Atten Percept Psychophys (2014) 76:81–97
DOI 10.3758/s13414-013-0545-7

Is the emotional Stroop task a special case of mood induction? Evidence from sustained effects of attention under emotion

Moshe Shay Ben-Haim · Yaniv Mama · Michal Icht · Daniel Algom

Published online: 17 September 2013

Abstract Sustained effects of emotion are well known in everyday experience. Surprisingly, such effects are seldom recorded in laboratory studies of the emotional Stroop task, in which participants name the color of emotion and neutral words. Color performance is more sluggish with emotion words than with neutral words, the *emotional Stroop effect* (ESE). The ESE is not sensitive to the order in which the two groups of words are presented, so the effect of exposure to emotion words does not extend to disrupting performance in a subsequent block with neutral words. We attribute this absence of a sustained effect to habituation engendered by excessive repetition of the experimental stimuli. In a series of four experiments, we showed that sustained effects do occur when habituation is removed, and we also showed that the massive exposure to negative stimuli within the ESE paradigm induces a commensurately negative mood. A novel perspective is offered, in which the ESE is considered a special case of mood induction.

Keywords Sustained effects · Emotion · Habituation · Mood induction · Emotional Stroop effect

The sustained effects of encounter with a negative, emotion-laden event are well known. Watching a depressing movie can ruin the rest of your day. An insult sustained in the workplace can loom in one's mind for months. Negative experiences, along with their long-term effects, are hardly avoidable in everyday life, as anyone visiting the emergency room or witnessing a traffic accident can readily attest. In the laboratory, too, sustained effects of exposure to emotion stimuli can be quite robust (e.g., Bradley, Cuthbert, & Lang, 1996; Pereira et al., 2006), inducing negative mood (Gomez, Zimmermann, Guttormsen-Schar, & Danuser, 2009; Smith, Bradley, & Lang, 2005). In view of the evidence for the presence of long-term effects of exposure to negative stimuli and of the negative mood thus induced, it is quite surprising that the most popular laboratory paradigm, the emotional Stroop task, lacks data and reference to sustained effects. Theories of the emotional Stroop effect (ESE) similarly lack consideration of sustained effects. Consequently, the goal of the present study was to probe the long-term effects of exposure to negative stimuli included within the ESE framework. Our conclusions implicated habituation as the source of the absence of order—or sustained–effects in many ESE studies. In the present study, we specified the conditions under which such effects do and do not obtain. Finally, we suggest a novel perspective for the ESE, by considering the phenomenon as a special case of mood induction.

The emotional Stroop task

The experimental setup of the emotional Stroop task is well known. Words in color are presented singly for view, and the participants' task is to name the ink color of each word as quickly and accurately as possible. Typically, it takes longer to name the ink color of emotion or negative words than that of neutral words, the ESE. The ESE has been employed in a gamut of pathologies (see Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & van IJzendoorn, 2007; Phaf & Kan, 2007; and J. M. G. Williams, Mathews, & MacLeod,

M. S. Ben-Haim (✉) · Y. Mama · M. Icht · D. Algom
School of Psychological Sciences, Tel-Aviv University,
Ramat-Aviv 69978, Israel
e-mail: Shay.mbh@gmail.com

Y. Mama
Department of Behavioral Sciences, Ariel University, Ariel, Israel

M. Icht
Department of Communication Disorders, School of Health,
Ariel University, Ariel, Israel

1996, for reviews). The ESE has also been studied in unselected (i.e., nonclinical) populations (e.g., Algom, Chajut, & Lev, 2004; Algom, Zakay, Monar, & Chajut, 2009; Ben-David, Chajut, & Algom, 2012; Chajut, Mama, Levy, & Algom, 2010; Chajut, Schupak, & Algom, 2010; Mama, Ben-Haim, & Algom, 2013; McKenna & Sharma, 1995, 2004). The popularity of the ESE is likely attributable to its promise as a diagnostic tool. The emotion words can be tailor-cut to match the specific pathology under test. Moreover, this tool is not based on self-report.

