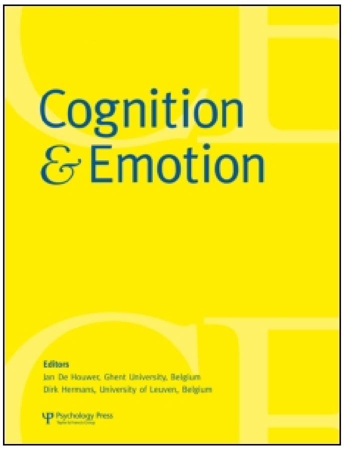

Cognition & Emotion
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713682755

Wheel chairs and arm chairs: A novel experimental design for the emotional Stroop effect

Daniel Algom [a]; Dan Zakay [a]; Ofer Monar [a]; Eran Chajut [b]
[a] Tel-Aviv University, Ramat-Aviv, Israel [b] The Open University, Raanana, Israel

First published on: 05 December 2008

To cite this Article Algom, Daniel, Zakay, Dan, Monar, Ofer and Chajut, Eran(2009) 'Wheel chairs and arm chairs: A novel experimental design for the emotional Stroop effect', Cognition & Emotion, 23: 8, 1552 — 1564, First published on: 05 December 2008 (iFirst)

To link to this Article: DOI: 10.1080/02699930802490243

URL: http://dx.doi.org/10.1080/02699930802490243

COGNITION AND EMOTION
2009, 23 (8), 1552–1564

Psychology Press
Taylor & Francis Group

Wheel chairs and arm chairs: A novel experimental design for the emotional Stroop effect

Daniel Algom, Dan Zakay, and Ofer Monar
Tel-Aviv University, Ramat-Aviv, Israel

Eran Chajut
The Open University, Raanana, Israel

The emotional Stroop effect demonstrates that people are slower to name the ink colour of emotionally charged words than that of neutral words. Attributing this slowdown to emotionality or threat requires that the emotion and neutral words presented are equal on all extraneous variables. However, it is impossible in principle to match the two types of words on all conceivable variables. To address the problem, we presented the *same* set of words, polarised to be negative in one condition and neutral in another by using appropriate prime words that produced the desired valence through idiomatic two-word expressions. Across two studies, participants took longer to name the colour of the same words when they were rendered negative than when they were rendered neutral. This difference reflects a true emotional Stroop effect given the control provided for all potentially confounding variables tied to the target stimuli. The new procedure also enables researchers to calculate item-specific emotional Stroop effects for the first time.

Keywords: Emotional and classic Stroop effects; Emotion stimulus control; Valence.

The ability to isolate the attribute of interest in a given object, excluding irrelevant or distracting attributes, is indispensable for adaptation and survival. In order to cross the street safely, one must focus on the velocity of the approaching car, ignoring momentarily its shape or colour. In order to report correctly the colour of a word, one must focus on the target ink colour

---

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

The research and the preparation of this report were supported by an Israel Science Foundation grant (ISF221–07) and by a grant from Tel-Aviv University Research Authority to DA.

We thank Dinkar Sharma and an anonymous reviewer for helpful comments on an earlier version of the article.

www.psypress.com/cogemotion DOI: 10.1080/02699930802490243

while ignoring the meaning of the word. By definition, each activity of everyday life entails an ability to attend selectively to the task-relevant attribute. How does one measure this vital ability? Consider again words written in colour with ink colour as the task-relevant attribute. If the words are colour words (i.e., words whose content refers to colours), then measurement profits from the fact that these stimuli divide into congruent (the word meaning names its ink colour) and incongruent (word meaning and ink colour conflict) combinations. Slower colour naming with incongruent than with congruent stimuli shows that the task-irrelevant words were noticed and affected performance. This failure of full selective attention to ink colour is known as the Stroop effect (Stroop, 1935; see MacLeod, 1991; Melara & Algom, 2003, for reviews).

If the words are not colour words (i.e., the meaning of the words does not refer to colour), then measurement differs. Suppose that the words written in colour are emotionally charged words such as DANGER or FAILURE and neutral words such as BOOK or TABLE. Slower colour naming with the emotion words shows that the meaning of the words was noticed, thereby compromising exclusive attention to the target ink colours. This slowdown with the emotional words is known as the emotional Stroop effect (see Williams, Mathews, & MacLeod, 1996, for a review). The phenomenon has been attracting a great deal of interest during the past two decades, with the emotional Stroop effect rivalling its namesake in sheer output of empirical research.

