Journal of Experimental Psychology: General
2004, Vol. 133, No. 3, 323–338

0096-3445/04/$12.00 DOI: 10.1037/0096-3445.133.3.323

A Rational Look at the Emotional Stroop Phenomenon: A Generic Slowdown, Not a Stroop Effect

Daniel Algom
Tel Aviv University

Eran Chajut
The Open University

Shlomo Lev
Tel Aviv University

The role of Stroop processes in the emotional Stroop effect was subjected to a conceptual scrutiny augmented by a series of experiments entailing reading or lexical decision as well as color naming. The analysis showed that the Stroop effect is not defined in the emotional Stroop task. The experiments showed that reading, lexical decision, and color naming all are slower with emotional words and that this delay is immune to task-irrelevant variation and to changes in the relative salience of the words and the colors. The delay was absent when emotional and neutral words appeared in a single block. A threat-driven generic slowdown is implicated, not a selective attention mechanism associated with the classic Stroop effect.

The Stroop effect (Stroop, 1935) is a prime example of the human failure to attend selectively to a particular aspect of a complex stimulus. When naming the print color of color words, people seem unable to prevent processes of reading even when such processes are irrelevant to the task at hand and can hurt performance. The task engages controlled processes (color naming) and more strongly automatic processes (reading) in a special setup. It is for these reasons that more than 60 years after its discovery, the Stroop effect continues to fascinate researchers, sustaining an impressive amount of research (for reviews, see C. M. MacLeod, 1991; Melara & Algom, 2003). Recently, another phenomenon, the emotional Stroop effect, has been attracting a great deal of attention; indeed, the emotional Stroop effect rivals its namesake in sheer output of experimental research (for a review, see Williams, Mathews, & C. MacLeod, 1996).[1] In the emotional Stroop task, people also name the print color of words, but the carrier words are not color words. They are emotionally charged words, threat words, or words related to the pathology of clinical patients, as well as neutral words. The emotional Stroop effect is the slowdown in naming the color of the emotional words compared with the naming speed of the neutral words, assuming that concurrent processes of reading are minimal and do not change across the different words. Both the Stroop and the emotional Stroop effects thus document the failure of fully selective focusing on color induced by the carrier words (Burt, 2002). Are the two effects related to one another? Are the two phenomena bearing the name of J. R. Stroop produced by the same mechanism and mediated by the same processes?

To answer these questions, we first examine the definitions of the Stroop and emotional Stroop effects. Although the respective procedures are routine (C. M. MacLeod, 1991; Williams et al., 1996), the underlying definitions have rarely been spelled out in a methodical fashion. Their explicit formulation shows that the emotional Stroop effect is not a Stroop effect. The conceptual overhaul is followed by a set of six experiments probing the emotional Stroop effect. A notable feature of these experiments is that the participants both named the print colors and read the various words (in separate tasks). The pervasive slowdown observed with emotional words was immune to changes in the salience of the carrier words and remained intact even in a condition in which only a single word (varying in color) was presented. However, the slowdown was absent in a condition in which the emotional words appeared along with the neutral words in the same block of trials. The results are commensurate with an interpretation of the emotional Stroop effect in terms of a general-purpose defense mechanism that reacts to threat (such as that conveyed by emotional words) by temporarily disrupting all ongoing activity (reading as well as color naming). In summary, all

Daniel Algom and Shlomo Lev, Department of Psychology, Tel Aviv University, Tel Aviv, Israel; Eran Chajut, Department of Psychology, The Open University, Ramat-Aviv, Israel.

This research is based on portions of a doctoral dissertation by Shlomo Lev in the Department of Psychology, Tel Aviv University, supervised by Daniel Algom. We are grateful to Yonatan Goshen-Gottstein for helpful suggestions. Thanks are also due to Boaz M. Ben-David, Danny Fitousi, and Leora Levy for their assistance. Many insightful comments by Derek Besner and Jenny Burt improved this project.

Correspondence concerning this article should be addressed to Daniel Algom, Department of Psychology, Tel Aviv University, Ramat-Aviv 69978, Israel. E-mail: algomd@freud.tau.ac.il

[1] Two of the major figures investigating the respective phenomena go by the same name, Colin MacLeod. The Colin MacLeod of the classic effect and the Colin MacLeod of the emotional effect are distinguished by a single character, the added middle initial of the former, often lost on the unprepared observer.

the analyses, conceptual and empirical, converged on the conclusion that the Stroop and emotional Stroop effects are independent phenomena.

Conceptual Review: Classic and Emotional Stroop Tasks, Definitions, and Effects

The Classic Stroop Effect: The Logical Structure of Stroop Stimuli

The Stroop effect is the difference in color-naming performance between congruent (e.g., the word *red* printed in red) and incongruent (e.g., the word *red* printed in green) stimuli. The presence of the Stroop effect documents the failure to focus exclusively on the target dimension of color. The effect shows that the participant noticed the task-irrelevant dimension of word, thereby compromising exclusive attention to color. It is immediately apparent that all multidimensional stimuli (e.g., colored noncolor words, colored nonwords, colored geometric forms, sounds of varying intensity and frequency) are not also Stroop stimuli. A red triangle is not a Stroop stimulus. It is neither more nor less congruent (or incongruent) than, say, a blue circle. The quality of congruity does not apply to non-Stroop stimuli. However, the question of selectivity can be pursued with equal force with non-Stroop stimuli, too. Can people attend selectively to color and ignore the form of geometric shapes (D. J. Cohen, 1997)? Obviously, the Stroop effect cannot be calculated for such stimuli. Other measures (notably Garner interference; Garner, 1974) serve to assess selective attention to attributes of non-Stroop stimuli. We return to discuss these measures.

What, then, is the essential property of Stroop stimuli, the one basic property telling Stroop and non-Stroop stimuli apart? The defining feature of all Stroop stimuli is the existence of a logical relationship, compatibility or incompatibility, between their components. Each and every Stroop stimulus falls into one of the mutually exclusive and exhaustive classes of congruent or incongruent combinations. Take the original Stroop dimensions of color word and print color for an example. All conceivable combinations of a color word and a print color must result in either a congruent (the word naming its color) or an incongruent (word and color mismatch) stimulus. Take the popular displays of picture–word compounds for another example. Again, all possible combinations of a picture and a word must result in either a congruent (the word naming the picture) or an incongruent (word and picture conflict) compound. Precluded is any other type of combination. Formally, any two dimensions whose values combine to yield stimuli characterized as either congruent or incongruent are Stroop-generating dimensions and the resulting stimuli are Stroop stimuli.

When elucidating the nature of Stroop stimuli, several qualifications apply. First, the logical relationship in question, the presence or absence of congruency, depends on the task at hand. An upward-pointing colored arrow is a Stroop stimulus in a top–bottom spatial location task (Clark & Brownell, 1975), but it is not a Stroop stimulus in a color-naming task. With the nature of the task taken into account, a congruent stimulus is one in which values from two dimensions are associated with the same task-appropriate response; an incongruent stimulus results when one dimension is associated with one task-appropriate response and the other with a different task-appropriate response. Again, a third possibility does not exist. Given a color-naming task, the word *black* printed in yellow (or in black) is a Stroop stimulus, but the word *cancer* printed in yellow (or in any other color) is not a Stroop stimulus because the latter word is not associated with any color-naming response. Yet another way to describe the essential property of Stroop stimuli is as follows: For any value along one dimension (say, a color word), there must be a value along the other dimension (a print color) with which it can be combined to create a congruent stimulus. Because the word *cancer* is not a color name, it cannot be a congruent item in a color-naming task, and hence it is not a Stroop stimulus. Note that color naming may nonetheless be slower to *cancer* than to *table,* but that slowdown is not a Stroop effect.

Second, consider the property that renders dimensions to be Stroop-generating dimensions. The basic stipulation is that at least one of the constituent dimensions be semantic in nature (i.e., the dimensional values possess meaning through associations with referent stimuli). The logical relationship linking the dimensions cannot exist without such semantic involvement. Concerning the original Stroop dimensions, color word is semantic although print color is not. With respect to the Stroop-like dimensions of picture and word, both are semantic. The dimensions of abstract form and color, by contrast, are not semantic; they do not bear a logical relationship to one another, and hence their combinations do not yield Stroop stimuli. The stipulation of semantic associations also means that the participant must possess the relevant semantic knowledge; obviously, the Stroop conflict does not exist in a situation in which the color-word stimuli come in a language unfamiliar to the participant. The classes of congruent and incongruent stimuli are meaningful only in a given culture and linguistic community.

Third, consider the source of the semantic associations at the basis of the Stroop effect. Whence those associations? First and foremost, they come from the “natural” meaning that words acquire through the normal course of linguistic development. Stroop (1935) capitalized on such deep semantic knowledge in his classic experiment. However, semantic associations may be based on environmental correlations, such as those existing between *sky* and blue or *grass* and green (e.g., Klein, 1964). Associations may be incidental, such as that typically existing between *chocolate* and brown. And associations can be created by symbolic interpretations that stimuli acquire through experiences in a given environment, such as that between the word *stop* and the shape of octagon (the corresponding traffic sign in some countries). One should be careful to discern such symbolic interpretations to detect the presence of Stroop dimensions. Thus, size and form are not Stroop dimensions when presented as attributes of geometric or of nonsense shapes, but they are Stroop dimensions when form is closely associated with known stimuli of characteristic sizes (e.g., elephant, mouse). The various sources of semantic associations likely create Stroop conflicts at different levels of intensity. However, the dimensions are Stroop dimensions in all cases by virtue of the fact that the pertinent stimuli divide into congruent and incongruent combinations.

We stress the logical structure of Stroop stimuli to bring into view two important points. First, because Stroop stimuli are classified by logical analysis, congruent and incongruent stimuli exist preexperimentally, independent of the person’s performance. The person may or may not exhibit a Stroop effect, but her or his performance does not alter the a priori classes of congruent and incongruent combinations. In this sense, the Stroop conflict or agreement is said to reside in the stimulus (Garner, 1974). Second, this a priori classification serves to define the behavioral Stroop effect, the difference in performance between congruent and incongruent combinations. The presence of a Stroop effect betrays the fact that the participant conferred psychological reality on the logical classes of congruent and incongruent combinations. An immediate corollary is that the Stroop effect is not applicable to dimensions and stimuli that lack the quality of conflict or correspondence. The dimensions and stimuli used in the emotional Stroop tasks belong in this latter class.

The Emotional Stroop Effect: A Listwide Phenomenon

The emotional Stroop effect is the difference in color-naming performance between emotional (e.g., the word *death* printed in red) and neutral (e.g., the word *door* printed in red) stimuli. Because the words are not color words, these stimuli lack the logical relationship of correspondence or conflict between their attributes. The word *disease* printed in blue is neither more nor less congruent than the word *lecture* presented in pink. The stimuli do not divide into congruent and incongruent combinations. Consequently, the Stroop effect cannot be calculated in studies of the emotional Stroop effect.