Its popularity granted, a worrisome feature of the ESE is the large variability witnessed across studies. In the review of 50 studies by Williams et al. (1996), effect size ranged between –1 and 400 ms. One source is the pathology tested (e.g., posttraumatic stress disorder patients produced the largest ESE, whereas panic disorder patients yielded the smallest effect). However, considerable variability remains when testing groups with the same or with similar pathologies. The method of presentation is implicated as a major source of variation (blocked, randomly mixed, pseudorandomly mixed, deliberately biased; see Frings, Englert, Wentura, & Bermeitinger, 2010; McKenna & Sharma, 2004; Phaf & Kan, 2007), In this study, we suggest a further methodological feature associated with long-term effects as a source of variability. Studies differed in the order, extent of repetition, and number of exposure to negative stimuli. These variables affected the habituation of responses to the negative stimuli, which, in turn influenced the observed ESE. Accounting for these effects can further reduce the heterogeneity of ESE results. As we shall also see, a further bonus is a ready explanation for (the absence of) long-term effects.

Two competing classes of accounts of the ESE are those of attention and of threat or freezing, which further divides into automatic and non-automatic response to threat. According to the attention account (e.g., J. M. G. Williams et al., 1996), the carrier word commands attention, thereby compromising full focusing on the target ink color. Attention to words laden with emotion is especially acute and the toll taken by this extra processing is expressed in longer color-naming latencies. According to the threat account (Algom et al., 2004; Fox, Russo, Bowles, & Dutton, 2001; McKenna & Sharma, 2004), the slowdown is related to exposure to threatening stimuli. This exposure mandates momentary reprioritizing of resources at the expense of performing all ongoing activity (Öhman, Flykt, & Esteves, 2001). Neither account gives a complete explanation to all known facets of the emotional Stroop phenomenon. On a first glance, the finding that the ESE is larger in high-anxious individuals than in low-anxious individuals is consistent with the threat account. However, Blanchette and Richards (2013) have recently reported that the main difference between the two samples is rooted in the amount and type of *attention* directed to negative stimuli and, in general, in the cognitive processing of threat. The finding that the ESE diminishes in subliminal conditions seems to favor the attention account and it is certainly inconsistent with an automatic threat account, but not with a non-automatic threat account. The attention account, by contrast, can be challenged by the trivial fact that ESE is larger with negative than with positive words. In sum, although attention and threat theories account for a fair number of major ESE findings, the examples cited also show that all issues are not yet resolved. Indeed, the ESE might ultimately derive from multiple sources (Frings et al., 2010). In this study, we wish to highlight the fact that neither the attention or the threat theories nor multiple-source accounts of the ESE concern long-term effects of emotion.

The absence of sustained effects: The role of within-block habituation

Within the framework of the ESE, the presence of sustained effects would mean that neutral stimuli that *follow* emotional ones yield longer response latencies than the same neutral items preceding the emotional ones. With sustained effects in force, the ESE is limited to one order of presentation, namely, that in which the block with neutral items comes first. Startlingly, existing ESE research does not evince an order-of-block effect. Investigators have typically balanced across orders of presentation and did not consider the effect further in their analyses (e.g., Becker, Rinck, Margraf, & Roth, 2001; Holle, Neely, & Heimberg, 1997; Richards, French, Johnson, Naparstek, & Williams, 1992; Segerstrom, 2001). Those who did (McKenna & Sharma, 1995) failed to discover an effect of order (but see McKenna, 1986; see also McKenna & Sharma, 2004, for an analysis of short term carryover effects). Apart from these few exceptions, order of blocks has been neglected with impunity across large portions of ESE research. Whether or not the participants experienced the block with emotion words first did not make a discernible difference. Why?

Our research implicates habituation to the repeated appearance of the negative words as the source for the absence of a long-term effect. In the typical ESE experiment, four words are presented each in four ink colors with each color-word combination appearing three times. Note that in this most common design, a given word repeats 12 times in a block of 48 trials. This number of repetitions cripples the power of the emotion words to interfere with the response to the color. By the end of the block, the responses to the ink color of emotion words become fast—as fast indeed as those to the color of neutral items. Consequently, the responses in a subsequent block with neutral items are free of any sustained effect of emotion. The upshot is that the absence of an order-of-block effect is rooted in within-block habituation. Because item repetition has been a fixture of ESE research, the habituation thus generated disabled sustained effects of the encounter with the emotion items. The absence of sustained effect, in turn, means that the temporal order of the block with emotion items (first, second) does not affect the resulting ESE.

The problem of habituation has been recognized in the ESE research. McNally, Riemann, and Kim (1990), as well as MacKay et al. (2004) found the ESE to be larger in the first 100 trials than in the last 100 trials. In a similar vein, Witthöft, Rist, and Bailer (2008) found the ESE larger in the first 20 trials than in the last 20 trials. In a dedicated study, McKenna and Sharma (1995) found an overall slowdown over successive blocks of trials, probably due to flagging of attention and fatigue. Notably, the ESE was only present in the first few blocks (of emotion and of neutral items), underscoring the effect of habituation. When the authors replaced the words in each successive block with new ones, thereby reducing habituation, the ESE reappeared.