Much of the work with the emotional Stroop effect concerned high-anxiety individuals or various patient populations with well-defined pathologies. The precise match that can be created between the patient's affective disorder and word content enabled to uncover valuable person-specific information (e.g., McNally, Kaspi, Riemann, & Zeitlin, 1990; Richards & Millwood, 1989; Rieman & McNally, 1995; Watts, McKenna, Sharrock, & Treize, 1986; see also MacKay et al., 2004). In the general population, too, the effect has proved a valuable tool in studying processes of attention, emotion, and cognition (Algom, Chajut, & Lev, 2004; McKenna & Sharma, 1995, 2004; Whalen et al., 1998).

Despite the difference in measurement, the designation of the phenomenon as "emotional Stroop effect" reflects a belief in a close kinship between the classic Stroop effects and the emotional Stroop effect. In this study, we show that the measurement model at the basis of the emotional Stroop effect differs in a qualitative way from that at the basis of the classic Stroop effect. We further show that the emotional Stroop model can be vulnerable on counts of control and validity. Most important, we introduce a new method for generating the emotional Stroop effect that is free of threats to validity and confounding. Elucidating the common and distinctive features of the two effects bearing the name of J. R. Stroop is important on both theoretical and pragmatic grounds.

The classic and the emotional Stroop effect: Common and distinctive features

The similarities between the emotion and classic effects are obvious. Both tasks include the presentation of words in colour and both document the influence of word-meaning—irrelevant to the task at hand—on naming the ink colour. Many investigators have found these and further procedural similarities instructive and argued for a common attention mechanism to underlie both phenomena (e.g., Williams et al., 1996). Other investigators found the differences equally compelling (Algom et al., 2004; De Houwer, 2003; Kuhl & Kazen, 1999; McKenna & Sharma, 2004; see also Wyble, Sharma, & Bowman, 2007, for a recent computational model). The hallmark of the classic Stroop effect is the partition of the stimuli into congruent and incongruent combinations. The difference in colour naming between the congruent and the incongruent stimuli defines the classic effect. In the emotion effect, by contrast, the stimuli do not divide into congruent and incongruent ones. The word GUILT in blue ink is neither more nor less congruent than the word HABIT in brown ink. The emotion effect is derived without appealing the concept of congruity: The difference in colour naming performance between the list of negative words and the list of neutral words defines the emotion effect. Do these differences between the two effects tap corresponding differences in the underlying cognitive processes?

Experimental and quasi-experimental designs with classic and emotional Stroop effects

In studies of the classic Stroop effect, the word RED sometimes appears in red ink (matching colour) and sometimes appears in green ink (conflicting colour). The difference in reaction time (RT) between the two combinations gives the *item-specific* Stroop effect for the word RED (Jacoby, Lindsay, & Hessels, 2003). The same calculation applies to *all* the other words in the set, and the classic Stroop effect is actually the sum total of the item-specific effects in the stimulus ensemble. The fact that the *same* words appear in congruent and incongruent combinations is of great importance in terms of experimental control. Differences in colour naming between the two combinations (the Stroop effect) cannot be attributed to differences between the words simply because they are the same words (Larsen, Mercer, & Balota, 2006). Consequently, the Stroop effect is measured within the framework of an experimental design with a singularly potent control.

In the emotional Stroop effect, by contrast, the critical difference in colour naming is calculated between *different* words. Suppose that a person is afraid of cockroach but not of dogs. When testing this person in the emotional Stroop task, the word COCKROACH may appear in the list of emotion items, and the word DOG in the list of neutral items. The objects of fear are not interchangeable for an individual, wherefore different items appear on the emotion and neutral lists. The words on the two lists are never the same. They cannot be, otherwise the task and the associated diagnosis is compromised.