The extent of the difference between the two effects is revealed upon further scrutiny. In the classic Stroop effect, semantic associations bond each value along one dimension with a specific value along the other dimension. As a result, one can actually calculate a Stroop effect at the level of each specific item (Jacoby, Lindsay, & Hessels, 2003). The typical listwide Stroop effect is the sum of the specific effects observed with the individual items. Consider the word *red* appearing in two possible colors, red (matching) and blue (mismatching). The difference in color-naming performance with the two colors gives the item-specific Stroop effect with the word *red.* The same calculation can be performed with all the other items in the stimulus ensemble. In a basic sense, the Stroop effect is an item-based phenomenon. In the emotional Stroop effect, by contrast, the dimensional values are not linked by unique semantic associations. The word *death* is not associated with (or dissociated from) any one particular print color. One cannot derive item-specific emotional Stroop effects! The emotional Stroop effect is a list-based phenomenon.

Finally, consider the implications of null effects obtained in each task. Observing the lack of a Stroop effect, one can safely conclude that the participant has ignored the carrier words. Observing the lack of an emotional Stroop effect, one cannot conclude that the participant has ignored the carrier words. It is eminently possible to exhibit no emotional Stroop effects with the participants fully attending to all of the words. The absence of the emotional Stroop effect means that the participants attended to or ignored emotional and neutral words to the same extent. To be able to conclude that the task-irrelevant words were unattended requires the additional control of nonword stimuli (e.g., colored strings of letters such as *XXXX*). This control is gratuitous in the classic task (in the classic task, incidentally, responses are typically faster on congruent stimuli than on *XXXX* stimuli).

The upshot is clear: The dimensions in the emotional Stroop task lack the semantic conflict or agreement that lies at the heart of the classic Stroop effect. The dimensions tested in the emotional Stroop paradigm are not Stroop-generating dimensions, and the stimuli are not Stroop stimuli. The Stroop effect is not defined in the emotional Stroop task.

Alternative Conceptions of the Origin of the Emotional Delay

The presence of an emotional Stroop effect shows that the participant noticed the meaning of the task-irrelevant words (the absence of the effect, we recount, is ambiguous). The presence of the effect shows that the content of the emotional word prolonged the time needed to name the color in which the word is printed. Whence this emotional delay? Conventional theorizing has at least tacitly relied on the same processes of attention that generate the classic Stroop effect. A competing account attributes the emotional delay to a general slowdown of activity in the presence of threat.

According to the first approach, the carrier word, irrelevant to the task at hand, commands attention (thereby compromising full focusing on color), but it does so to a different extent depending on its content. Selective attention fails for color, but it fails more for the color of words with an emotional content. Attention to emotional words is especially acute and the toll taken by this extra processing is expressed in longer color-naming latencies. The processing of emotional words is said to be even more strongly automatic than that of the other words; the former are processed in an obligatory fashion and intrude on other concurrent activities (McKenna & Sharma, 1995). On this account, the emotional Stroop effect bears special kinship with the dimensions of word and color much the same way that the classic effect is associated with these particular dimensions. The composition of the dimensions (one semantic, the other physical) and their respective modes of processing (automatic vs. controlled) account for the failure to focus on color in the emotional task. It is the fine interplay between word and color that produces the effect much the same way that this interplay generates the classic effect. In their influential review, Williams et al. (1996) suggested that the same connectionist network generates both effects. The attribution to J. R. Stroop thus is interpretative, reflecting the operation of a common attention mechanism as the basis of the emotional and classic effects.

According to the alternative account, the slowdown observed in naming the color of emotional words is unrelated to the Stroop effect. The slowdown is related instead to the exposure to threatening stimuli. Converging evidence from the cognitive, social, and physiological realms shows that humans share a disposition to preferentially direct resources toward threatening stimuli at the expense of performing an ongoing activity (Öhman, Flykt, & Esteves, 2001). A dedicated system (Öhman, 1993) captures threatening stimuli in an automatic fashion and prioritizes processing, thereby precipitating the temporary interruption of ongoing activity (e.g., color naming). This automatic response to biologically relevant threats shares many features with the well-known orienting response (Sokolov, 1963) as well as with "automatic vigilance" (Pratto, 1994; Wentura, Rothermund, & Bak, 2000), documenting the omnipotence of negative information and the concomitant slowdown in ongoing activity. The emotional Stroop effect reflects the activity of a general-purpose defense mechanism that responds to threat by temporarily freezing all ongoing activity.

On this account, the delay observed in the emotional Stroop task does not bear any particular relationship to words and print colors. Any ongoing activity is delayed by the presence of threatening stimuli. The delay has been observed with a range of ongoing behaviors apart from color naming (attitude evaluation, visual search, motor approach and avoidance) in response to a range of threatening stimuli other than emotional words (faces, animals, pictures). By the same token, there does not exist an interplay between words and colors, because all concurrent activities are disrupted by the emotional stimuli, color naming as well as reading. To recap, the emotional Stroop effect is mediated by a preattentive or early-attentive inhibitory mechanism associated with threat. This mechanism is independent from the selective-attention mechanism mediating the Stroop effect.

How can one decide between the theoretical alternatives? The following five contrasts are helpful by way of a theoretical resolution.

Five Diagnostics

1. Variation Along the Task-Irrelevant Dimension of Word

The classic Stroop effect is eminently sensitive to the presence of irrelevant variation. The effect virtually collapses when such variation is absent. Let the participant name print colors in a block of trials (baseline) in which the irrelevant dimension of word is held constant (e.g., the word *red* is always presented with its print color varying from trial to trial in a random fashion). Let the participant name the colors in another block of trials (filtering) in which the carrier words change from trial to trial, too. When the Stroop effect is calculated in each of the two blocks, that in the filtering condition is appreciably larger than that in the baseline condition (Algom, Dekel, & Pansky, 1996; Melara & Algom, 2003; Melara & Mounts, 1993; Pomerantz, 1983). The extreme sensitivity of the Stroop effect to irrelevant variation reflects its nature as a measure of selective attention. Is the emotional Stroop effect similarly sensitive to irrelevant variation? Note that the emotional task itself does not mandate trial-to-trial variation in the carrier words. A single emotional word varying in color can be presented in one block and a single neutral word varying in color in another block. The difference in color-naming performance between the two blocks yields the emotional Stroop effect at the baseline condition. This effect can be compared with that derived from another pair of blocks (one emotional, another neutral) in which the words do vary from trial to trial (the filtering condition, the traditional way for calculating the effect). Does the presence (filtering) or absence (baseline) of irrelevant variation influence the emotional Stroop effect the way it shapes the classic effect? Does the emotional effect vanish in the absence of irrelevant variation?

2. Asymmetry of Interference to Color and Word

A hallmark of the classic Stroop effect is the asymmetric pattern of interference observed. Color naming is affected by the irrelevant words, but word reading is not similarly affected by the irrelevant print colors (although reversed Stroop effects are obtained under some conditions). Does a similar asymmetry in interference characterize performance in the emotional task? The Stroop effect cannot be calculated in the emotional task, but another measure of selective attention, Garner interference, can be calculated and serve as a valuable diagnostic tool. Garner interference is the difference in performance between the baseline and the filtering blocks of trials. Its presence (i.e., worse performance in the filtering than in the baseline condition) attests to the failure of full selective attention to the target dimension. The detriment in the filtering condition reflects the toll taken by task-irrelevant variation (the absence of Garner interference attests to good selective attention). For the original Stroop dimensions of color words and print colors, Garner interference reproduces the asymmetry observed with the Stroop effect: It is sizeable when the task is color naming but small when the task is word reading (e.g., Melara & Mounts, 1993; Sabri, Melara, & Algom, 2001). Therefore, gauged by Garner interference, too, selective attention to word is better than that to color. Is Garner interference larger for color naming than for word reading in the emotional task, too?

3. The Relative Salience of the Colors and Words Presented

Relative salience, or relative dimensional discriminability, refers to the psychological differences separating stimulus values along a dimension. Word and color are matched in discriminability if the speed and accuracy in classifying words (with color held constant) equal those in classifying the colors (with word held constant). The classic Stroop effect is eminently sensitive to relative salience (matched or mismatched in favor of word or of color). In large portions of classic Stroop research, the two dimensions have not been matched but rather mismatched in favor of word. This mismatch alone could have engendered the word–color asymmetry in interference because a more discriminable dimension intrudes on a less discriminable dimension more than vice versa (Pomerantz, 1983).[2] Indeed, when Stroop (1935) trained

[2] It is important to realize that manipulations of relative dimensional discriminability affect selective attention to the relevant dimension, not perception of the pertinent stimuli. Changing dimensional discriminability does not alter psychophysical performance with the pertinent stimuli. All of the stimuli are separated well beyond the constraints posed by imperfect resolution or the difference threshold. A glance at the accuracy obtained in the cited studies, including the present one, suffices to clarify this often misconceived point: It is uniformly close to 100%. Stimulus values along the less discriminable dimension are also perfectly read, named, classified, or discriminated. Modifications of dimensional discriminability make one of the dimensions more salient (and hence attention catching) than the other.

Another subtle point should also be appreciated. Dimensional discriminability is manipulated willy-nilly in every experiment. When specifying the experimental stimuli (e.g., selecting the fonts and the colors used), the experimenter unwittingly determines relative dimensional discriminability. Garner's (1962, 1974) insight made this fundamental feature explicit along with its sundry ramifications. The main point to note is that the experimenter, by creating the stimuli and by conceiving the design, indirectly governs attentional processing and influences the outcome.

Another point worth noting is that in studies of the emotional Stroop effect, the words are often more numerous than the colors. For example, the study by Gotlib and McCann (1984; see also Gotlib & Cane, 1987) entailed 150 different words but merely four print colors. The difference in set size, another potent factor of context (Melara & Mounts, 1994; Sabri et al., 2001), can itself engender the failure of selective attention to color.

participants for several days on color (rendering it more salient than word), the usual pattern of interference reversed. Recently, Algom, Melara, and their colleagues produced Stroop effects, reverse Stroop effects, and zero effects virtually at will by manipulating the relative salience of the words and colors presented (e.g., Algom, Dekel, & Pansky, 1993; Algom et al., 1996; Arieh & Algom, 1997, 2002; Chajut & Algom, 2003; Melara & Mounts, 1993, 1994; Pansky & Algom, 1999; Sabri et al., 2001). When the dimensions were mismatched in favor of word, the usual Stroop (and Garner) effects obtained. When the dimensions were matched, selective attention was good for both word and color and Stroop and Garner effects vanished. Finally, when colors were made more discriminable than words, the reverse Stroop effect obtained, with colors intruding on words but not vice versa. Is the emotional Stroop effect similarly sensitive to the relative discriminability of the words and colors presented? Does the emotional effect diminish (or vanish) when the colors are made as discriminable as the words? Is there a reverse emotional Stroop effect by which colors intrude on word reading more than do words on color naming?

