AN ALTERNATE CHANNEL THEORY OF THE STROOP EFFECT

Rani Moran and Daniel Algom
Department of Psychology, Tel-Aviv University
rani.moran@gmail.com , algomd@freud.tau.ac.il

Abstract

Virtually all existing accounts of the Stroop Effect (SE) rely on a coactive architecture, wherein information from both the color and word channels is integrated. We challenge this assumption by presenting a new formal alternate-channel theory of the SE. The theory assumes that on each trial exactly one, i.e. color or word but not both, channel determines response. Performance (accuracy, RT) emerges from a mixture of color and word operative trials. The theory is shown to account for Stroop facilitation, interference and the total effect, with respect to both accuracy and mean RT. Nevertheless the model is limited by the strong quantitative bond between accuracy and latency: Given a low error rate, 'small' temporal effects are predicted.

When asked to name the ink color in which color words are printed, people are unable to ignore the meaning of task-irrelevant words. It takes people longer to name the color of a conflicting color word (e.g., the word RED printed in ink color green) than that of a matching color word (GREEN in green), the Stroop effect (SE; Stroop, 1935). Whence the Stroop effect? Virtually all existing models of the Stroop effect espouse an interactive architecture to govern the processing of the ink color and the color word. In other words, color and word cross-talk to determine the response. Activation from the color and the word channels is summed up, or more generally integrated, in the process of response selection.

The present theory is radical in challenging precisely the tacit assumption of a color-word interaction in the psychological theater of operations. Consider what *is* and what *is not* mandated by the presence of a Stroop effect. The presence of the Stroop effect mandates the conclusion that the word was processed so that it was not completely ignored. However it does not follow that the word was processed on <u>all</u> of the experimental trials. It is eminently possible that the word was processed only on a subset of the trials. Now it is uncertain that word and color were processed simultaneously on those trials in which the word was processed. If they were not, an interactive architecture is precluded. We argue that, the observed behavioral effects (facilitation or interference) do not logically mandate a corresponding facilitation or interference in the underlying information processing. The hallmark of the present theory is the lack of a cross-talk between word and color in processing the Stroop stimulus. The color horse does not know the position, speed, indeed the very existence of the word horse. It is in the face of such a completely independent architecture that we derive the Stroop effect.

The present theory suggests that, on a given trial, a single information channel is processed: either that of the color or that of the word, but not both. On the majority of the experimental trials, the participant complies with the instructions and ignores the irrelevant word. However, on some small portion of the trials, the participant (erroneously) responds to the word and ignores the color. Clearly, it is impossible for color and word to interact simply because both are not processed on the same trials. It is shown that the Stroop effect obtains as a natural consequence of this simple part-time-processing scenario.

An Alternate Channel Theory of the Stroop Effect

The alternate channel theory is a formal theory of the effect. The approach is axiomatic unlike those based on modeling, computation and simulation. A small set of axioms or basic assumptions served us to *prove* that the Stroop effect *must* ensue given the validity of the axioms. Space limits shall not allow a fully formal description so we confine ourselves to an informal description of the theory.

We consider the standard Stroop task of naming the color of a color word in a 2 (WORD: RED, GREEN) x 2 (color: red, green) stimulus matrix. We use capital letters to denote words. The following assumptions form the basis of our theory (The reader is referred to Figure 1.)

Assumption 1. On each trial a single stimulus dimension is processed (we will also say that a single channel is operative): either word or color but not both. Full compliance with the task instruction mandates color processing on all trials. Nevertheless, we assume that on some (small) portion of the trials the word is processed and hence the color is ineffective. We refer to trials on which the color *or* the word is operative as color-operative or word-operative trials, respectively. We denote by α the constant probability that on a trial the color channel is operative. Our part-time word processing assumption translates into $0< \alpha <1$. We also assume that the operative channels are independent across trials.

Assumption 2. For the operative channel (word or color), all possible stimulus values for that channel are activated, regardless of the particular value presented for view.). Consider the word RED printed in ink color green (hence RED in green). If the color channel is operative, red and green are activated both, green from perception and red from memory. Otherwise, the word channel is operative, in which case RED and GREEN are activated both, RED from perception and GREEN from memory. This is a stochastic modeling instantiation of a well-known tenet of information theory. Perception depends not only on the stimulus presented for view but also on its alternatives, those stimuli that could have been presented although were not presented on that particular trial (Garner, 1974). A stimulus has information properties in addition to its energy properties.

Assumption 3. The architecture of each channel is that of a strictly parallel independent race. On color-operative trials, red races against green and on word-operative trials RED races against GREEN.