The last point is important, posing a difficult if not intractable problem for experimental control. The observed slowdown can be attributed to emotional content only if the words in the emotional and neutral lists are matched on all other variables. However, it is impossible in principle to match the two lists of words on all conceivable variables. Inevitably, the measurement of the emotion effect is accomplished within a quasi-experimental design (Larsen et al., 2006). To illustrate the threat to validity, Larsen et al. (2006) found systematic differences between the emotion and neutral words in length, frequency, and orthographic neighbourhood in a sample of more than one thousand words culled from 32 published studies. Larsen and his colleagues thus identified one class of confounding variables, lexical features. However, even if the two lists are equal on lexical features, the threats to the validity of the emotional Stroop effect still exist. It is impossible to exhaust all extraneous variables that can account for the effect apart from emotional content. Only an experimental design is capable of accomplishing this feat. We developed such a design in this study.

A true experimental design for the emotional Stroop effect

We used the *same* words in the emotion and neutral lists. Our tactic was to polarise a given word to be negative or neutral by placing it in idiomatic expressions of different valence. The target word always appeared last in the two-word expressions used, so that the first word rendered it negative or positive (once the idiomatic expression was apprehended in full). Consider the target word, CHAIR. On the control list, this word was preceded by the word, ARM, to form the phrase, ARM CHAIR. On the emotion list, the same word was preceded by the word, WHEEL, to form the expression, WHEEL CHAIR. The participant was requested to name the ink colour of CHAIR and ignore the prime (which appeared in uniform black). Note that the prime words, ARM or WHEEL, are not negative (or positive) themselves (neither, incidentally, is CHAIR), so that the participant was unable to use the primes as cues. Indeed, the to-be-responded word, as well as the entire expression, assumed valence only upon the presentation of the target word.

A feature of the Hebrew language in which the study was conducted greatly enhanced the power of our procedure. In Hebrew, unlike in English, the adjective *follows* the noun (in Hebrew one says "man good" for "good man" in English). Notably, in our study we deviated from the standard Hebrew usage and presented the adjective before presenting the noun (illustrating the procedure in English, the observer was exposed to a sequence like CHAIR, WHEEL rather than to a sequence with the standard word order). Consequently, only upon the presentation of the entire sequence did the presence of an idiomatic expression become apparent, rendering the shared noun neutral or negative. Our procedure thus emulated common priming in which the consecutive words, NURSE, DOCTOR, do not coalesce into a single perceptual whole.

In sum, because the same words appeared in the emotion and neutral lists, all lexical and other possible confounds are eliminated. Any difference in colour naming can now be safely attributed to the (momentary) valence of the target word. For example, the word CHAIR was a negative item when preceded by WHEEL, but was a neutral item when preceded by the word ARM. We could also calculate item-specific emotional Stroop effects for the first time. The overall emotional Stroop effect, like the classic Stroop effect, is the sum total of the item-specific effects. Therefore, this study entails the demonstration of a pure emotion-induced slowdown in performance.

EXPERIMENT 1

Method

Participants. The participants were 24 young men and women, volunteers from the Tel-Aviv University community. There was an equal number of males and females. The age of the participants ranged between 20 and 29 years. All were native speakers of Hebrew, and all reported normal or corrected-to-normal vision.

Materials. The materials were drawn from the rich repertoire of idioms in the Hebrew language to be found in commonly used scientific dictionaries (e.g., *The Concise Sapphire Dictionary*, 1997; *The New Even-Shoshan Dictionary*, 1997). An initial sample of 72 two-word idioms was selected. The list consisted of pairs of idioms with a shared noun. The complementary words created a negative expression in one case and a neutral one in the other case. For example, the shared noun, SHEEP, was completed into the phrase, BLACK SHEEP (negative), in one case and into the phrase, WHITE SHEEP (neutral), in the other case. We selected the phrases such that *all* of the component words were neutral (i.e., neither particularly negative or positive) when viewed in separation. Only when presented jointly did the phrase don a distinctly negative or neutral-to-positive meaning.