4. Emotional Delay With Colors and Words

In the classic Stroop task, colors and words are yoked by a complementary pattern. Good performance with the words comes in tandem with poor performance with the colors and vice versa. Large interference in one task comes in tandem with a lesser interference in the other. For example, words intrude on color naming when they compose the more salient dimension (the Stroop effect), but colors intrude on word reading when the colors compose the more salient dimension (the reverse Stroop effect). Theoretically, too, word–color complementarity is a basic ingredient of all major accounts of the Stroop effect, including models of horse race or relative speed of processing (Posner & Snyder, 1975), automaticity (Logan, 1980), parallel distributed processing (J. D. Cohen, Dunbar, & McClelland, 1990), and tectonic theory (Melara & Algom, 2003). Does the same seesaw relationship characterize the emotional Stroop effect? Alternatively, is the effect typically confined to color naming? Does the automaticity of reading salvage it from the adverse effects of the emotional content of the words? The (tacit) acceptance of this last idea (a) fostered the theoretical allegiance with the classic Stroop effect and (b) helped narrow scrutiny to color naming in large portions of pertinent research. Color naming commanded interest in a natural fashion in the face of a presumably invariant (perhaps minimal and/or “unconscious”) reading (e.g., C. MacLeod & Hagan, 1992; Mathews & Harley, 1996; Mogg, Bradley, Williams, & Mathews, 1993; Williams et al., 1996).

However, as Contrasts 2 and 3 already imply, the examination of color naming as well as word reading is indispensable for a theoretical resolution. One must include word reading (or lexical decision) in the experiment to test whether the emotional delay affects this activity, too. A complementary relationship between word reading and color naming is expected if the emotional effect is an attention phenomenon associated with the classic effect. A general slowdown (i.e., the absence of complementarity) is expected if the emotional effect is associated with a (preattentive?) mechanism responding to threat.

5. Separate Versus Mixed Presentation of the Emotional and Neutral Words

Is the delay in processing color limited only to the emotional word presented or does it extend to affect that of the words presented after the emotional word, too? If the emotional effect is the result of word–color complementarity (that governs the classic effect), then the delay obviously is limited to the emotional word in question. However, if the delay is part of a general slowdown, then it may well extend in time to affect subsequent stimuli. Several investigators (McKenna, 1986; McKenna & Sharma, 2002; Sharma & McKenna, 2001; Suslow, Ohrmann, & Arolt, 2001; see also Öhman et al., 2001) suspected that the threat-induced slowdown persists in time to affect the processing of stimuli appearing proximately after the detection of threat. Note that these stimuli are affected regardless of their valence (e.g., the color of a neutral word appearing after an emotional one also suffers a delay in processing). The sustained effect may explain why many demonstrations of the emotional Stroop effect entail testing the emotional and neutral words in separate blocks of trials. In the separate-blocks design, each word in the emotional block (except the first) is preceded by an emotional word, but none in the neutral block is preceded by an emotional word. The difference in performance across the two conditions (the emotional Stroop effect) may derive from the effect that emotional words exert on the processing of the subsequent words. If the main determinant of the emotional Stroop effect is the persistence of this automatic vigilance, then the effect should diminish or dissipate altogether in a design in which emotional and neutral words are presented randomly mixed within a single block of trials.

These five contrasts entailing conflicting predictions were subjected to empirical examination in a series of six experiments.

Empirical Study: Color Naming and Reading in the Emotional Stroop Task

The first four experiments followed the same procedure. For color naming of emotional words, the participants performed in a baseline task in which they named the color, changing from trial to trial, of a single emotional word (a different constant word for different participants). The participants then performed in a filtering task, in which they again named the color of emotional words, but the carrier words also changed from trial to trial in an orthogonal fashion. For color naming of neutral words, the participants performed in the same two tasks, baseline and filtering, this time with neutral words. For word reading, the same four conditions were repeated. For reading emotional words, in the baseline condition, the participant read emotional words, changing from trial to trial, printed in an invariant color (different for different participants). In the filtering condition, the participants again read emotional words, but the words varied in print color. For reading neutral words, the participants performed in the same two tasks, baseline and filtering, reading neutral words. Therefore, each participant performed in eight tasks: naming colors in four and naming words in four. For each dimension, two tasks (baseline and filtering) entailed emotional words and two tasks (baseline and filtering, again) entailed neutral words. The color-naming and word-reading tasks were performed together as a set. Order of presentation was counterbalanced, with approximately half of the participants first naming colors, and half first reading words. We manipulated the amount of irrelevant variation within experiments (baseline vs. filtering) and dimensional salience across experiments (matched vs. mismatched in favor of word).

The aim of the last two experiments was to clarify the role of (a) type of semantic processing and (b) mode of presentation in generating the emotional Stroop effect. In Experiment 5, we used a lexical decision task in place of the word-reading task to test whether the emotional Stroop effect obtains for this type of semantic processing, too. In Experiment 6, we compared the emotional Stroop effects derived in separate and mixed presentations of the emotional and neutral words.

Collectively, therefore, in the present experiments we attempted to address the following six questions: (a) Is the emotional Stroop effect sensitive to task-irrelevant variation? (b) Does the emotional effect display a word–color asymmetry in interference? (c) Is the emotional Stroop effect sensitive to the relative salience of the words and colors presented? (d) Do emotional words hinder reading as well as color naming? (e) Is the emotional Stroop effect sensitive to the way that the emotional and the neutral words are presented (i.e., separately or mixed in a random fashion)? (f) Does the emotional delay extend to lexical decision (apart from color naming and reading)?

Experiment 1

Method

Participants. The participants were 23 Tel Aviv University undergraduates, 15 women and 8 men, paid to perform in the experiment. All had normal or corrected normal vision (a stipulation that applied to participants in all the subsequent experiments) and all were native speakers of Hebrew.

Stimuli, apparatus, and design. The emotional stimuli were the Hebrew equivalents of the words *crisis, fear, betrayal, failure,* and *danger*. The neutral stimuli were the Hebrew words for *field, avenue, footway, neighborhood,* and *passage*. Note that the neutral words selected formed a single category (living environment), as do the emotional stimuli. This control rules out an explanation of a performance difference between emotional and neutral stimuli as being a category effect. The two sets of words also matched in word length and frequency on the basis of an informal cumulative list compiled over the last couple of years at Tel Aviv University and on the ratings of 14 judges on a 0 (*not familiar, or barely familiar*) to 5 (*very, very familiar*) scale. We included words of average frequency (i.e., those rated 3 or 4 by all of the judges). We also used ratings by the same judges to decide category membership of the neutral stimuli (affirmative responses by at least 10 judges was required for inclusion). The words appeared in the following print colors: red, green, brown, gray, and blue. Note that we used as many words as colors (five each) to keep set size comparable across the two dimensions.

From the entire set of stimuli, we created eight experimental tasks: four entailing naming colors and four entailing reading words. Of each set of four tasks, two involved emotional words and two involved neutral words. For each task with each type of stimuli, one condition was the Garnerian baseline block, the other the Garnerian filtering block. To recap, in the baseline tasks, participants judged the stimuli on the criterial dimension (color or word) with the irrelevant dimension held constant. In the filtering tasks, the participants again judged the stimuli on the criterial dimension, but the stimuli also varied on the irrelevant dimension. Because the emotional and neutral words were grouped in separate blocks, each participant performed in two baseline tasks and in two filtering tasks for both color naming and word reading (once with emotional and once with neutral stimuli). Therefore, each participant performed in eight tasks in all.

The four color-naming tasks and the four word-reading tasks were performed together as a set. Of the 23 participants, a random 12 first named colors and the other 11 first read words. Within each set, the two tasks including emotional words and the two tasks including neutral words were performed together. Thus, for roughly half of the participants, all the emotional stimuli were presented first, whereas for the other participants, all the neutral stimuli were presented first. Finally, within each of these subsets, condition order (baseline, filtering) was again randomly determined.

A given stimulus (i.e., a particular word in a particular print color) was presented twice in each task. Therefore, there were 50 stimuli (5 words $\times$ 5 ink colors $\times$ 2 repetitions) in each of the filtering tasks and 10 stimuli (1 word $\times$ 5 ink colors $\times$ 2 repetitions) in each of the baseline tasks. Prior to the experimental trials and unbeknownst to the participant, he or she performed in 5 trials of that task as practice. Intervals of approximately 2 min separated the various tasks. All in all, each participant performed in 140 trials of naming colors and in another 140 trials of naming words.

The stimuli were generated in Microsoft Word (in Hebrew font Ramat-Gan, 48-point type) by a Macintosh G3 300 computer and displayed on a 17-in. color monitor set at a resolution of 1,024 $\times$ 768 pixels. Using the standard palettes, we created prototypical colors of red, green, brown, gray, and blue. The words appeared in color within the invisible frame of a 118- $\times$ 40-pixel rectangle; viewed from a distance of approximately 60 cm, they subtended 4.48° of visual angle in width and 1.52° in height. The stimuli appeared over the light gray background of the screen, approximately at its center. To avoid adaptation or strategic responding (e.g., fixating on a small portion of the print to avoid reading when naming the colors), we introduced a trial-to-trial spatial uncertainty of 50 pixels around the target location.

Procedure. The participants were tested individually in a dimly lit room. Each performed in the eight tasks in a prescribed, randomly varying order. For tasks of color naming, the participants were instructed to name the print color of the word by speaking into an Apple external microphone as fast and as accurately as possible. For tasks of word reading, the participants were instructed to name the word aloud into the microphone, again as quickly and accurately as possible. Stimuli were terminated by the participant's response. Each vocal response triggered an interrupt, and latency from stimulus onset to voice-activated interrupt was measured in milliseconds. An experimenter sitting in the room but having no visual contact with the participant and the screen recorded the vocal response on each trial. The interval between the participant's response and the appearance of the subsequent stimulus was 900 ms. Trials with response times slower than 2 s or faster than 180 ms were repeated later in the task; fewer than 1.5% of the trials were repeated.

Data analysis. The error rates were very low throughout the first five experiments; they were 1.5% for color naming and 1.3% for word naming in this experiment. As a result, we do not discuss accuracy further in this article. Another very important result of the preliminary analysis was the total absence of order effects in the data. There was no main effect of task order, whether color naming versus word reading, emotional versus neutral words, or baseline versus filtering. Order did not interact with any other variable ($F < 2.91$, $p > .05$, in all cases; see McKenna & Sharma, 1995, for a similar outcome, but see also McKenna, 1986). The absence of an order effect was found in all five experiments ($F < 3.49$, $p > .05$, in all cases). This pair of (null) results greatly simplifies the discussion of the main data with respect to the various effects found in color naming and word reading.