This much granted, habituation has been tested to date at a relatively coarse level of granularity. The smallest unit examined was that of a block (however small) or a rough-cut across the first versus the last few trials of an experiment. Missing are data at the level of the *individual* words. What is the fate of the response to an individual word over its successive presentations in the course of the experiment? How early does habituation occur? How many repetitions are needed produce an appreciable effect of habituation? Are there ways to avoid or bypass habituation? These questions were addressed in the present study.

The ESE: A case of mood induction?

In the present analysis, the absence of sustained effects derives from excessive repetition of the same set of negative items. An immediate corollary is that sustained effects *do* emerge when within-block repetition is removed. Eschewing repetition, the negative items become immune to habituation and are bound to exert considerable effect on subsequent responding. The magnitude of the order-of-block effect is thus proportional to within-block repetition of items. In this study we show that sustained effects do appear within the ESE framework, too, once repetition of items is avoided. If so, how does one explains their presence, especially as existing accounts of the ESE are silent on sustained effects?

We suggest a novel perspective, which can accommodate long-term effects. Our point of departure is the sheer fact of exposure to dozens of negative words during the course of a given ESE experiment. This procedure is bound to induce a commensurately negative mood in the observer. A voluminous literature on mood induction (for reviews, see Clark, 1983, and Gerrardshesse, Spies, & Hesse, 1994) reveals a depression-like slowdown in responding following induced negative mood in healthy individuals. The slowdown affects the gamut of behaviors from lever pulling (Gouaux & Gouaux, 1971) to word association (Coleman, 1975; Matheny & Blue, 1977) to speech (Banse & Scherer, 1996; Barrett & Paus, 2002; McKenna & Lewis, 1994; Sobin & Alpert, 1999) to writing (Natale, 1977, 1978) or counting (Teasdale, Taylor, & Fogarty, 1980). In fact, reading negative statements suffices to engender the negative mood along with its deleterious consequences (Velten, 1968) as does even exposure to individual words (Higgins, Rholes, & Jones, 1977; see also Gilboa-Schechtman, Revelle, & Gotlib, 2000, on the effect of congruent and incongruent mood on performance).

Therefore, we suggest that the ESE is a special case of mood induction. Our hypothesis is not incontestable. First, single-item presentations might prove a weak vehicle for inducing a specific mood state. Second, in many (though by no means all, see Parrott & Sabini, 1990, for natural mood induction) procedures, the participants are asked to experience the required mood. In the ESE paradigm, by contrast, the participants are asked to process a nonemotion feature. In the following empirical section, we provide evidence that neither concern is substantive and show that the ESE is a genuine case of mood induction.

The present study

In Experiment 1, we report the global or mean ESE—the statistics routinely reported—but also the novel statistics of the momentary ESE, which changes as a function of item repetition. We show that the mean ESE actually derives from the very first repetitions of an item. In Experiment 2 we demonstrate the presence of sustained effects by creating a block of trials in which each item is presented only once. This tactic removes the possibility of habituation. We test sustained effects further in Experiment 3, applying a within-subjects design. In our final experiment, we test directly the presence of negative mood following performance in blocks of emotion words.

Experiment 1

The goal of Experiment 1 was to generate the ESE within the framework of the standard paradigm. A small set of words was repeatedly presented in the block with emotion items and another small set of words was repeatedly presented in the control block with neutral items. A unique feature of this experiment was the measurement of reaction time (RT) to each individual word over its repeated presentation in the block. In this way we followed the evolution in time of the ESE. Because both habituation curves (one for emotion words and another for neutral words) approximate exponential functions that asymptote at the same minimum (Rankin et al., 2009; Thompson, 2009), the ESE is expected to be large in the beginning (due to the slower, habituation-free responses to the emotion words), but to decrease with repetitions. Because repetition takes its full toll by the end of the block, we did not expect an order-of-blocks effect.

Method

Participants

The participants were 40 Tel Aviv University undergraduates, 34 women and six men, with a mean of 26 years of age. The participants performed against course credit. All had normal or corrected-to-normal vision and all were native speakers of Hebrew. Each participant was assigned in a random fashion into one of two groups defined by the order of exposure to the emotion and the neutral items. Half of the participants performed first in the neutral block, and the remaining half first performed in the emotion block. Regardless of order, the mean RT difference in color naming between the two blocks defined the ESE.

