations

Simulations of the model have been designed to allow comparison with the human data. The same number of individuals and trials per individual were administered as in the actual study.

In general the model was able to account for a variety of Stroop effects and task manipulations as shown in previous modeling work (Lovett, 2005). Here a preferential treatment is given to those results suggesting answers to the questions posed above, that is, whether or not a TIB mechanism is necessary and plausible. Two versions of the model have

been run separately and their results compared against the human data. These two versions differ only with regard to the presence or absence of the TIB mechanism.

Figure 1 shows RT for each condition for the two versions of the model and for the data.

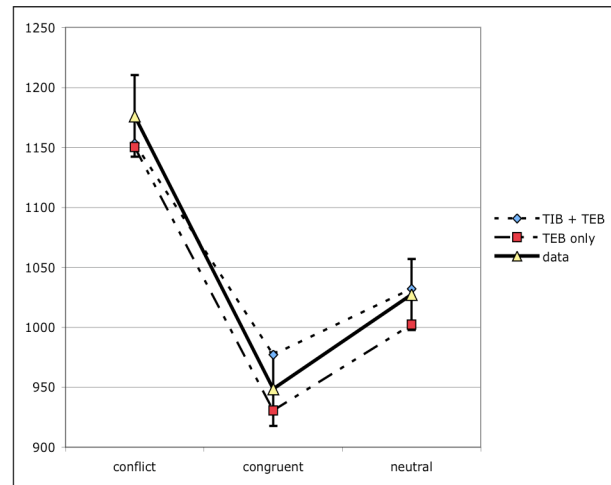


Figure 1: Reaction time per condition showing interference and facilitation effects (vertical bars represent standard errors of the means in the data).

The two versions of the model, TEB and TIB+TEB, seem to account equally well for the within-trial effects. RTs of the TIB+TEB model are slightly higher overall because of a generalized reduction in activation caused by TIB.

Figure 2 shows RT for the W-C and C-C trials as compared with an average RT for no-prime trials, again evaluating the two versions of the model against the human data.

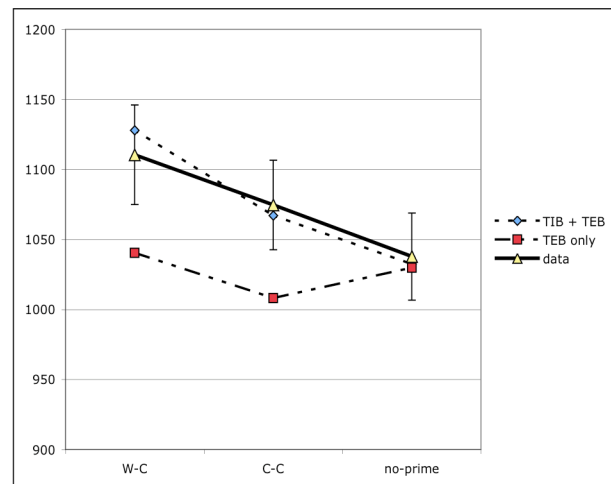


Figure 2: Reaction time for W-C, C-C, and no-prime trials, showing that only the TIB+TEB model is able to account for both negative priming effects (vertical bars represent standard errors of the means in the data).

While the TIB+TEB version accounts reasonably well for the data, the TEB version of the model is unable to account for the between-trial effects observed in the data. For the no-prime trials the two versions behave exactly the same the difference between them becoming visible only for the prime trials.

### Discussion of the Simulation Results

The simulation results show that a TIB control mechanism is indeed necessary and a TEB mechanism alone is insufficient to account for the behavioral effects observed in our study. It can be argued that the TEB model could have been enriched with a within-trial selective inhibition or an episodic retrieval mechanism and made able to account for the W-C effect. However, such a model would still be unable to account for the C-C effect.