Results

The data for color naming and word reading in the four conditions of each task are shown in Figure 1. Plotted are the mean reaction times for correct responses in the various conditions. A glimpse at Figure 1 reveals that word reading was considerably

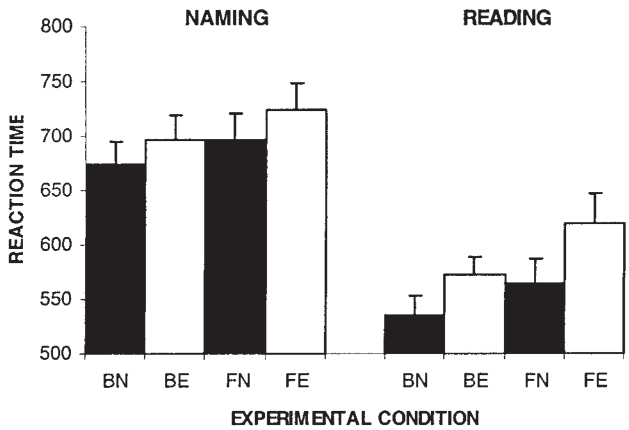

Figure 1. Mean reaction time in milliseconds needed to read or to name the print color of neutral and emotional words in each of two experimental conditions, baseline and filtering, in Experiment 1. The respective differences between (a) the second and first columns and (b) the fourth and the third columns give two values of the emotional Stroop effect in both reading and naming. The respective differences between (a) the third and the first columns and (b) the fourth and the second columns in each task give values of Garner interference for neutral and emotional words. BN = baseline task with neutral words; BE = baseline task with emotional words; FN = filtering task with neutral words; FE = filtering task with emotional words. Error bars represent one standard error around the mean.

faster than color naming, with an overall difference of 126 ms favoring reading, $F(1, 22) = 78.37$, $MSE = 9{,}139.02$, $p < .0001$, $\eta^2 = .87$. The lack of interaction of task with any of the other variables ($F < 1$ for both) confirms the speed advantage of reading in all of the experimental conditions. In the baseline conditions with neutral words, reading was faster by 138 ms than naming, $t(22) = 7.83$, $p < .001$, $\eta^2 = .731$. In the baseline conditions with emotional words, the corresponding difference was 125 ms, $t(22) = 7.75$, $p < .001$, $\eta^2 = .732$. The differences in favor of reading amounted to 132 ms, $t(22) = 7.1$, $p < .001$, $\eta^2 = .696$, and to 104 ms, $t(22) = 5.64$, $p < .001$, $\eta^2 = .591$, in the filtering conditions with neutral and emotional words, respectively.

Another notable feature of the data for both naming and reading is the presence of Garner interference with both the neutral and the emotional stimuli, $F(1, 22) = 15.4$, $MSE = 2{,}811.56$, $p < .001$, $\eta^2 = .411$, for the main effect of experimental condition (baseline, filtering). For a given task, performance in the filtering condition was always worse than performance in the baseline condition. For color naming of neutral stimuli, Garner interference amounted to 23 ms, $t(22) = 3.4$, $p < .01$, $\eta^2 = .344$. For color naming of emotional words, the respective means for baseline and filtering yielded a Garner interference of 26 ms, $t(22) = 3.1$, $p < .01$, $\eta^2 = .304$. For reading neutral words, Garner interference was 29 ms, $t(22) = 1.99$, $p = .059$. Finally, for reading emotional words, Garner interference amounted to 48 ms, $t(22) = 2.8$, $p < .02$, $\eta^2 = .263$.

The most striking result to emerge from this experiment—the slowdown in color naming and in reading the emotional stimuli—is readily apparent in Figure 1. The same staircase gestalt (created by the difference between emotional and neutral stimuli) characterized performance with the two tasks despite the large difference in speed. There emerged a main effect of stimulus type (emotional, neutral) in the omnibus analysis, $F(1, 22) = 11.99$, $MSE = 8{,}085.95$, $p < .0001$, $\eta^2 = .353$. Notably, the lack of an interaction between stimulus type and task ($F < 1$) confirmed that the emotional slowdown applied with equal force to color naming and to word reading.

Consider first color-naming performance. In the baseline tasks, a single word was presented in all of the experimental trials, and the participant's task was to name its print color in the different trials. Color naming was faster by 24 ms when the invariant word was neutral than when the invariant word was emotional, $t(22) = 2.05$, $p < .05$, $\eta^2 = .16$. This emotional Stroop effect is the first derived for invariant word stimuli. In the filtering tasks, color naming was faster by 27 ms with neutral words than with emotional words, $t(22) = 5.24$, $p < .001$, $\eta^2 = .555$. This latter emotional Stroop effect is the one routinely reported in the literature. Consider next word-naming performance. In the baseline tasks, the participants read neutral words printed in an invariant color faster by 36 ms than they did emotional words printed in an invariant color, $t(22) = 2.2$, $p < .05$, $\eta^2 = .18$. This slowdown is an emotional Stroop effect with a single color for word reading. In the filtering tasks, too, mean reading latency was faster by 55 ms with neutral words than with emotional words, $t(22) = 5.15$, $p < .001$, $\eta^2 = .547$. This threat-induced slowdown in reading was the largest emotional Stroop effect observed in the entire experiment.

Discussion

The presence of an emotional Stroop effect for reading is an important feature of the present results. The emotional stimuli adversely affected color naming, but they also adversely affected reading. The effects were comparable despite the fact that reading was much faster than naming. Augmenting the reading data by 126 ms (the average advantage of reading over color naming), the data on the right half of Figure 1 reproduce those in the left half. The results further indicate that variation of the carrier words is not a necessary condition for producing the emotional Stroop effect. Naming the print color of a single emotional word suffices to engender the slowdown in performance. Prior exposure to emotional stimuli or prior experiences of reading cannot account for the present results because of the absence of any order effects. Finally, the large mismatch in difficulty of the two tasks shows that words are more discriminable than colors in the standard emotional Stroop experiment.

Of the conflicting predictions recounted in the introduction, the data support the hypothesis that the emotional Stroop effect is not sustained by the same mechanism of attention that produces the classic Stroop effect. Several features of the data converge on this conclusion. Foremost, of course, is the presence of an emotional effect in reading. This outcome indicates that threatening stimuli disrupt all ongoing activity. A generic inhibitory mechanism is implicated, not a selective attention mechanism. The latter entails color–word complementarity, a feature conspicuously missing from the data. The insensitivity of the emotional slowdown to relative dimensional discriminability is another strong indicator for the absence of a deliberate attention mechanism (cf. Öhman et al., 2001). A mismatch of the size recorded in Experiment 1 exerts appreciable influence on the classic Stroop effect. Additional evidence comes from the results obtained in the baseline conditions.

The absence of irrelevant variation in word did not affect the emotional Stroop effect. The classic Stroop effect is strongly affected by the absence of irrelevant variation.

The conventional construal of the emotional Stroop effect rests on the twin components of (a) a slowdown in color naming accompanied by (b) an invariant semantic processing ("unconscious reading"?). The present results challenge this conception, showing the second component to be inconsistent with the data. Given the theoretical burden of these results, we decided to attempt a replication using a fresh group of participants performing only in the filtering tasks, the typical preparation used in the literature.

Experiment 2

Method

Participants. The participants were 15 Tel Aviv University undergraduates, none of whom had performed in Experiment 1.

Stimuli, apparatus, and design. The stimuli and the apparatus were those of Experiment 1. However, the participants performed only in the filtering tasks in which both the words and the colors varied in a random fashion. There were two tasks of color naming: one with neutral words, the other with emotional words. And there were two complementary tasks of reading: one with neutral words, the other with emotional words. Therefore, each participant performed in four tasks in all. A randomly selected group of 8 participants did color naming first, whereas the other 7 participants did word reading first. Within each task, order of performance (neutral, emotional) was randomly determined.

Procedure. The procedure was that for the filtering tasks in Experiment 1.

Results

Figure 2 depicts the results of Experiment 2. Salient to visual inspection, again, is the large asymmetry in performance across the two tasks. On average, word reading was faster by 104 ms than color naming, $F(1, 14) = 49.75$, $MSE = 3{,}246.31$, $p < .0001$, $\eta^2 = .78$. The advantage of reading over naming applied equally to neutral (103 ms), $t(14) = 6.3$, $p < .001$, $\eta^2 = .739$, and to emotional (105 ms), $t(14) = 5.44$, $p < .001$, $\eta^2 = .678$, stimuli.

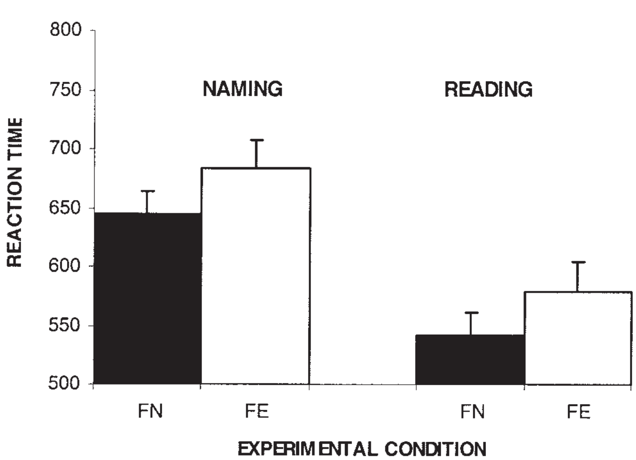

Figure 2. Mean reaction time in milliseconds needed to read or to name the print color of neutral and emotional words in Experiment 2. The differences between the second and first columns give values of the emotional Stroop effect in both reading and naming. FN = filtering task with neutral words; FE = filtering task with emotional words. Error bars represent one standard error around the mean.

The large difference in performance between the two tasks contrasts with the presence of an invariant emotional Stroop effect within each task. The main effect of stimulus type (neutral, emotional), $F(1, 14) = 23.42$, $MSE = 952.60$, $p < .001$, $\eta^2 = .626$, did not interact with that of task ($F < 1$), underscoring the uniformity of the emotional slowdown across color naming and word reading. The participants named the color of neutral words faster by 39 ms than they did those of emotional words, $t(14) = 2.7$, $p < .05$, $\eta^2 = .342$. Virtually the same emotional Stroop effect (37 ms) emerged for reading, the much faster performance notwithstanding, $t(14) = 3.47$, $p < .005$, $\eta^2 = .462$. The symmetry of the emotional Stroop effects amid the asymmetry in dimensional discriminability is remarkable.

Discussion

The results of this experiment show that the emotional Stroop effect for reading is a dependable phenomenon. It is probably as robust as the emotional Stroop effect for color naming.