Stimuli and design

We used five negative emotion words—the Hebrew equivalents of *injured*, *terrorist*, *suicide bomber*, *danger*, and *terrorist act* (the words were borrowed from the validated list of Ben-David et al., 2012; see also Ben-David, Calderon, & Algom, 2005). The five neutral words were *street*, *table*, *bench*, *cupboard* and *neighborhood*. All the words were matched for length and for average frequency using the Word Frequency Database for Printed Hebrew (Frost & Plaut, 2001).

The five emotion words and the five neutral words were presented in separate blocks. Within a block, each word appeared ten times (twice in each of the ink colors, red, brown, orange, blue, and green), making for 50 trials per block in all. The participants reported the ink color orally by speaking the name of the color into a microphone. They were given five practice trials with the Hebrew equivalent of the word *example*. The emotion and the neutral blocks were separated by a break of 30.

Apparatus

The stimuli were generated in Microsoft Word (in Hebrew font Ariel, 28-point size) via a Pentium 4, 2.8-GH computer and displayed on a 15-in. color monitor set at a resolution of 1,024 × 768 pixels. Using the standard palettes, we created prototypical colors for red (255R, 0G, 0B), brown (128R, 64G, 0B), orange (255R, 128G, 0B), blue (0R, 0G, 255B), and green (0R, 128G, 0B). The words appeared in color over the white background of the screen within an invisible frame of a 61 × 19 pixels rectangle. Viewed from a distance of approximately 60 cm, the words subtended 1.53º of visual angle in width and 0.48º of visual angle in height. In order 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 approximately 50 pixels around the center location. A microphone headset (Teac HPX-8 brand) served to transduce the oral responses.

Procedure

The participants were tested individually in a dimly lit room. They were instructed to report the ink color of the word as quickly and accurately as possible by speaking its name into the microphone headset. The computer software recorded the responses and their timing. Stimulus exposure was response terminated. The interval between the participant's response and the appearance of the next stimulus was 500 ms. Following the experiment, the participants were debriefed and thanked for their participation.

Data analysis

We exercised extra caution to ascertain that any systematic change (speed increase) in responding over trials was the result of habituation. In order to rule out an effect of (recovery from) surprise, five practice trials were run with each participant. Moreover, the very first trial of each participant was removed from the analyses due its long latency (Tukey's post hoc analysis revealed that the first trial was slower than all subsequent 49 trials, $p < .015$). We alert the reader to the fact that the fine-grained analyses below are based on *repetition*, and not on mere order of appearance.

The error rates and microphone/speech failures were low [3.9% and 4.6% for emotion and neutral words, respectively; $t(1, 39) = 1.14$, $p = .26$]; we do not discuss errors further on in this report. For RTs, responses longer than 2,200 ms or shorter than 250 ms were discarded from the analysis.

Results

Figure 1 gives the results with respect to the global ESE. Plotted are the mean RTs for naming ink color as a function of the valence of the carrier words and order of presentation. The left-hand half of Fig. 1 gives the results for the group performing first with the block with neutral items (N) and performing then with the block with emotional items (E; order NE). The right-hand half of Fig. 1 gives the comparable data for the group performing in the reverse order (EN). These data were analyzed in a repeated measure analysis of variance (ANOVA) with block valence (E, N) as a within-subjects variable and group order (EN, NE) as a between-subjects variable. Clearly, an appreciable ESE is discernible in both groups: Naming the ink color of words was always slower in the emotion block than in the neutral block. In the NE group, the mean RT for the emotion items was 797 ms, whereas for the neutral items it was 776 ms. The respective means in the

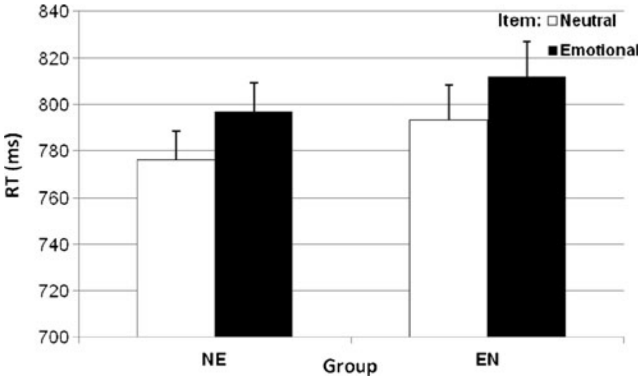

Fig. 1 Mean reaction times (RTs) to name the ink color of emotion (black columns) and neutral (white columns) words, presented in separate blocks of trials in different orders. In the NE group, the block of neutral words preceded the block of emotion words, and in the EN group, the block of emotion words preceded the neutral words. The error bars depict one standard error around the mean

EN group were 812 and 793 ms. The overall ESE amounted to 20 ms [$F(1, 38) = 4.1$, $p = .049$]. The lack of an Order × Valence interaction ($F < 1$) confirmed the presence of the same pattern of performance in the two groups.