Assumption 4. In each channel, processing is stochastic accomplished with imperfect accuracy. Each racer is modeled as a Poisson counter ("racer") and hence accumulates counts at an exponential rate v (i.e., the times between counts on a counter have an exponential distribution). The first counter to reach a threshold K wins the race and determines the response. Thus, the response is "red" whenever red or RED wins the race and the response is "green" whenever green or GREEN wins the race. The correct response is produced when the counter entailing the presented print color or the counter entailing its matching color word (whether or not that word is presented on that particular trial) wins the race and determines the response. An error occurs when any of the other counters wins the race. We categorize the trials according to the operative channel (color, word) and the winner's origin (perception of the presented color/word, memory of the non-presented color/word). For example, a "word-memory trial" refers to an operative-word trial in which memory is the winner.

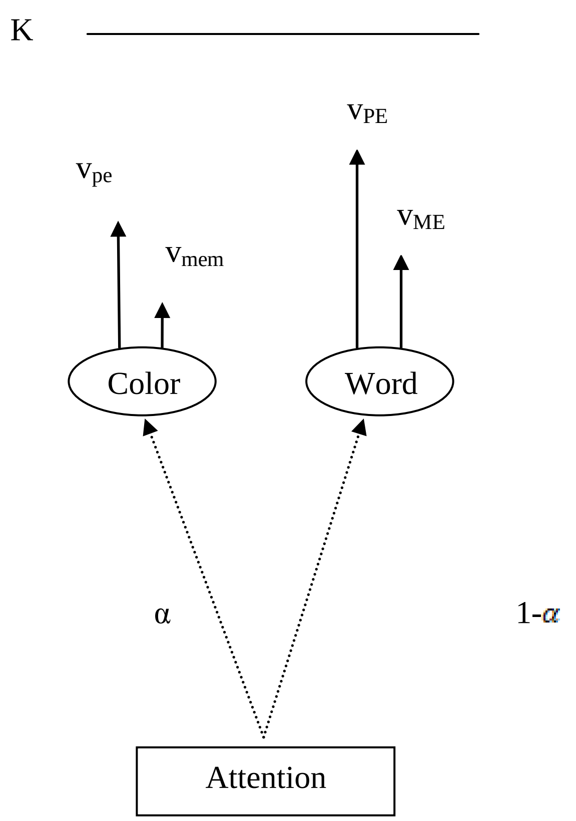

Figure 1: The alternate channel model architecture. On each trial one and only one channel is operative, color, with probability α or word, with probability 1- α . The architecture of each channel is that of a strictly parallel and independent Poisson race between information activated from perception and memory. The first racer to reach threshold K, represented by the top horizontal line, determines the threshold. The rates of the racers are represented by the lengths of the upward pointing arrows, with longer arrows for higher rates. The perception rates are higher than the memory rates corresponding to Assumption 5. The word rates (both for perception and memory) are higher than the color rates, corresponding to Assumption 6.

Note that correct and erroneous responses can ensue from either word or color processing. For congruent stimuli, perceived word and perceived color are conductive to the correct response and remembered word and remembered color are conductive *to* the erroneous one. For incongruent stimuli, perceived color and remembered word are conductive to the correct response, whereas remembered color and perceived word are conductive to an error. When the word channel is operative, the observer is actually responding to information on the task-irrelevant channel. However, her or his response may still be correct, in which case an "unobservable error" has occurred. Our theory depends crucially on these "unobservable errors;" they are necessary for Stroop effects to occur.

Assumption 5. Perceived attributes (i.e., the attributes exposed for view on a particular trial) are processed more efficiently than memory-based attributes (the attributes remembered from previous trials).

Assumption 6. The word channel is processed more efficiently than the color channel. This assumption may reflect the fact that observers are trained in reading words more extensively than they are in naming colors or that, in most Stroop experiments, words are more salient than colors (Melara & Algom, 2003).

Generating the Stroop effect

Assumptions 1-6 converge on a natural theory that yields the quintessential Stroop phenomenon in a straightforward fashion. Let us consider the trials in which the ink color is red. On each trial, either the color or the word channel is operative (Assumption 1), with the response determined solely by the operative channel (Assumption 4). Because congruent and incongruent stimuli carry identical color information (red), color-produced responses are invariant across the two classes of stimuli (hence cannot generate a Stroop effect). Inevitably, the Stroop effect is generated on word operative trials. Consider the values of word on those trials. For congruent stimuli, RED (the matching color word) is activated by perception and GREEN is activated by memory. For incongruent stimuli, RED is activated from memory and GREEN is activated by perception. By Assumptions 2-4, each word-operative trial is a strictly parallel and independent race between two Poisson counters, RED and GREEN. This race occurs with congruent and incongruent stimuli alike. However, on congruent trials the word RED, counting for the correct response, is *perceived*, whereas it is merely *remembered* on incongruent trials. The superior processing of perceived over memory-based attributes (Assumption 5) readily generates commensurately better performance on congruent than on incongruent trials (= the Stroop effect). The reader is referred to Figure 2.