Several independent groups of participants were engaged to test various lexical properties of the sample of the idioms. Thus, a group of 42 graduate students from Tel-Aviv University (26–38 years of age) were approached to rate individually the idioms for familiarity on a 1 (*not familiar, rare*) to 4 (*very familiar, common*) scale. This screening process resulted in the deletion of approximately half of the original items. The sample thus was reduced to 46 equally familiar items of 23 negative and 23 neutral phrases with the target word shared across pairs of negative and neutral expressions. Another group of 18 graduate students from Tel-Aviv University (25–33 years of age) evaluated the phrases on a 1 (*positive*) to 7 (*negative*) scale. The means were 5.3 and 2.7, respectively, for the negative and neutral-to-positive phrases, $t(17) = 5.11$, $p < .001$. For the two idioms sharing a given target word, 20 of the 23 pairs differed in the planned direction, for these phrases, $t(17) \geq 1.94$, $p \leq .06$. The three deviant pairs were deleted such that the sample was further reduced to 20 negative and 20 neutral-to-positive phrases. Notably, when the component words were presented singly, we did not find reliable differences in valence across various subsets of the words (targets vs. primes, primes in negative phrases vs. primes in neutral phrases; $t < 1$, in all comparisons).

Finally, we wished to rule out the possibility that the two prime words associated with a given noun activated the meaning of that noun to a different extent. A group of 20 students from Oranim College (19–26 years of age) produced two-word expressions to each of the *target* words presented singly. We then selected those expressions that included the original prime words and found that none of the pair of expressions produced for a given target word (negative, neutral) differed in frequency of selection. In a complementary experiment, we presented the *prime* words singly and asked a group of 26 senior high-school students from the greater Tel-Aviv area to produce to each as many verbal associates "as come to mind". We compared then the frequency with which each of the original target word appeared as a response to a given prime. None of the pair-wise differences including the original target words was reliable.

Apparatus and design. The stimuli were displayed on a 21-inch colour monitor set to a resolution of 1024×768 pixels. Using the standard palettes, we created the prototypical colours of red, blue, green, and brown. The prime word was presented in black over the light grey background of the screen 30 pixels below or above its centre. The target word was presented in one of the colours, red, blue, green, or brown. To avoid adaptation or strategic responding to the target word (e.g., fixating on a small portion of the print to avoid reading), we introduced a trial-to-trial spatial uncertainty of 50 pixels in appearance around the centre location. The words were presented in Ramat-Gan, 48 font. Viewed from a distance of approximately

60 cm, the words subtended a maximum of 5.4 degrees of visual angle in width and 1.52 degrees in height. An Apple external microphone, fitted to each participant and held at a constant distance from the mouth, collected the vocal responses.

Each trial entailed the following sequence. A fixation cross appeared in the centre of the screen for 500 ms, followed by a blank interval of 200 ms. A prime word in black was then presented slightly above or below the centre of the screen for 200 ms. After another blank interval of 200 ms, the target word printed in colour appeared at approximately the centre of the screen and remained visible until the participant's vocal response.

The 46 stimuli were presented in a random order subject to the proviso that phrases with the same target word did not appear consecutively. The four print colours were used with equal frequency to create the target words.

Procedure. The participants were tested individually in a dimly lit room. They were instructed to name the colour of the second word (the first word was always black) by speaking into the microphone as fast and accurately as possible. Stimuli were terminated by the participant's response. The interval between the participant's response and the appearance of the subsequent stimulus (i.e., the next fixation cross) was 600 ms. Trials with response times slower than 2 s or faster than 180 ms were repeated later in the task; 2.4% of the trials were repeated. The error rates were very low and did not exceed 2.2% with all of the individual participants. As a result, we do not discuss accuracy further in this article.

Results

The mean response time (RT) for correct responses was 679.5 ms for target words embedded in negative phrases and 660.9 ms for the same target words embedded in neutral-to-positive phrases. The difference of 18.5 ms amounted to a reliable emotional Stroop effect, $t(23) = 2.18$, $d = 0.2$, $p < .05$. Apart from the omnibus effect, we also calculated the item-specific effects for each of the 20 target words. We found a reliable emotional Stroop effect for 16 of these words—for the 16 items, all pair-wise tests had values larger (t, d) or smaller (p) than $t(23) \geq 1.92$, $d \geq 0.17$, $p \leq .08$. For each of these words, the ink colour was named slower when the word completed a negative expression than when the word completed a neutral expression.