The presence of the ESE is salient to visual inspection. However, the most revealing feature of the data is the absence of an order-of-blocks effect. The participants responded to neutral words faster than they did to emotion words *regardless* of whether the neutral words preceded or followed exposure to emotion words. In particular, we did not record a slowdown in performance to neutral words that were presented after exposure to the emotion words. How can one explain this lack of an order-of-block effect?

In order to gain insight, we recorded color-naming performance for repeated presentations of items within each block for each group. The results of this fine-grain analysis are presented in Fig. 2. The two panels present the mean RTs per repetition in the two groups for the emotion (squares) and neutral (diamonds) blocks of items.

The repetition functions exhibit largely the same pattern: Color-naming latencies fall precipitously beyond the first repetition. The effect of repetition is confirmed by the results of separate ANOVAs performed in each of the four conditions depicted in Fig. 2 [$F(9, 171) \geq 1.92$, $p \leq .05$, for both sets of data in the left-hand panel, and for the emotion set in the right-hand panel; repetition did not produce a statistically reliable effect in the neutral set in the right-hand panel].

For a more powerful analysis on the temporal locus of the ESE, we considered (a) the two emotion blocks together as a single set of data, and similarly (b) the two neutral blocks together as another set. Habituation was very strong in the emotional set [$F(9, 351) = 7.52$, $p < .0001$]. Tukey's HSD test revealed that the main difference in color naming was between the first appearance and all subsequent repetitions ($p < .0001$). For the neutral blocks considered as a group, habituation was also present [$F(9, 351) = 1.98$, $p < .05$], but Tukey's HSD test failed to identify a specific locus ($p > .069$). Nevertheless, in order to gain a clue as to the source of the effect with neutral items, we applied Fisher's LSD test (the lack of correction for multiple comparisons notwithstanding). This test revealed that the first appearance differed from Repetitions 1–8, which, in turn, differed from the last one ($p < .05$).

Consider Fig. 2 again. Closer inspection reveals that the drop beyond the first appearance of the words is more noticeable in the blocks with emotion items than in those with neutral items (especially as the starting latencies were longer in the former than in the latter). This initial difference is the major factor producing the global ESE. The ESEs, partitioned by repetition, are presented in Fig. 3. For each group of participants, the difference in performance between emotion and neutral items is presented as function of the repeated presentation of items. A glimpse at Fig. 3 reveals that the RT

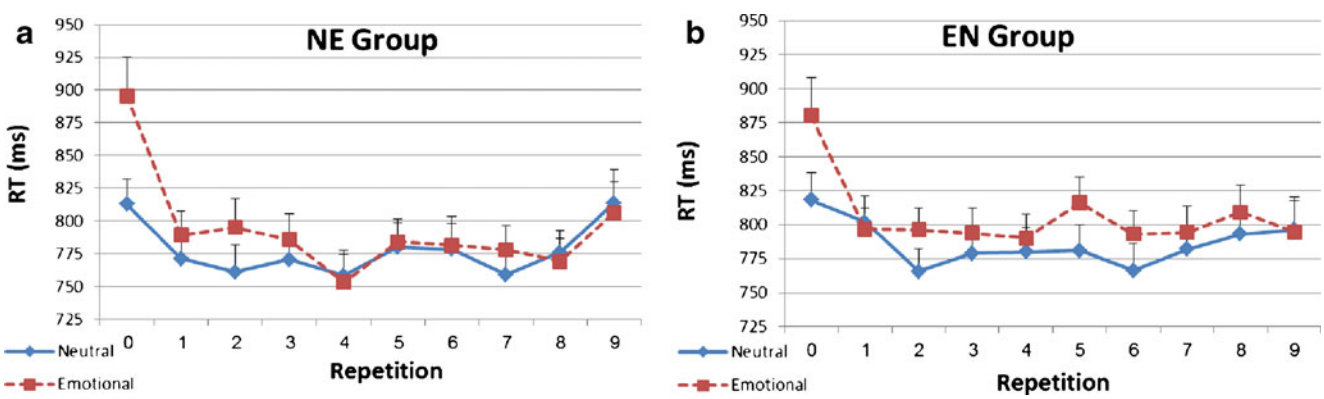

Fig. 2 Mean reaction times (RTs) plotted against word repetition. *Repetition* denotes the ordinal number of presentation of a word. It should be distinguished from the trial number of the experiment: If a given word first appeared in the fourth trial, its repetition number would then be 0. The ordinate thus gives the mean RTs for color naming as a function of the number of repetitions of the words in the experimental block. The left-hand panel presents the data for the participants in the NE group, whereas the right-hand panel presents the data for the participants in the EN group. In each panel, the squares denote the data for emotion words, and the diamonds the data for neutral words. The bars denote one standard error around the momentary mean