Formally, espousing Assumption 5 is conductive to two important results. First, perceptual information is endowed with a higher probability of winning the race compared with that based on memory information. In other words, the perception horse beats the memory horse most of the time. Second, the winning processing times are shorter for perception than for memory. When perception wins, the processing ends sooner than when memory wins. Assumption 6 asserts that the word channel is more efficient than the color channel. Espousing Assumption 6, it follows that winning times on the word channel are faster than those on the color channel.

We are now in a position to derive the Stroop effect for both accuracy and reaction time. Consider first mean accuracy rates. For congruent stimuli, the correct response obtains when the perceptual channel for color (red) wins (on color-operative trials) or when the perceptual channel for word (RED) wins (on word operative trials). For incongruent stimuli, by contrast, the correct response obtains when the perceptual channel for color (red) wins (on color-operative trials) or when the memory channel for word (RED) wins (on word-operative trials). For the latter trials, the word counting for the correct response is perceived with congruent stimuli, but is remembered with incongruent stimuli. The probability of winning the word-channel race is larger for perceptual than for memory information and hence, accuracy is better for congruent than for incongruent stimuli.

Next, consider the mean reaction times for correct responses. For congruent stimuli it is the weighted average of the RTs obtained on trials in which the perceptual channel for color wins (on color-operative trials) and when the perceptual channel for word wins (on word-operative trials). For incongruent stimuli, the mean RT is the weighted average of the RTs obtained on the perceptual channel for color (on color-operative trials) and on the memory channel for word (on word- operative trials). Again, congruent and incongruent RTs differ only on word-operative trials. Notice that the mean RT for congruent stimuli benefits from the contribution of the word on word-operative trials because, the perceptual channel for word is faster (i.e. has shorter winning times) than the perceptual channel for color. For incongruent stimuli, the word (on word-operative trials) is not generally supportive of fast performance because it is memory based. But even if it is (if word-memory trials are generally faster than perception-color trials) it necessarily yields a smaller gain than the word-perception gain for congruent stimuli. The reason for this is two-fold: In the word channel, perception is both faster and wins more often than memory. Thus, congruent trials <u>more often</u> enjoy a <u>larger</u>

word-gain than incongruent trials. It is the speed advantage generated for congruent over incongruent stimuli that appears as the behavioral Stroop effect.