In Experiment 3, we attempted another replication to rule out a possible confounding. In Experiments 1 and 2, we were careful to match the neutral and emotional words in length, frequency, and category. Recent studies of the emotional Stroop effect have increasingly applied such controls. In Experiment 3, we applied an additional control: We also matched the two sets of words on concreteness or "context availability" (Schwanenflugel, Harnishfeger, & Stowe, 1988). It is well known that concrete words or words high in context availability tend to be processed faster than abstract words or words low in context availability (e.g., Holmes & Langford, 1976; Klee & Eyesenk, 1973; Schwanenflugel & Shoben, 1983). In many studies of the emotional Stroop effect, the emotional words have perhaps been more abstract than the neutral words were. If so, then the emotional content effects may simply be context availability effects. To address the issue of context availability, in Experiment 3 we used another set of neutral words with context availability equal to that of the emotional words.

Experiment 3

Method

Participants. An independent sample of 14 Tel Aviv University undergraduates took part in the experiment.

Stimuli, apparatus, and design. The apparatus and design were those of Experiment 2. To select the neutral stimuli, we recruited a new group of 24 judges to decide the context availability of sets of neutral words sharing category (as well as that of the emotional words from Experiments 1 and 2). To define context availability, we followed the guidelines of Schwanenflugel et al. (1988). We decided to gauge context availability and not concreteness per se because Schwanenflugel et al. have shown the former concept to be more general than the latter. On the 0- to 5-point scale used, both the emotional words from Experiments 1 and 2 and the new neutral words—the Hebrew equivalents for *trial, law, state, regime,* and *leader*—averaged between 3.5 and 4.6 on context availability (i.e., were relatively highly available or abstract). Note that the new set of neutral words forms a single category (political science). Apart from the new set of neutral words, the design was precisely that of Experiment 2.

Procedure. The procedure was that of Experiment 2.

Results

Figure 3 gives the results. Notable, again, is the sizeable difference between reading and color naming for these stimuli strictly matched on abstractness. Overall, word reading was faster by 84 ms than color naming, $F(1, 13) = 47.61$, $MSE = 3{,}656.64$, $p < .0001$, $\eta^2 = .785$. In the face of this difference in difficulty, the intratask effects of threat were remarkably similar. Emotional stimuli took longer to name and read than neutral stimuli, $F(1, 13) = 36.72$, $MSE = 1{,}062.72$, $p < .001$, $\eta^2 = .738$, and the lack of an interaction between emotionality and task ($F < 1$) confirmed the symmetry of the emotional delays. For color naming, the emotional Stroop effect amounted to 32 ms, $t(13) = 2.64$, $p < .05$, $\eta^2 = .349$. For word reading, the toll exacted by emotional content was 27 ms, $t(13) = 2.99$, $p < .05$, $\eta^2 = .407$.

Discussion

Reading again yielded the same emotional delay observed with color naming. One does not reap a theoretical gain by shunning reading in experiments of the emotional Stroop effect. Effects of word length, frequency, category, and abstractness cannot account for the robust emotional delay observed in this experiment.

Does the emotional delay remain intact when one changes the relative salience of the colors and words presented? The most visible feature of the collective data of Experiments 1–3 is the large mismatch in discriminability favoring word. In Experiment 4, we made a concerted effort to redress the dimensional imbalance by making color and word equally salient.

Experiment 4

The words (and colors) presented in Experiments 1–3 have been widely used in the literature. The data show that, without taking precautions, the routine preparation results in stimuli that are much more easily read than named. Can the usually observed emotional disruption be explained, partly or fully, by this large mismatch in discriminability? Does the emotional delay diminish when precautions are taken to make the words and colors equally salient to perception? In Experiment 4, we rendered reading and color naming equally difficult (or easy) by slowing down the former. The classic Stroop effect is malleable (to the point of elimination) in response to manipulations of relative dimensional discriminability. Does the emotional Stroop effect display the same sensitivity to the relative difficulty of the two tasks?

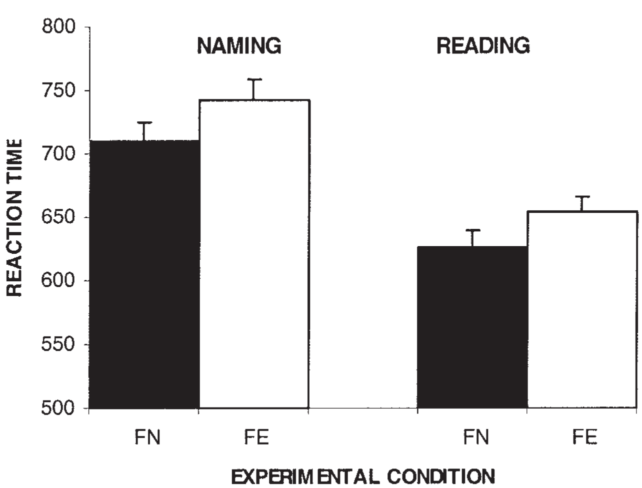

Figure 3. Mean reaction time in milliseconds needed to read or to name the print color of neutral and emotional words in the filtering condition in Experiment 3. The differences between the second and first columns give values of the emotional Stroop effect in both reading and naming. FN = filtering task with neutral words; FE = filtering task with emotional words. Error bars represent one standard error around the mean.

Method

Participants. An independent sample of 15 Tel Aviv University undergraduates were paid to perform in the experiment.

Stimuli, apparatus, and design. The stimuli, apparatus, and design were those of Experiment 1. We slowed down reading through changes in font and visual angle. We now presented words in the Hebrew font Gilboa (using italics, bold print, and 12-point type). In addition to these surface changes, we introduced spaces of 0.8 cm between the consecutive letters of a word. This resulted in a larger width (5.22° of visual angle) of the imaginary rectangle enwrapping the stimuli. Apart from these changes in word appearance, the design followed precisely that of Experiment 1.

Procedure. The procedure was that used in Experiment 1.

Results

Figure 4 gives the summary of the results. A glimpse at Figure 4 reveals a greater balance in performance across the two tasks. It turned out that we were singularly successful at matching the difficulty of word reading and color naming: Both averaged precisely 656 ms. At the level of the individual conditions, reading was faster than color naming in the baseline conditions for neutral words, $t(14) = 4.4$, $p < .001$, $\eta^2 = .58$, and color naming was faster than reading in the filtering conditions for emotional words, $t(14) = -3.8$, $p < .001$, $\eta^2 = .507$. One should realize that differences among the individual conditions are expected in the face of the global matching because of the presence of the emotion and Garner effects.

Consider the toll taken by task-irrelevant variation. For naming the color of neutral words, Garner interference was 33 ms, $t(14) = 3.21$, $p < .01$, $\eta^2 = .424$. For naming the color of emotional words, the respective interference was 36 ms, $t(14) = 3.1$, $p < .01$, $\eta^2 = .407$. For reading neutral words, Garner interference amounted to 65 ms, $t(14) = 6.22$, $p < .001$, $\eta^2 = .734$. The comparable value for reading emotional words was 60 ms, $t(14) = 3.25$, $p < .01$, $\eta^2 = .43$. Regardless of the experimental task, word reading or color naming, the presence of variation in the irrelevant dimension impaired performance, $F(1, 14) = 41.57$, $MSE = 1{,}686.03$, $p < .005$, $\eta^2 = .748$.

The most important feature of the data was the presence of emotional disruption for both color naming and word reading, the matched difficulty notwithstanding. The main effect of stimulus type, $F(1, 14) = 29.96$, $MSE = 3{,}742.30$, $p < .0001$, $\eta^2 = .681$, is modified somewhat by a borderline interaction with task, $F(1, 14) = 4.45$, $MSE = 5{,}768.83$, $p = .053$, $\eta^2 = .241$, indicating greater disruption in reading than in color naming. For color naming in the baseline tasks, the difference between the neutral and emotional words was 30 ms, $t(14) = 1.45$, $p = .17$. For color naming in the filtering tasks, the respective value of emotional

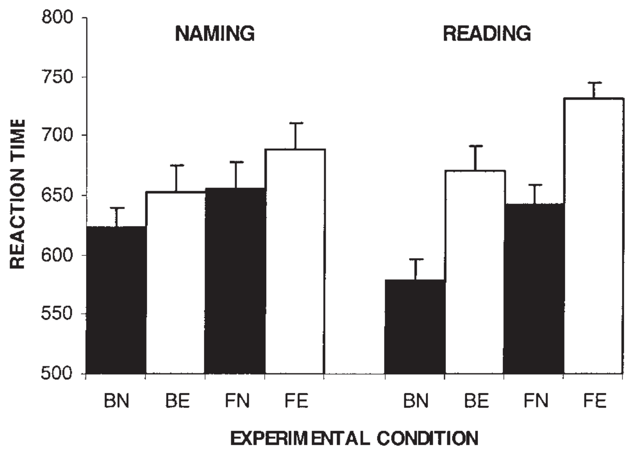

Figure 4. Mean reaction time in milliseconds needed to read or to name the print color of neutral and emotional words in each of two experimental conditions, baseline and filtering, in Experiment 4. The respective differences between (a) the second and first columns and (b) the fourth and the third columns give two values of the emotional Stroop effect in both reading and naming. The respective differences between (a) the third and the first columns and (b) the fourth and the second columns in each task give values of Garner interference for neutral and emotional words. BN = baseline task with neutral words; BE = baseline task with emotional words; FN = filtering task with neutral words; FE = filtering task with emotional words. Error bars represent one standard error around the mean.

disruption was 33 ms, $t(14) = 1.88$, $p < .05$, $\eta^2 = .201$. For reading in the baseline tasks, the emotional effect amounted to 93 ms, $t(14) = 4.73$, $p < .001$, $\eta^2 = .615$. The respective value was 88 ms for word reading in the filtering tasks, $t(14) = 3.92$, $p < .005$, $\eta^2 = .523$. The tendency for greater emotional disruption in reading noted in Experiment 1 was more visible in the present experiment.

Discussion

Methodological issues have proven both critical and limiting in research on the emotional Stroop effect. This experiment entailed a novel manipulation by which the colors and the words were equally salient to perception. Despite the altered pattern of task difficulty, there was again clear evidence that the emotional disruption affected both color naming and reading. Figure 4 illustrates roughly the same pattern of disruption found in Figures 1–3. Notable is the invariance of the emotional Stroop effect in color naming: The effects at the baseline and filtering tasks were 30 and 33 ms in Experiment 4 compared with 23 and 27 ms, respectively, in Experiment 1. Redressing the dimensional imbalance of Experiments 1–3 (and, presumably, of the vast majority of experiments published on the emotional Stroop effect) did not change the pattern of emotional interference. The results support the idea that the emotional Stroop effect is not truly associated with the classic Stroop effect.

Experiment 5

We have demonstrated substantial modulation of reading as a function of word meaning. The slowdown in reading emotional words duplicated that observed for color naming, the latter routinely documented in the literature. The emotional modulation is thought to be a nonspecific process affecting all ongoing activity. Given the theoretical implications of this conclusion, it is important to rule out potential confounding and artifacts. This was the goal of Experiment 5.