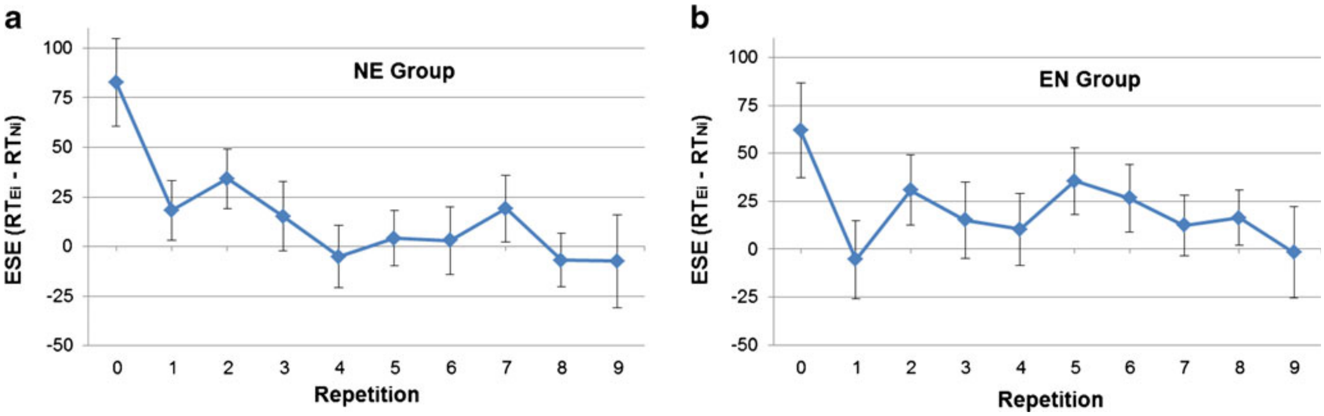

Fig. 3 Momentary emotional Stroop effect (ESE) plotted against word repetition. For each repetition, the momentary ESE is the difference in color-naming latencies between emotion and neutral items for that repetition. Presented in the left panel are the data for the participants in the NE group; presented in the right panel are the comparable data for the participants in the EN group. The bars denote one standard error around the mean

difference—the *momentary ESE*—diminishes after the first few repetitions.

All of the previous observations were confirmed in a three-way ANOVA with word valence (emotional, neutral) and repetition (0–9) as within-subjects variables and block order (NE, EN) as a between-subjects variable. The marginally significant interaction of word valence and repetition [$F(9, 342) = 1.8$, $p = .066$] confirmed that the difference between emotion and neutral items did not remain constant across repetitions. Tukey's post hoc analysis showed that the locus of the interaction resided in the difference between the first appearance of the emotion items and all subsequent repetitions across emotion and neutral items, $p < .0001$.

Discussion

We used the standard procedures of the ESE experiment. In each block, a small set of carrier words was presented repeatedly for color naming. These procedures yielded the standard ESE result. We recorded a slowdown in naming the color of emotion or threat items compared with that of neutral items. The slowdown amounted to an overall ESE of 20 ms. Another customary result was the lack of an order effect. The slowdown with emotion items occurred regardless of whether they preceded or followed the neutral items. We argue that this null effect of order is a natural outcome of the standard procedures used in the great bulk of ESE research (including Exp. 1). The item-by-item analysis performed in this experiment bore out the validity of our argument.

One should realize that the ESE of 20 ms obtained in Experiment 1 is a global measure, based on *all* trials in both blocks. This global measure is what is routinely reported in ESE studies. What we do not know is the contribution of the individual trials to this measure. Are those contributions equal? Is the global ESE mainly attributable to the last couple of items? Or, is the ESE mainly arising during the first few presentations? The novel contribution of Experiment 1 therefore was the item-by-item dissection of the ESE. By calculating the momentary ESEs, we followed the temporal evolution of the global ESE.