Discussion

Because colour naming was performed on identical words, any lexical influences are conclusively ruled out. The observed difference in performance can be attributed solely to the momentary valence of the words, emotional or neutral. This study was also the first to derive item-specific emotional Stroop effects. Given the importance of the current results, we attempted a full-scale replication on a fresh group of participants in Experiment 2.

EXPERIMENT 2

Method

Participants. The participants were 38 volunteers from the Tel-Aviv University community (24 women and 14 men) with age ranging between 17 and 41 years. Again, all were native speakers of Hebrew, and all reported normal or corrected-to-normal vision.

Materials. The same two-word phrases from Experiment 1 were used, but we deleted the 4 pairs of phrases that failed to show a reliable item-specific effect. Consequently, the list of items included 32 phrases, 16 negative and 16 neutral sharing 16 common target words.

Design and procedure. All aspects of the design and procedure were the same as those in Experiment 1.

Results

The responses took appreciably longer than those in Experiment 1 (probably due to the extended range of age), but the same pattern reappeared. The mean (correct) RTs for the target words were 822.2 and 776.2 ms, respectively, for negative and neutral phrases, revealing an appreciable emotion effect of 46 ms, $t(37) = 5.62$, $d = 0.29$, $p < .05$. For the individual target words, each of the 16 comparisons between negative and neutral-to-positive expressions with the same item yielded a dependable emotional Stroop effect, $t(37) \geq 1.99$, $d \geq 0.16$, $p \leq .06$. The results of an overall ANOVA yielded a highly dependable effect of valence, $F(1, 36) = 32.61$, $\eta^2 = .22$, $p < .0001$, but also a smaller effect of target word, $F(15, 540) = 3.33$, $\eta^2 = .09$, $p < .01$. All of the effects were in the predicted direction (i.e., longer RTs for emotion than for neutral items). In particular, the effect found for the individual target words justifies the quest for item-specific emotional Stroop effects.

Discussion

The results of this experiment reinforced the evidence for the existence of an emotion- or threat-induced slowdown in performance. The results show that it is emotional content and not any other factor that produces the observed slowdown. Because the present method is new, we wished to apply a further experimental control to ascertain that the different primes are processed at roughly the same speed. We thus engaged the current participants in another condition.

EXPERIMENT 3

The different primes used in the negative and positive phrases were not negative or (overly) positive themselves. Although we tested these words quite extensively for equivalence in valence and on a range of lexical features in advance (see the method section of Experiment 1), we still wished test them in a speeded task in the laboratory. The prime words, drawn from the emotion and the control phrases, were presented singly. In this experiment, all the words appeared in colour, and the participant's task was to name the ink colour. If there are systematic differences in speed of responding between the two sets of primes, then this difference alone can produce the results. The two sets of words were judged to be of roughly comparative valence (neutral) in our pilot studies, yet there is still a (remote) possibility that one set engenders slower responding than the other set. It is this possibility that we wished to be ruled out by the testing of Experiment 3.

Method

Participants. The same group of participants from Experiment 2 also performed in this experiment. In fact, a random half of the participants first performed in Experiment 2 with the other half first performing in Experiment 3.

Apparatus, design and procedure. In this experiment, the participant named the print colour of words presented singly for view. A fixation cross was presented for 500 ms, followed by a blank screen for 200 ms. One of the prime words from Experiment 2 was then presented in one of the colours, red, blue, green, or brown. The word remained visible until the vocal response by the participant. The inter-trial interval remained 600 ms.

There were 32 words, 16 borrowed from the negative phrases and 16 from the positive ones. The words were intermixed in a single block and presented in a random and different order for each participant.

Data analysis. We subjected the data to a test aimed at detecting effects of order. We did not find any, $t(37) = 0.44$, $d = 0.07$, $p > .1$. Therefore, the results of Experiment 3 cannot be attributed to effects of habituation or learning in Experiment 2.

Results

We did not find a significant difference between the prime words that served (in a separate presentation) to create the negative and the positive expressions when conjoined with common target words. The mean RT for the words in the negative phrases was 775.1 ms, whereas the mean for the words in the neutral-to-positive phrases was 784. 8 ms, $t(37) = 0.35$, $d = 0.11$, $p > .1$.