Although the sets of words used in Experiments 1–4 varied somewhat (i.e., a new set of neutral words was presented in Experiment 3), the key finding of reading modulation was established on one set of materials. This raises the possibility that the modulation is due to some confounding differences between the two types of words. Testing with additional materials is needed for cross-validation. A second concern is the possible presence of spelling-to-sound regularity that affects word-naming responses. Another related concern is pronunciation. Suppose that by fortuitous selection, the emotional words were more difficult (and slower) to pronounce than the neutral words were. Pronunciation differences are presumably not a problem in color naming, because the same color names are used equally across word types. Yet another problem concerns the oral mode of responding. People may be reluctant to say taboo and emotional or threat words aloud. If so, then the longer reading latencies reflect this impediment, not a general slowdown. Finally, repeated presentation might have exacerbated the artifacts (assuming that they were present).

To bypass all such influences, we selected a new set of words and used a lexical decision task in place of the word-naming task. The lexical decision task is “not affected by [spelling-to-sound] regularity/consistency. . . . [because] none of the core processes involved in making lexical decisions . . . are sensitive to spelling-to-sound regularity/consistency” (Hino & Lupker, 2000, p. 167; see also Hino & Lupker, 1996; Waters & Seidenberg, 1985). The possibility that threat-induced modulation affected processes of phonological encoding is also ruled out because there is evidence to suggest that the lexical decision task is performed without such processes (Hino & Lupker, 2000; Waters & Seidenberg, 1985; see also Masson, Bub, Woodward, & Chan, 2003). Does threat affect word processing when this processing is gauged by lexical decision (a task for which factors of phonology and spelling-to-sound regularity are irrelevant)?

Method

Participants. An independent group of 14 Tel Aviv University undergraduates was tested.

Stimuli, apparatus, and design. A set of 30 threat words was collected from the pertinent literature. Another set of 30 neutral words was constructed from the general category of household items. The words were all three to five letters in length. After these 60 words were collected, a group of 18 judges was asked to rate the familiarity for each word on a 5-point scale with rating values ranging from 1 (*very unfamiliar*) to 5 (*very familiar*). On the basis of the rating data, we selected 32 words, 16 emotional and 16 neutral. The words were of medium to good familiarity (average judgments were 3.5 and 4.1 for the emotional and neutral items, respectively, $t < 1$), with ratings matched carefully across the emotional and the neutral words. In addition to the 32 word stimuli, we also constructed 32 pronounceable nonword stimuli. Preserving the sounds of the word stimuli but changing two to three of the consonants created the nonwords. Using this method, we constructed an initial set of 60 nonwords. They were presented to an undergraduate class of 122 students for judgments on “wordlikeness.” The 32 items with the highest average judgments were selected. The entire experimental set consisted of 64 stimuli.

There were two separate blocks of trials. One included the 16 emotional words and 16 nonwords interspersed in a random order (subject to the proviso that no more than 3 items of the same category appear in succession). The other block included the 16 neutral words and the remaining 16 nonwords appearing in a random order (subject to the same proviso). Order of block presentation was counterbalanced across participants. The assignment of each of the 32 nonwords to either the emotional-word or the neutral-word condition was randomly determined for each participant.

The method of stimulus generation and presentation (font, size, color, screen, visual angle, and spatial uncertainty) was that used in Experiment 1. The various words and nonwords appeared printed in the colors used in Experiment 1 to render the stimuli in the task of lexical decision comparable to those in the task of word reading from the previous experiments. Response mode was changed from oral to manual as described next.

Procedure. Participants were tested individually in a dimly lit room. They were asked to make a word–nonword discrimination for each letter-string stimulus by pressing either the word or the nonword key on the computer board (labeled keys *K* and *S*, respectively). The participants were asked to respond as rapidly and accurately as possible. To avoid emotional alert or arousal prior to the experiment and a performance advantage for neutral words, as well as to ensure that each word was seen only once by each participant, no practice trials were given. However, we were careful to verify that each participant understood well the meaning of the word–nonword classification, providing aid in the form of a pair of examples presented on cards. No feedback was given during the experimental trials. As in Experiment 1, stimuli were response terminated, but we increased the intertrial interval to 2 s. Lexical decision latencies as well as errors were recorded by the computer.

Results

Lexical decision latencies less than 250 ms or greater than 1,600 ms were removed from the analyses, affecting a mere 0.4% of the data. Performance with nonwords was less accurate than with words (mean error rates of 5.6% vs. 3.0%, $t < 1$), but the difference was not statistically dependable. Plotted in Figure 5 are mean lexical decision latencies for words and nonwords for correct responses. The slowdown of 48 ms for the emotional words was dependable, $t(13) = 4.24$, $p = .001$, $\eta^2 = .58$. Apparent in Figure 5 are the longer latencies for nonwords than for words, $t(13) = 9.11$, $p < .001$, $\eta^2 = .79$. Notably, correct responses for the nonwords were directionally slower when appearing in the block with emotional words than when appearing in the block with neutral words, but the slowdown of 26 ms was not statistically dependable, $t(13) = 1.68$, $p > .10$.

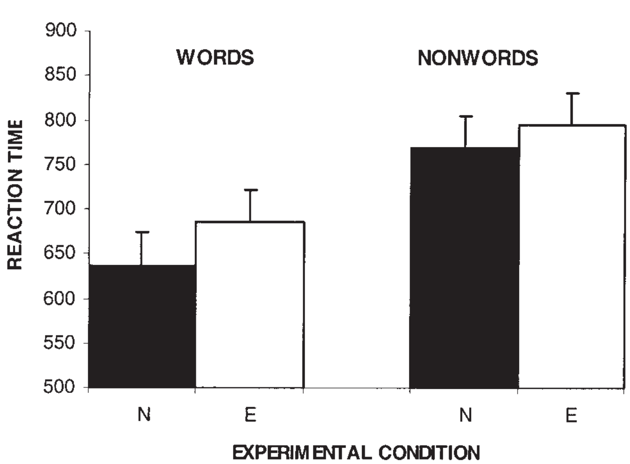

Figure 5. Mean lexical decision latencies in milliseconds to emotional and to neutral words (left-hand panel) and to nonwords appearing along with neutral and with emotional words (right-hand panel) in Experiment 5. E = block with emotional words; N = block with neutral words. Error bars represent one standard error around the mean.

Discussion

The slowdown observed in reading emotional words (Experiments 1–4) reappeared when participants were making lexical decisions with such words. Because factors of spelling-to-sound regularity and pronunciation do not play a role in the lexical decision task, the present results confer validation support on the previous results with reading. They affirm the prediction of the threat hypothesis by which modulation of word processing is an example of a generic process affecting all ongoing activity.

The present results agree with findings reported in earlier research suggesting that emotionally negative words give rise to longer lexical decision latencies (e.g., C. MacLeod, Tata, & Mathews, 1987; see also C. MacLeod & Mathews, 1991). Earlier yet, in the New Look literature, the thresholds for tachistoscopically presented threat and taboo words were reported to be higher than those for neutral words (McGinnies, 1949). The emotional slowdown was not found in a pair of lexical decision studies (C. MacLeod & Mathews, 1991), but this result

> may possibly reflect the characteristics of the word sets employed. The . . . threat words probably share closer associations than do the . . . neutral words, possibly leading to greater associative priming and/or a general expectancy for such items, and thus to a corresponding reduction in decision latency. (C. MacLeod & Mathews, 1991, p. 606)

In this experiment, we were careful to select the neutral words, too, from a single semantic category, thereby removing a confounding due to category. In C. MacLeod and Mathews's (1991) procedure, sometimes more than a single letter string per trial was presented; in this study, we used the standard preparation of a single letter string presented on each trial. We also note that in several studies (e.g., C. MacLeod & Mathews, 1991; C. MacLeod et al., 1987), the threat and neutral words were intermixed in a single block. Our participants made lexical decisions of emotional and neutral words (and nonwords) presented in separate blocks of trials. We examine the role of mode of presentation (separate blocks vs. single block) in the next experiment.

A final feature of the data deserves attention. The participants tended to make lexical decisions to nonwords appearing along with emotional words slower than they made lexical decisions to nonwords appearing along with neutral words. This tendency raises the possibility of a carryover effect in the course of processing threat-related stimuli. An emotional stimulus may affect the processing of the subsequent stimulus regardless of the nature of that stimulus. In the emotional block, many of the nonwords appeared after an emotional word (some nonwords appeared after another nonword); in the neutral block, none of the nonwords appeared after an emotional word. The persistence in time of threat-induced freezing (effective for 1–2 s; Sharma & McKenna, 2001) provides a natural explanation to the observed trend. In the next and final experiment, we further explored the cumulative effect of threat-induced modulation. Does the effect of threat survive random mixtures of emotional and neutral words?

Experiment 6

Öhman et al. (2001) described the common experience of automatic focusing "on the snake that is resting in the grass" such that one may "freeze a fraction of a second later" (p. 466) to assess the true nature of the stimulus. This freezing or slowdown of ongoing activity is sustained in time, if for a fraction of a second. Suslow et al. (2001) found that negative primes produced subsequent slowing regardless of the valence of the target words, demonstrating the persistence of "automatic vigilance" (Pratto & John, 1991; Wentura et al., 2000). Given that the effect of threat develops in time, it is in fact questionable whether the carrier words themselves are affected. The fear interpretation of the emotional Stroop effect thus entails a unique prediction: The emotional slowdown affects the stimuli that appear immediately after the presentation of the emotional words rather than the emotional words themselves.[3]

McKenna and Sharma (2002; see also McKenna, 1986; Sharma & McKenna, 2001) have drawn on the consequences of this notion to question the conventional conception of the emotional Stroop effect. They argued that the emotional Stroop effect appears when the emotional and the neutral words are presented in separate blocks of trials, but often the effect is absent or attenuated when the two types of words are mixed within a single block. Why does this difference due to procedure exist? The sustained effect of the emotional stimulus explains the difference. In a block with emotional stimuli, all stimuli but the first stimulus follow an emotional stimulus. Because the emotional delay affects the stimuli that follow an emotional stimulus, the emotional effect is large in this block. In a block with neutral stimuli, none of the words follow an emotional stimulus. As a result, an emotional delay is absent in this block. In mixed presentation within a single block, stimuli appear equally often after emotional and neutral words. Therefore, the effect of emotionality evaporates over trials. Confirming these predictions, McKenna and Sharma (2002) also recorded color-naming disruption for neutral words if they followed emotional words but not if the same words followed other neutral words. In a complementary fashion, performance was good for emotional words when they followed neutral words but was poor when the same words followed other emotional words.