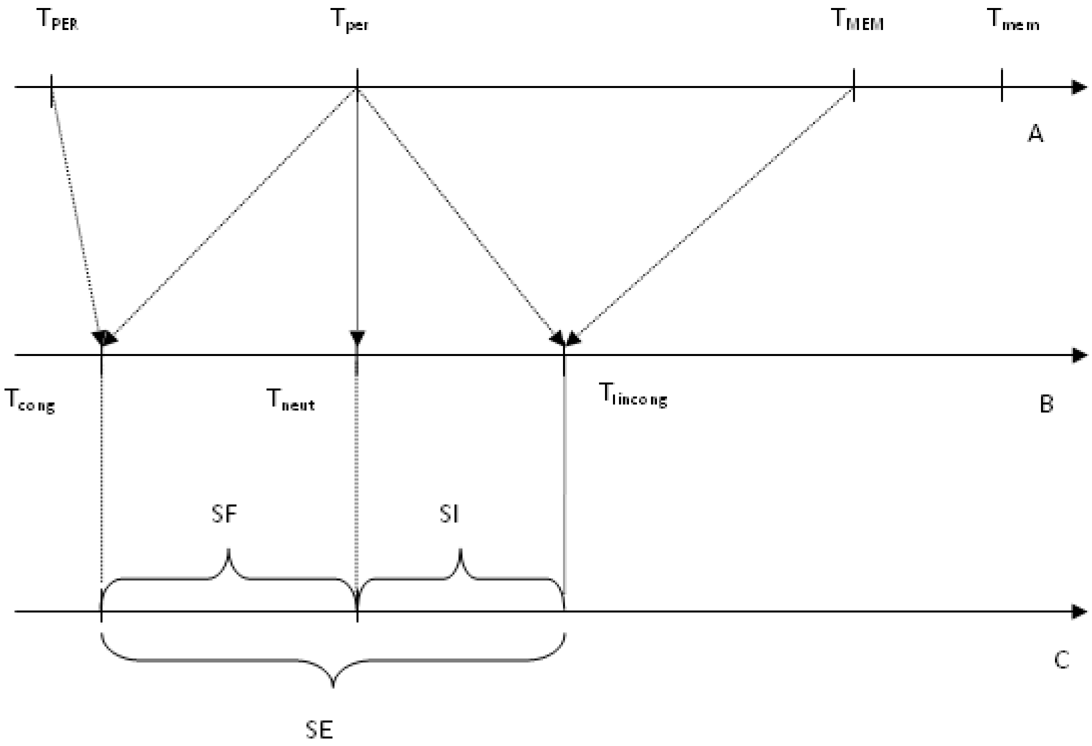

Figure 2: The RT Stroop Phenomenon: The top horizontal axis (A) displays the typical ordering of mean winning times (denoted by T) for the different racers. Within a channel, the perception racers (per, PER) have shorter wining time than the corresponding memory racer (mem,MEM) and word racers (PER,MEM) have shorter winning times than the corresponding color racers (per,mem). The middle horizontal axis (B) displays the resultants mean RT for different stimuli type (congruent, neutral and incongruent). The dashed arrows demonstrate that 1) mean RT for congruent stimuli is a weighted average of mean RT in the color-perception and the word-perception trials and, 2) mean RT for incongruent stimuli is a weighted average of mean RT in the color-perception and the word-memory trials. 3) mean RT for neutral stimuli equals mean RT of color-perception trials. The bottom axis (C) shows the emergent differences between mean RT for congruent neutral and incongruent stimuli, the Stroop effect (SE).

Stroop Facilitation and Stroop Interference

The Stroop effect can be subdivided into two components. Stroop facilitation (SF) is the difference in performance between congruent and neutral stimuli. Stroop interference (SI) is the complementary difference between neutral and incongruent stimuli. The Stroop effect (SE) is thus the algebraic sum of interference and facilitation. Typically, performance is better for congruent than for neutral items and for neutral items than to incongruent items. Alternate channels theory readily explains these components of the Stroop effect, too. Since potential responses to the carrier are completely irrelevant to the Stroop races that determine the requested color response, we assume that the color channel is always operative on trials entailing neutral stimuli. Within this channel, there is still a competition between the Poisson racers for perception and memory. It can be shown, that under this assumption the theory generalizes to account for both SI and SF, both with respect to accuracy rates and mean response-times (see Figure 2).

Model Limitations

The alternate channel theory *is* successful in *qualitatively* predicting the Stroop phenomena both with respect to accuracy and reaction-time. Nevertheless the model is limited in the sense that is displays a strong bond between the magnitudes of the accuracy and the SE: Error rate constraints the mean-RT effect, posing an upper bound on the SE. Given a low error rate, the model can *quantitatively* predict only a "small" SE. The fundamental issue is the dual role of unobservable errors (UE) in the model: On the one hand, UE's are the cause of mean-RT facilitation and interference. Thus, 'large' temporal effects mandate a high rate of UE's. But on the other hand, a high UE rate implies that the word channel is highly operative, which results also in a substantial (observable) error rate. Therefore, 'large' temporal effects are necessarily accompanied by a substantial error rate. For example, if the influence of the irrelevant word channel is strong, then word-perception trials will greatly improve performance on congruent trials. However the same influence will generate a high error rate in incongruent trials.

An explicit example will hopefully illustrate the model's limitation: Consider a Stroop experiment with 5% error rate in the incongruent condition. Assume that the perception-memory trials account for 1% of the errors (those trials are solely responsible for errors in the neutral condition and also yield errors in both the congruent and the incongruent conditions). It follows that word-operative trials account for the rest, 4% of the errors, in the incongruent condition. Now, assume in addition that on word-operative trials perception beats memory 75% of the time. Since, word-perception trials yield errors for incongruent trials, we obtain $\alpha \sim 0.04/0.75=0.053$. Hence, there are about (25% of 5.33=) 1.33% of word-memory trials. But these trials are the solely responsible for the mean-RT interference. Hence, if word-memory trials are 50% slower than color-perception trials the mean-RT interference will be less than 1%. Of course, with the inclusion of a base (non-decision) time component in empirical model fits, the effect decreases further. Therefore, the model cannot realistically account for the combination of a large temporal interference effect and a 'small' error rate.

The upshot is clear. An SE is possible, indeed mandated, within an independent scheme of color-word processing. However, in such a framework, the magnitude of the SE is too small to account for much human performance.

References

Garner, W.R. (1974). *The processing of information and structure.* Potomac, MD: Erlbaum.

Melara, R.D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110*, 422 - 471.

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