Discussion

The absence of a systematic difference between the prime words used in the negative and control expressions further reinforces the validity of our manipulation. All the prime words used were neutral, free of emotional content. In a commensurate fashion, they produced comparable naming times when presented singly for view. Members of this uniform set of stimuli created pairs of expressions with distinct valence *only* when conjoined by shared target words.

GENERAL DISCUSSION

Full equivalence between the emotion and control words used in the emotional Stroop task is important from a scientific point of view. Mismatched lists of emotion and control items can create spuriously large effects. A difference in performance between the two types of items can be due to true emotional effects or due to lexical (or other) differences between the lists of words. Typically, both components contribute to the observed slowdown with emotion items, engendering an overestimation of the size of the true emotion effect. In the extreme, it is conceivable that, with grossly mismatched words, the lexical imbalance itself generates the effect.

We believe that a true emotion effect is likely present in the great majority of existing studies. However, it is difficult to assess its size. The effect may be absent though in some studies (cf. MacKay et al., 2004). In the aforementioned analysis of 32 emotional Stroop studies by Larsen et al. (2006), a “striking pattern of non-equivalence” (p. 67) emerged in lexical features between the emotion and control items used. When the authors removed these lexical differences statistically, the emotional Stroop effect vanished (except for a group of fairly rare words). Such data jeopardise validity. Can the words on the two list be balanced to enable a valid measure of the emotional Stroop effect?

One strategy has been to create balanced lists by statistical means. Any observed difference between the two lists can serve as a covariate in a subsequent analysis. Thus, Wentura, Rothermund, and Bak (2000) applied multiple regression analyses to correct for differences between their lists before examining the effect of emotion. Another strategy has been to select items known in advance to be equal on important lexical characteristics. Balota and his colleagues (Balota et al., 2002) have assembled a database containing lexical characteristics and experimental measurements for a large corpus of English words (the English Lexicon Project; ELP). The investigator can enter the relevant lexical parameters and the ELP will generate words matched on those parameters. A third and probably the most popular strategy has been to test the to-be-presented words for equivalence in dedicated pilot studies.

However, none of these strategies is capable of creating completely balanced lists of words that are equivalent on all extraneous variables (lexical and non-lexical). All of the methods require prior identification of the to-be-matched variables. This requirement compromises full control. There are always variables that the experimenter is unaware of at the time of testing or that are not included in collections such as the ELP. Even after lexical (and other) matching, the presence of residual extraneous variance cannot be ruled out. One cannot match the items on all conceivable variables using any one of the traditional methods. It is at this juncture that the current procedure accomplishes the feat of establishing completely balanced lists of emotion and control items. The present procedure is probably one of only two methods capable of producing a radical control of *all* lexical and other extraneous variables (including those where the experimenter is unaware of their existence).

A close relative of the present method is one that imbues some of the words with negative emotional tone using classical conditioning techniques (Kelly & Forsyth, 2007; Richards & Blanchette, 2004). We acknowledge the power of conditioning methodologies to achieve the common goal of lexical control and removal of possible confounds. The studies by Richards and Blanchette (2004) and Kelly and Forsyth (2007) are particularly notable for their convincing demonstrations. We note, though, than many other attempts at conditioning with humans proved failures. Richards and Blanchette (2004) themselves succeeded in conditioning non-words but not words. Of more importance, *different* words are conditioned to be positive and negative and/or *different* participants are exposed to the words in studies of conditioning. Indeed, Kelly and Forsyth (2007) express the need to use other supporting techniques due to these features of conditioning. In the present method, by contrast, the same participants are exposed to the same words, thereby providing the most stringent control conceivable.

Our point is not to argue that investigators should invariably use the present method. Clinical considerations mandate the use of different words in the two lists. These cannot be fully matched, nor should they be. The emotion- or threat-words presented are often tailor-made to test the pathology in question. Despite the absence of exact match, the effects thus derived carry considerable diagnostic value. In addition, in many clinical applications it is the group by stimulus interaction that is of main interest so that exact match of the stimuli is less critical. Judicious application of the methods mentioned does provide for reasonable control.