Dedicated research on separate- versus single-block presentation is still scarce, and existing data concern color-naming responses. Experiment 6 extends the pertinent database by examining word naming. In Experiment 6, we tested the effect of mode of presentation—separate blocks of emotional and neutral words versus a random mixture—recording word-naming responses. Is threat-induced slowdown in reading confined to separate presentation of the two types of words? An affirmative answer would further support the threat hypothesis of the emotional Stroop effect.

Method

Participants. An independent group of 56 Tel Aviv University undergraduates took part in the experiment. They were randomly assigned to either the separate-blocks or the random-mixture condition.

Stimuli, apparatus, and design. The same set of words from Experiment 5 was used. In the separate-blocks condition, the two types of words were presented separately in two blocks of trials. One block included the 16 emotional words and the other included the 16 neutral words. There were no repeated presentations. Within each block, the words appeared in a random and different order for each participant. Order of block presentation was counterbalanced across participants. In the random-mixture condition, the same emotional and neutral words appeared interspersed in a random fashion within a single block of trials. The method of stimulus generation, presentation, and response measurement was that used in the reading condition of Experiment 1.

Procedure. Participants were tested individually in a dimly lit room. They were asked to read aloud into the microphone each word as quickly and accurately as possible. Those performing in the separate-blocks condition paused for at least 1 min between the two blocks. As in Experiment 1, stimuli were terminated on the participant's response. And, as in Experiment 1, errors were extremely rare, averaging at 0.6% and 0.4% for neutral and emotional words, respectively. We further consider reaction times in our analyses.

Results

A glimpse at Figure 6, which summarizes the data, reveals that mean word-naming latencies were longer for emotional than for neutral words in the separate-blocks condition. The emotional slowdown of 24.2 ms was dependable, $t(27) = 2.03$, $p < .05$, $\eta^2 = .132$ (one-tailed test). In contrast, word-naming latencies for the same stimuli did not differ across emotional and neutral words in the random-mixture condition.

Discussion

The slowdown in naming emotional words is eliminated when such words are mixed with neutral words in random succession. The current results with word naming duplicate those with color naming by McKenna and Sharma (2002). The results in the separate-blocks condition replicate those observed in Experiments 1–5. The present results demonstrate the same effect with another set of words, with each word seen only once by each participant. The elimination of the emotional effect when the words are mixed shows the boundary condition for obtaining the effect. This stipulation is predicted by the fear arousal interpretation of the emotional Stroop effect but not by an attention explanation akin to the classic Stroop effect.

General Discussion

The results provide for a clear decision between the contrasting predictions developed in the introduction. Thus, (a) the emotional Stroop effect is not sensitive to task-irrelevant variation; (b) the effect does not display word–color reciprocity in interference; (c) the emotional effect is not sensitive to changes in the relative

[3] There are sequential effects in the classic Stroop task, too. However, they are stimulus-specific directional effects, not a generic slowing of any subsequent response. For example, color naming on a given trial is slowed if the name of that color was the word presented in the previous trial. Conversely, color naming is faster (i.e., there is facilitation) if the to-be-ignored color word is the name of the print color presented on the previous trial. Recall also that the classic effect is always measured in a single block of trials entailing congruent and incongruent stimuli in a random order.

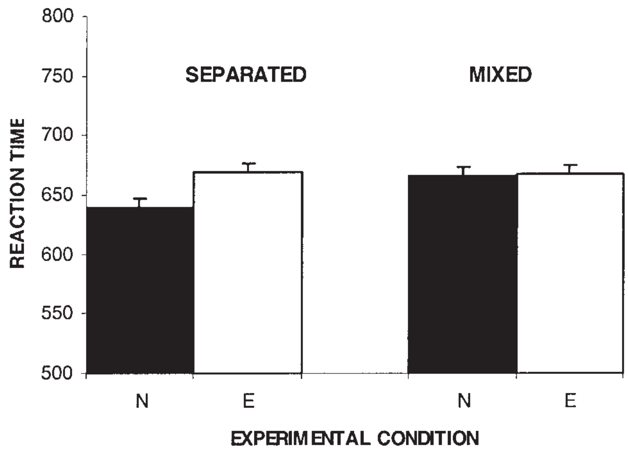

Figure 6. Mean reaction time in milliseconds needed to read emotional and neutral words presented in separate blocks of trials (left-hand panel) and randomly mixed in a single block of trials (right-hand panel) in Experiment 6. E = block with emotional words; N = block with neutral words. Error bars represent one standard error around the mean.

salience of the words and colors presented; (d) emotional words retard reading to the same extent that they do color naming; (e) the effect is sensitive to the way that the emotional and neutral items are presented, and it diminishes (or vanishes) when the different items appear in the same block; finally, (f) the emotional delay is present when the task of reading is replaced by that of lexical decision. Each of these results contrasts with the respective results obtained with the classic Stroop effect in response to the same manipulations. For (a) and (c), the classic effect is eminently sensitive to irrelevant variation and to relative salience. For (b) and (d), a hallmark of the classic effect is the asymmetry between reading and color naming. For (f), the classic effect is sensitive to alternative tasks (Smith & Magee, 1980). And for (e), Stroop-induced attention does not mandate the observed difference. In summary, the empirical dissociation joins the results of the conceptual analysis in challenging claims for a kinship between the two phenomena.

What is the nature of the threat-driven mechanism implicated by the present results? The generic slowdown observed does not necessarily imply that the underlying mechanism is that of inhibition. Resources can be allocated in particular ways such that threat-irrelevant stimuli receive fewer resources. This depletion in turn precipitates the disruption of the pertinent behaviors. There seems to be a consensus (Chen & Bargh, 1999; De Houwer, 2003; Öhman et al., 2001; Pratto & John, 1991; Wentura et al., 2000) that the threat-driven mechanism acts early and in an automatic fashion, hence the popular designation of automatic vigilance. Öhman et al. (2001) suggested a distinction between “preattentive and postattentive visual attention. The former is fast, automatic, and parallel. . . . The latter is slow, deliberate, and serial” (p. 466). We take exception to the oxymorons “preattentive attention” and “postattentive attention,” but the distinction itself is well taken and seems well established (e.g., Johnston & Dark, 1986; Jonides, 1981). Öhman et al. (2001) claimed that the early preattentive system is stimulus driven (i.e., involuntary), hence the temporary freezing of ongoing activity. It is likely that threatening stimuli are processed by the former system, Stroop stimuli by the latter system.

The most popular term used to depict the system in question seems to be that of *automatic capture of attention.* We find it noteworthy that automaticity is combined with its former nemesis, attention. In earlier conceptions, a process was either automatic or attentional. Given the convoluted history of the concept of automaticity (Bargh, 1992; Logan, 1988; Pansky & Algom, 2002), their current wedding does not come as a total surprise. On our part, we suggest using the term *automatic activation* to depict the general-purpose defense mechanism, saving *attention* for processes of attention associated with the classic Stroop effect and with spatial and object perception (see Shalev & Algom, 2000, for a theoretical distinction between various types of attention). Despite the uncertainty in the optimal nomenclature, it is clear that the processes sustaining the classic and the emotional effects differ in a qualitative fashion.

The difference between the two phenomena has not been sufficiently appreciated, producing confusion at the level of rhetoric, experimental procedure, and theory alike. For rhetoric, there is an indiscriminate use of the terms *Stroop effect, Stroop interference, Stroop experiment,* and *Stroop task* when the emotional effect is actually being tested. For experimental practice, reading has seldom been tested in studies of the emotional effect in part because of the analogy claimed with the Stroop effect. In point of fact, reading is neither invariant nor immune to interference in the classic task. In several studies of the emotional effect, color words are included in the experimental set, presumably reflecting a close association with the classic effect. An amalgamation of emotional and classic effects results.

Concerning conceptual development, theories developed for the classic effect are imported to account for the emotional effect without realizing the extent of the qualitative difference separating the two phenomena. The Stroop model of J. D. Cohen et al. (1990) is a good example of the problems met importing Stroop theory to the foreign territory of the emotional task (Williams et al., 1996). The input units of this three-layered network are assigned to the values on each of the two stimulus components (e.g., words: *red* and *green*; colors: red and green). The strength of connections along the color pathway or the word pathway is set in a training phase during which the network learns through back propagation to categorize color or word stimuli. On each trial of a later test phase, during which no learning occurs, a Stroop stimulus activates only its corresponding input units; other input units are set to zero. Attention units appropriate to task demands (i.e., color naming or word reading) enhance the target’s input to the intermediate layer. Output units corresponding to the correct target response receive excitation from the target dimension’s pathway (e.g., color) and either excitation or inhibition from the distractor dimension’s pathway (e.g., word), depending on whether the distractor is congruent (excitation) or incongruent (inhibition) with the target. The magnitude of total activation received by the output units (from both pathways) determines the network’s speed to reach an evidence threshold and to produce a response. The direction and magnitude of activation from the distractor dimension is the basis of the Stroop congruity effect in the model.

This brief depiction suffices to show that the J. D. Cohen et al. (1990) theory is tailored to model the processing Stroop stimuli. The input units are values of Stroop dimensions, the output units are the legitimate responses for both dimensions, and response production and speed depend on whether the stimulus presented is congruent or incongruent. Because congruent and incongruent stimuli drive response production in the model, the model does not apply in cases in which the two classes of stimuli are absent. Clearly, the theory cannot be applied to explain the emotional Stroop effect. The stimuli in the emotional task are not Stroop stimuli. The attributes in the emotional task do not share response options: There is no overlap between the responses to the colors and the responses to the words. The stimuli do not divide into congruent and incongruent combinations. The output units in the J. D. Cohen et al. model ("red," "green") cannot be fed by the noncolor words of the emotional task. The J. D. Cohen et al. theory was developed to model genuine Stroop phenomena: the Stroop asymmetry, the reverse Stroop effect, the effect of stimulus onset asynchrony, and, in general, the role of attention and automaticity. These are largely irrelevant in the environment of the emotional task. Therefore, we must take exception to the claim by Williams et al. (1996) that the J. D. Cohen et al. model "offers a way of understanding the mechanisms underlying emotional Stroop performance" (p. 21), and we certainly disagree with the general quest "to have models that can apply to both the original color-conflict situation and the emotional Stroop" (p. 22). The Stroop and the emotional Stroop effects are unrelated phenomena.