It can also be argued that the TIB mechanism that we propose could produce other effects that are not observed in the data. For example, a W-W effect would be expected. When the word feature of the preceding stimulus re-occurs as the word feature of the current stimulus RT should be decreased. Since the re-occurring word feature has been suppressed it should have less interfering potential. This effect has not been observed in our data. A C-W could also be predicted by the proposed TIB mechanism. Since the color feature of the preceding stimulus has been suppressed, it has less interfering potential when it returns as the word feature of the current stimulus, thus causing reduced RT. This effect has been observed in our data but its magnitude is small and it doesn't reach statistical significance. Although we don't have a complete solution to this issue yet, a mechanism is already in place that is responsible for the differences in magnitude observed between the different negative priming effects. This mechanism is based on the well-known asymmetry between word reading and color naming (Cattell, 1886). Word reading is a more dominant and automatic process than color naming. In some of the trials the color feature is not encoded. For example, in the congruent condition a fast response can be generated based only on the word feature. In these cases the color feature is also not inhibited at the end of the trial, thus losing a chance to increase the magnitude of the C-W or C-C effects. As for a possible W-W effect, three scenarios are conceivable:

- It may actually exist but we failed to measure it. The set size of the word feature (50 words) was much larger than the set size of the color feature (4 colors). Thus the chance that a word is repeated is relatively small (see Table 2).
- It may be that the word reading routine is so strong that a temporary top-down suppression has little to no effect on it.
- Individual differences may exist with regard to the ratio of inhibitory vs. excitatory control. Individuals using TIB to avoid between-trials interference would show decreased RT, whereas individuals using TEB-only would show increased RT in W-W trials. Given

these circumstances, averaging across all individuals might cause the W-W effect to disappear altogether.

The simulation results presented above show that within- and between-trial effects are independent of each other. Although a small increase in RT was observed due to the TIB mechanism, it is quite evenly distributed across conditions thus not affecting the magnitudes of interference and facilitation. The classical accounts of between-trial effects assume that they are by-products of within-trial effects. This assumption makes them fail to explain the whole range of negative priming effects. When a separate process is assumed to be responsible for between-trial effects as in the SIOR account, the whole range of results is well accounted for.

### General Discussion and Conclusion

The empirical and modeling results presented above suggest that both excitatory and inhibitory processes are involved in the sub-symbolic dimension of cognitive control. In accord with previous models of the Stroop task (Herd, Banich, & O'Reilly, 2006), a TIB mechanism was indeed not necessary to account for within-trial effects. However, such a mechanism was necessary to account for between-trial effects.

TEB and TIB have been implemented by representing task relevant information in two distinct goal representations. The excitatory goal holds a representation of the current stimulus and the inhibitory goal holds a representation of the previous stimulus. Arguably, such alternation of activation and suppression has a functional role in tasks involving rapid serial presentations of stimuli: the current stimulus should be activated and the previous stimulus should be prevented to interfere with the current stimulus. When it happens that features of the previous stimulus (word, color or both) return as features of the current stimulus, they need to be reactivated. Thus, in a time interval of about 1s, a mental representation might be exposed to a complete activation-suppression-reativation sequence. Such a fast adaptation mechanism is one of the characteristics of cognitive control (Gilbert & Shallice, 2002). Mental representations and processes involved in performing a task suffer only temporary top-down influences that bias their regular state or activity.

A TIB mechanism has been rejected by Herd et al. (2006) on basis of biological implausibility. They cited evidence that inhibitory projections in the human brain are strictly local (White, 1989). Our model suggests that the inhibitory signal does not need to travel across the brain: its target might be located in the same brain area. The effects we have described happen at retrieval. PFC has been thought to be involved in both retrieval (Anderson, Albert, & Fincham, 2005; Rugg, Henson, & Robb, 2003) and in holding mental representations involved in cognitive control (Miller & Cohen, 2001).

There are perhaps alternative ways to explain and model the effects observed in our study. Fuentes (1999) proposed a temporary inhibitory tagging mechanism to account for a

variety of IOR effects. They argued that such a tagging mechanism extended to semantic and response relevant properties of stimuli (Fuentes, 1999), thus it could explain our results. ACT-R has an inhibitory tagging mechanism called *finst* that could be used to model inhibitory processes. However, IOR seems a universal and stable phenomenon that affects perhaps any search or foraging process. Correspondingly, the *finst* mechanism consists of a set of parameters that are usually set by the modeler and affect the behavior of the model in a rather deterministic manner. The inhibitory mechanism we propose here has the potential to be adaptive and flexible, thus appropriate to model cognitive control. Arguably, the effects we have described are influenced by properties of the task environment as well as individual strategies and learning effects. For example, the magnitude of W-C and C-C effects would probably diminish with practice because they work against the adaptive role of SIOR, which is to reduce interference. If the frequency of the prime trials increases people might change their strategies and actually reverse the negative priming effects. These predictions are to be addressed in further empirical and modeling work.