Instead, our point is that the current method is uniquely suited to derive *pure* emotion effects. When researchers are interested in a rigorous estimate of a genuine emotional Stroop effect, they ought to use some variant of the present method. For this scientific purpose, the present method lacks alternatives. The method is user friendly. The selection of items is less daunting a task than it may seem at first glance. Language (e.g., English, Hebrew, German) is rich in idiomatic expressions of distinct undertones that share a (key) word. However, one should consider in experimental planning the constrains posed by the particular word order mandated by the particular language. With precautions taken, the present method is of quite large generalisability.

Having established a true experimental design for the emotional Stroop effect (the classic effect has always been tested by one), one is in a better position to ponder the nature of the relationship between the classic and emotional Stroop effects (cf. Algom et al., 2004; Chajut, Lev, & Algom, 2005). Accessing the meaning of the task-irrelevant words is the fundamental generative condition for both phenomena. However, their relationship should be carefully scrutinised beyond this point. The fundamental manipulation in studies of the emotional Stroop effect is the instantiation of threat or emotion. The fundamental manipulation in studies of the classic Stroop effect is the instalment of conflict in the stimulus. The emotional effect is stimulus bound (fears and attractions are not interchangeable), whereas the interchangeability of the stimulus attributes (words and their colour) is a hallmark of the classic effect. One should be cognizant of these features when importing cognitive paradigms and theories to the realm of emotion and personality. Beware of cognitive psychologists bearing gifts!

Manuscript received 25 March 2008
Revised manuscript received 9 September 2008
Manuscript accepted 17 September 2008
First published online 5 December 2008

REFERENCES

Algom, D., Chajut, E., & Lev, S. (2004). A rational look at the emotional Stroop phenomenon: A generic slowdown, not a Stroop effect. *Journal of Experimental Psychology: General*, *133*, 323–338.

Balota, D. A. Cortese, M. J. Hutchison, K. A. Neely, J. H. Nelson, D. Simpson, G. B. et al. (2002). *The English Lexicon Project: A web-based repository of descriptive and behavioral measures for 40,481 English words and non-words.* (Available at: http://elexicon.wustl.edu)

Chajut, E., Lev, S., & Algom, D. (2005). Vicissitudes of a misnomer: Reply to Dalgleish (2005). *Journal of Experimental Psychology: General*, *134*, 592–595.

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

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

Kelly, M. M., & Forsyth, J. P. (2007). Observational fear conditioning in the acquisition and extinction of attentional bias for threat: An experimental evaluation. *Emotion*, *7*, 324–335.

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.

Larsen, R. J., Mercer, K. A., & Balota, D. A. (2006). Lexical characteristics of words used in emotional Stroop experiments. *Emotion*, *6*, 62–72.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate recall tasks. *Memory & Cognition*, *32*, 474–488.

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

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. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: The role of fast and slow components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 382–392.

McNally, R. J., Kaspi, S. P., Riemann., B. C., & Zeitlin, S. B. (1990). Selective processing of threat cues in posttraumatic stress disorder. *Journal of Abnormal Psychology*, *99*, 389–402.

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

Richards, A., & Blanchette, I. (2004). Independent manipulation of emotion in an emotional Stroop task using classical conditioning. *Emotion*, *4*, 275–281.

Richards, A., & Millwood, B. (1989). Color-identification of differentially valenced words in anxiety. *Cognition and Emotion*, *3*, 171–176.

Riemann, B. C., & McNally, R. J. (1995). Cognitive processing of personally relevant information. *Cognition and Emotion*, *9*, 325–340.

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

The Concise Sapphire Dictionary. (1997). E. Avenyon (Ed.). Jerusalem: Hed Arzi/Eitav.

The New Even-Shoshan Dictionary. (1993). A. Even-Shoshan (Ed.). Jerusalem: Kiryat Sefer.

Watts, F. N., McKenna, F. P., Sharrock, R., & Treize, L. (1986). Color naming of phobia-related words. *British Journal of Psychology*, *77*, 97–108.

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

Whalen, P. J., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., et al. (1998). The emotional counting Stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry*, *12*, 1219–1228.

Williams, J. M. G., Matthews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin*, *120*, 2–24.

Wyble, B., Sharma, D., & Bowman, H. (2007). Strategic regulation of cognitive control by emotional salience. *Cognition and Emotion, 22*, 1019–1051.