Our analysis revealed that the global ESE—the statistics typically reported—is actually produced by the first few emotion and neutral items presented, if not on the very first ones. The momentary ESE amounted to 73 ms upon the first appearance of the emotion and the neutral items, but it diminished noticeably beyond those presentations. In fact, none of the momentary ESEs beyond that first one was reliable statistically. So, the dilution over trials of an initially large ESE is what produced the global value of 20 ms.

The fact that the ESE arises only from the first (few) trial(s) is striking and important. On a first glance, this finding is at odds with explanations of a deep seated attention bias to negative information with patients with high anxiety disorders. Note though that our participants were drawn from the general (student) population, and were not clinically selected. We also note that Amir, Najmi, and Morrison (2009) have reported apparent habituation of attention bias to threat words in a dot-probe experiment with subclinical OCD patients, so that "a comparison of early and late blocks of trials revealed an attenuation of attention bias . . . potentially reflecting habituation to threatening information over the course of the experiment" (p. 153).

Why does the ESE diminish over repetitions? We attribute the trend to habituation to the word stimuli over their repeated presentation. Responses to both types of items undergo habituation so that the difference between the respective curves diminishes with repetition. Isolating habituation provides the student of ESE with a further bonus. It explains the absence of an effect for order. Repeated presentation of the carrier words renders them less salient. For emotion words in particular, excessive repetition cripples their emotive power, leaving responses in a subsequent block with neutral items unaffected.

The upshot is that habituation, a within-block effect, eliminates the influence of the particular sequence of block presentations, a between-block effect.

If the absence of an order-of-block effect derives from habituation, the removal of habituation should restore the sustained effect of emotion often observed in everyday life. The natural (only?) way to get rid of habituation is to eliminate word repetition. Thus, the block of trials would still include 50 presentations, but none of the words appears more than once This tactic was used in Experiment 2. We predicted that, with habituation eliminated, order of blocks *does* make a difference. In particular, performance with neutral words that follow exposure to emotion words would be poorer than that with neutral words that do not follow such an exposure. Removal of habituation permits one to observe sustained effects of emotion within the ESE context.

Experiment 2

Method

Participants

The participants were 32 Tel Aviv University undergraduates none of whom took part in Experiment 1. Each participant was assigned into one of two groups differing in order of presentation. In the NE group, the participants first performed in the block with neutral items and performed then in the block with emotion items. For the participants in the EN condition, the blocks were presented in the reverse order.

Stimuli, apparatus and design

The apparatus and design were the same as in Experiment 1 except for the following notable exception. In Experiment 2, 50 different words were presented in the emotion block, and, similarly, 50 different words were presented in the neutral block. Because each block entailed 50 words, a word was presented only once (in one of the five possible ink colors). The words in the emotion and neutral blocks were matched for length and average frequency using the Word-Frequency Database For Printed Hebrew (Frost & Plaut, 2001). The emotion words were drawn from the list of 322 *negatively encoded* words used in 32 published emotion Stroop studies as complied by Larsen, Mercer, and Balota (2006). We will provide a list of these words on request.

Procedure

The procedure was the same as in Experiment 1. After random assignment into one of the order conditions, the participant proceeded with naming the ink color of words presented singly for view. As in Experiment 1, a break of at least 30 s separated the two blocks.

Data analysis

The same procedures used in Experiment 1 were applied.

Results

Figure 4 gives the results. Again, the data were analyzed in a repeated measure ANOVA with block valence (E, N) as a within-subjects variable and group order (EN, NE) as a between-subjects variable. Ignoring block order, responses to emotion words (787 ms, on average) took longer than did those to neutral words (763 ms, on average). This difference in color performance favoring neutral words amounted to an overall ESE of 24 ms [$F(1, 30) = 4.82$, $p = .036$]. However, the most revealing feature of the data depicted in Fig. 4 is the potent effect of order. The ESE appeared only in the group performing in the NE order. In this group, the ESE was 45 ms, whereas in the group performing in the reverse order (EN), the ESE amounted to a minuscule 2 ms. The interaction of block order and valence, $F(1, 30) = 4.1$, $p = .05$, confirmed the effect of order, confining the ESE into the sequence in which emotion items *follow* neutral items.