The dissociation accomplished in this study shows that qualitatively different processes are involved in slowing color naming to incongruent color words (e.g., *white*) and to emotional words unrelated to color (e.g., *cancer*) relative to neutral words (e.g., *table*). Color naming to *table,* in turn, can be slower than to a letter string such as *XXXX,* but that slowdown is neither a Stroop nor an emotional Stroop effect. The idea that the three processes form a continuum of speeds of attentional disengagement from words forms an attractive theoretical possibility, but one that is not supported by the data. The well-known gradient of interference derived in Klein's (1964) classic study is informed by this idea. In Klein's study, color-naming latencies were as follows (from long to short): color words, color words that were not in the response set, words with a color association (e.g., *grass*; note that all three classes entail genuine Stroop stimuli because the items divide into congruent and incongruent stimuli), high-frequency words, low-frequency words, unpronounceable nonwords, and meaningless letter strings (which served as the common yardstick to measure interference). Note that the last four classes of items do not entail Stroop stimuli. Adding to Klein's list another class of non-Stroop stimuli, emotional words, can complete a strongly ordered continuum. However, subsequent research did not replicate Klein's (1964) findings. Color naming to noncolor words is only slightly slower than to letter strings (color naming to letter strings is slower than to color-congruent words), it is slower to nonwords than to words, and it is slower to low-frequency than to high-frequency words (Burt, 2002). High-frequency words are both read and color named faster than low-frequency words. Emotional words are another example of this pattern of parallel reading and color naming: They are both read and color named more slowly than neutral words. On the basis of these and additional findings, Burt (2002) concluded that "models of the standard Stroop task offer little explanatory utility" (p. 1035) for the variations associated with color naming of noncolor words (anticipating our conclusion with respect to the emotional Stroop effect).

Finally, it is arguable that the comparable performance across reading and color naming is limited to healthy, unselected groups of participants. Our participants were drawn from the general (student) population, whereas the majority of emotional Stroop studies concerned patient populations. In these studies, color naming to emotional and to neutral words is compared (the emotional Stroop effect), and the difference is sometimes compared with that observed with normal cohorts. Typically, the emotional Stroop effect is larger with the highly anxious individuals of the special groups than with the normal cohorts. In a preliminary experiment with patients with posttraumatic stress disorder, we tested both reading and color naming. We found a larger emotional delay for color naming than for reading (although reading latencies also tended to be slower to emotional than to neutral words). The limited data available leave open the possibility that color-naming interference is greater than reading interference in selected, emotionally disturbed groups. An alternative possibility (one that we repeatedly observed with our patients) is that such individuals are reluctant to read aloud words associated with their problems and suffering. If so, then this provides a practical justification for the use of color naming in patient populations. Another problem is the markedly different latencies involved. Reading was longer by 163 ms on average with the posttraumatic stress disorder patients than with the healthy participants of Experiments 1–4. The respective difference for color naming was 367 ms. A meaningful comparison of interference across patient and healthy populations is complicated by the large difference in generic speed of responding. A dedicated study entailing patient and healthy populations tested on a variety of tasks is needed to resolve this issue.[4]

Science progresses by a never-ending process of differentiation and refinement. Phenomena considered identical or unitary are discovered to tap separate processes. Nuances and subtleties are noticed and elaborated. Eventually, the distinctions attained are reflected in the pertinent terminology and theory. In physics, light was once considered to be a wave phenomenon of vibration or interference in a medium just like sound (ether was postulated as the medium of propagation). Sound and light are fundamentally distinguished in today's physics. In cognitive science, the Stroop and the emotional Stroop effects must be similarly separated. Noticing some of the differences discussed in this article, a few authors have recently suspected that the emotional Stroop effect is not a Stroop effect at all (e.g., De Houwer, 2003; Kuhl & Kazen, 1999). We provide conceptual and empirical grounds for separating the two effects. It will be critical in future research to make explicit the assumptions underlying the emotional Stroop effect and to distinguish it carefully from the classic Stroop effect.

[4] Regardless of the resolution of this issue, the incommensureability of the Stroop and the emotional Stroop effects derived on the basis of the conceptual development apply with equal force to patient populations as well. For these populations, too, the emotional Stroop effect observed is not a Stroop effect.

References

Algom, D., Dekel, A., & Pansky, A. (1993). On the perception of number. In A. Garriga-Trillo, P. R. Minon, C. Garcia-Gallego, C. Lubin, J. M. Merino, & A. Villarino (Eds.), *Fechner Day '93* (pp. 1–6). Palma de Mallorca, Spain: International Society for Psychophysics.

Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The Stroop effect revisited. *Memory & Cognition, 24,* 557–572.

Arieh, Y., & Algom, D. (1997). Kriat milim bemesimot Stroop: Haspaat hatelut ben hamarkivim hasemantiim vehahazutiim [Reading words in Stroop-like tasks: The effect of the contingency between semantic and visual components]. *Psychologia, 6,* 7–19.

Arieh, Y., & Algom, D. (2002). Processing picture–word stimuli: The contingent nature of picture and of word superiority. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 221–232.

Bargh, J. A. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105,* 181–199.

Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance, 28,* 1019–1038.

Chajut, E., & Algom, D. (2003). Selective attention improves under stress: Implications for theories of social cognition. *Journal of Personality and Social Psychology, 85,* 231–248.

Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or to avoid the stimulus. *Personality and Social Psychology Bulletin, 25,* 215–224.

Clark, H. H., & Brownell, H. H. (1975). Judging up and down. *Journal of Experimental Psychology: Human Perception and Performance, 1,* 339–352.

Cohen, D. J. (1997). Visual detection and perceptual independence: Assessing color and form. *Perception & Psychophysics, 59,* 623–635.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

De Houwer, J. (2003). A structural analysis of indirect attitude measures. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). Mahwah, NJ: Erlbaum.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York: Wiley.

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

Gotlib, I. H., & Cane, D. B. (1987). Construct accessibility and clinical depression: A longitudinal investigation. *Journal of Abnormal Psychology, 96,* 199–204.

Gotlib, I. H., & McCann, C. D. (1984). Construct accessibility and depression: An examination of cognitive and affective factors. *Journal of Personality and Social Psychology, 47,* 427–439.

Hino, Y., & Lupker, S. J. (1996). The effects of polysemy in lexical decision and naming: An alternative to lexical access accounts. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1331–1356.

Hino, Y., & Lupker, S. J. (2000). Effects of word frequency and spelling-to-sound regularity in naming with and without lexical decision. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 166–183.

Holmes, V. M., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Learning and Verbal Behavior, 15,* 559–566.

Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin & Review, 10,* 638–644.

Johnston, W. A., & Dark, V. J. (1986). Selective attention. *Annual Review of Psychology, 37,* 43–75.

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. B. Long & A. D. Baddeley (Eds.), *Attention and performance IX* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Klee, H., & Eyesenk, M. W. (1973). Comprehension of abstract and concrete sentences. *Journal of Learning and Verbal Behavior, 12,* 522–529.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *American Journal of Psychology, 77,* 576–588.

Kuhl, J., & Kazen, M. (1999). Volitional facilitation of difficult intentions: Joint activation of intention memory and positive affect removes Stroop interference. *Journal of Experimental Psychology: General, 128,* 382–399.

Logan, G. D. (1980). Attention and automaticity in Stroop and priming tasks: Theory and data. *Cognitive Psychology, 12,* 523–553.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

MacLeod, C., & Hagan, R. (1992). Individual differences in the selective processing of threatening information and emotional responses to a stressful life event. *Behaviour Research and Therapy, 30,* 151–161.

MacLeod, C., & Mathews, A. (1991). Biased cognitive operations in anxiety: Accessibility of information or assignment of processing priorities? *Behavioural Research and Therapy, 29,* 599–610.

MacLeod, C., Tata, P., & Mathews, A. (1987). Perception of emotionally valenced information in depression. *British Journal of Clinical Psychology, 26,* 67–68.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163–203.

Masson, E. J., Bub, D. N., Woodward, T. S., & Chan, J. C. K. (2003). Modulation of word-reading processes in task switching. *Journal of Experimental Psychology: General, 132,* 400–418.

Mathews, G., & Harley, T. (1996). Connectionist models of emotional distress and attentional bias. *Cognition and Emotion, 10,* 561–600.

McGinnies, E. (1949). Emotionality and perceptual defense. *Psychological Review, 39,* 242–255.

McKenna, F. P. (1986). Effects of unattended emotional stimuli on color-naming performance. *Current Psychological Research and Reviews, 5,* 3–9.

McKenna, F. P., & Sharma, D. (1995). Intrusive cognitions: An investigation of the emotional Stroop task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1595–1607.

McKenna, F. P., & Sharma, D. (2002). *Reversing the emotional Stroop effect: The role of fast and slow components.* Manuscript submitted for publication.

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

Melara, R. D., & Mounts, J. R. W. (1993). Selective attention to Stroop dimension: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21,* 627–645.

Melara, R. D., & Mounts, J. R. W. (1994). Contextual influences on interactive processing: Effects of discriminability, quantity, and uncertainty. *Perception & Psychophysics, 56,* 73–90.

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. M. (1993). Subliminal processing of emotional information in anxiety and depression. *Journal of Abnormal Psychology, 102,* 304–311.

Öhman, A. (1993). Fear and anxiety as emotional phenomena: Clinical phenomenology, evolutionary perspectives, and information-processing mechanisms. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 511–536). New York: Guilford Press.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478.

Pansky, A., & Algom, D. (1999). Stroop and Garner effects in comparative judgment of numerals: The role of attention. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 39–58.

Pansky, A., & Algom, D. (2002). Comparative judgment of numerosity and numerical magnitude: Attention preempts automaticity. *Journal of Ex-*

perimental Psychology: Learning, Memory, and Cognition, 28, 259–274.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General, 112,* 516–540.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Pratto, F. (1994). Consciousness and automatic evaluation. In P. M. Niedenthal & S. Kitayama (Eds.), *The heart's eye* (pp. 115–143). New York: Academic Press.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology, 61,* 380–391.

Sabri, M., Melara, R. D., & Algom, D. (2001). A confluence of contexts: Asymmetric versus global failures of selective attention to Stroop dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 515–537.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decision for abstract and concrete words. *Journal of Memory and Language, 27,* 499–520.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 82–102.

Shalev, L., & Algom, D. (2000). Stroop and Garner effects in and out of Posner's beam: Reconciling two conceptions of selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 997–1017.

Sharma, D., & McKenna, F. P. (2001). The role of time pressure on the emotional Stroop task. *British Journal of Psychology, 92,* 471–481.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General, 109,* 373–392.

Sokolov, E. N. (1963). *Perception and the conditioned reflex.* New York: Pergamon Press.

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

Suslow, T., Ohrmann, P., & Arolt, V. (2001). Ausgestaltung sogennanter affektiver Voraktivierungseffekte in der evaluativen Entscheidungsaufgabe: Hinweise auf automatische Vigilanz fuer negative Informationen [Direction of so-called affecting priming effects in the evaluation tasks: Evidence for automatic vigilance for negative information]. *Zeitschrift fuer Psychologie, 209,* 137–152.

Waters, G. S., & Seidenberg, M. S. (1985). Spelling–sound effects in reading: Time course and decision criteria. *Memory & Cognition, 13,* 557–572.

Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78,* 1024–1037.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3–24.

Received October 15, 2002
Revision received April 7, 2004
Accepted April 7, 2004 ■