The proposed mechanism to model the sub-symbolic dimension is complementary to the symbolic dimension of cognitive control. Our model employed a symbolic dimension of control to ensure a proper sequencing of actions and resist the influence of pre-potent responses. This way, for example, the model avoided making too many errors. This symbolic dimension of control is believed to be localized in ACC (Anderson, 2005), whereas the sub-symbolic dimension seems to be localized in PFC (Herd, Banich, & O'Reilly, 2006; Miller & Cohen, 2001). The sub-symbolic dimension of control seem to be useful in modeling dynamic and temporary aspects of task performance such as rapid serial presentation and repetitions of stimuli and task switching. They are likely to be influenced by practice, reward, cognitive strategies and individual differences.

This paper has argued for a composite nature of the sub-symbolic dimension of cognitive control. An alternation of activation and suppression modulates performance in tasks in which information rapidly changes its relevance.

### Acknowledgments

The research reported here was supported from ONR grant no. N00014-06-1-0055. We would like to thank Andrea Stocco, and Jelmer Borst for their comments on a previous draft of this paper.

### References

Altmann, E. M. (2002). Functional decay of memory for tasks. *Psychological Research*, 66 287-297.

Altmann, E. M., & Davidson, D. J. (2001). *An integrative approach to Stroop: Combining a language model and a unified cognitive theory* Paper presented at the Twenty-Third Annual Conference of the Cognitive Science Society, Hillsdale, NJ

Anderson, J. R. (1989). Human Memory: An Adaptive Perspective. *Psychological Review*, 96, 703-719.

Anderson, J. R. (2005). Human symbol manipulation within an integrated cognitive architecture. *Cognitive Science*, 29(3), 313-341.

Anderson, J. R., Albert, M. V., & Fincham, J. M. (2005). Tracing Problem Solving in Real Time: fMRI Analysis of the Subject-Paced Tower of Hanoi *Journal of Cognitive Neuroscience*, 17 1261-1274.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, 11, 63-65.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience*, 8(1784-1790).

Fuentes, L. J. (1999). Inhibitory Tagging of Stimulus Properties in Inhibition of Return: Effects on Semantic Priming and Flanker Interference *The Quarterly Journal of Experimental Psychology: Section A*, 52(1), 149 - 164.

Gilbert, S. J., & Shallice, T. (2002). Task Switching: A PDP Model. *Cognitive Psychology*, 44, 297-337.

Herd, S. A., Banich, M. T., & O'Reilly, R. C. (2006). Neural Mechanisms of Cognitive Control: An Integrative Model of Stroop Task Performance and fMRI Data *J. Cognitive Neuroscience MIT Press*, 18(1), 22-32

Houghton, G., Tipper, S. P., Weaver, B., & Shore, D. I. (1996). Inhibition and interference in selective attention: Some tests of a neural network model. *Visual Cognition*, 3(119-164).

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences*, 4(4), 138-147.

Lovett, M. C. (2005). A Strategy-Based Interpretation of Stroop. *Cognitive Science*(29), 493-524.

MacLeod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review *Psychological Bulletin*, 109(2), 163-203.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.*, 24, 167-202.

Neill, W. T. (1997). Episodic Retrieval in Negative Priming and Repetition Priming. *Journal of Experimental Psychology; Learning, Memory, and Cognition* 23(6), 1291-1305.

Rugg, M. D., Henson, R. N., & Robb, W. G. (2003). Neural correlates of retrieval processing in the prefrontal cortex during recognition and exclusion tasks. *Neuropsychologia*, 41(1), 40-52.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.

Taatgen, N. A. (2007). The Minimal Control Principle. In W. Gray (ed.), *Integrated models of cognitive systems*. Oxford: Oxford University Press.

White, E. L. (1989). *Cortical circuits: Synaptic organization of the cerebral cortex, structure, function, and theory*. Boston: Birkhauser