Another noteworthy feature of the data in Fig. 4 is the way that the ESE evaporated in the group that was first exposed to emotion items (EN). Clearly, this exposure made subsequent responses to *neutral* items sluggish. As is evident in Fig. 4, responses to neutral items that followed exposure to emotion ones were as slow as those given to the emotion items themselves (and much slower than the responses given to the same items without prior exposure to emotion stimuli

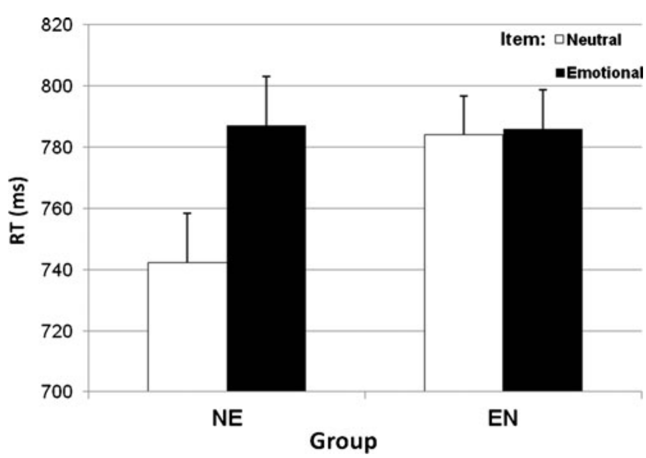

Fig. 4 Mean reaction times (RTs) to name the ink color of emotion (black columns) and neutral (white columns) words, presented in separate blocks in different orders. In the NE group, the block of neutral words preceded the block of emotion words, and in the EN group, the block of emotion words preceded the neutral words. The error bars depict one standard error around the mean

as in group NE). The parity eliminated the ESE in the EN condition, demonstrating the sustained effect of emotion in the laboratory, too.

The last point is notable. In Experiment 2, we uncovered a sustained effect of exposure to emotion stimuli, an effect absent in large portion of studies in the laboratory (although present in everyday life). We implicate habituation as the root cause of the difference. In Experiment 2, we avoided habituation by presenting each word only once. This tactic preserved the full force of the menacing content of each item throughout the block.

In order to test our account, we followed the evolution of responding within each block by measuring performance in successive subsets of five trials (the closest analogue to the unit repetition from Exp. 1; recall that in this experiment, there were not any repetitions of items). A glimpse at Fig. 5 shows the lack of habituation in all conditions ($p > .05$, Tukey HSD, for set of trials). In the two blocks in the NE group (left-hand panel), we observed a tendency for the responses to become *slower* rather than faster, as would be mandated by habituation. For the comparable blocks in the EN group (right-hand panel), the mean RTs remain fairly constant (in the face of considerable variation).

Statistical analysis revealed a main effect of valence (= ESE) in the NE group [$F(1, 15) = 4.75$, $p = .046$] but none in the EN group ($F < 1$). The momentary ESEs are plotted in Fig. 6. In the NE group, the ESE was mostly positive, whereas in the EN group, the ESE hovered around zero throughout the block. Neither trend was reliable statistically though.

The collective results of Experiments 1 and 2 implicate habituation to be the root cause for the absence of an order-of-blocks effect in large portions of ESE studies. Sensory adaptation and/or fatigue, unlike habituation, do not exhibit stimulus specificity (Rankin et al., 2009). In contrast, our results do document stimulus specificity. In Experiment 1, the responses were facilitated as a function of repeating the *same* carrier words. In Experiment 2, by contrast, *different* emotion-word carriers continued to disrupt color performance throughout the block. These features along with our methodological precautions conspire to implicate habituation as the sole determinant of the data.

Discussion

The hallmark of the present results is the demonstration in the laboratory of a sustained effect of emotion. Recurrent encounter with emotion or threat stimuli whose potency is preserved by warding off habituation affects responses to non-emotion stimuli presented after the encounter. As we recounted, encountering an emotional event in everyday life often carries long term consequences. We now have emulated this experience in the laboratory. Our participants reacted to neutral stimuli that *followed* emotion ones almost as sluggishly as they did to the emotion stimuli themselves. As a result, the ESE—the difference in color performance between emotion and neutral items—evaporated under those circumstances. An ESE was present only when the neutral stimuli *preceded* the presentation of emotion stimuli.

Our conclusion is threatened however by the following possible confound. In Experiment 1 (as well as in the great bulk of existing ESE studies), a small number of different emotion words (five) were presented (each word appearing many times). In Experiment 2, by contrast, a large number of different emotion words were presented (50; each word appearing once). Although the number of trials was the same in both experiments (50), their composition differed. Only five *different* emotion words were used in Experiment 1, whereas 50 *different* emotion words were used in Experiment 2. Therefore, an alternative explanation of the diverging results would implicate this difference in set size of different emotion items rather than habituation.

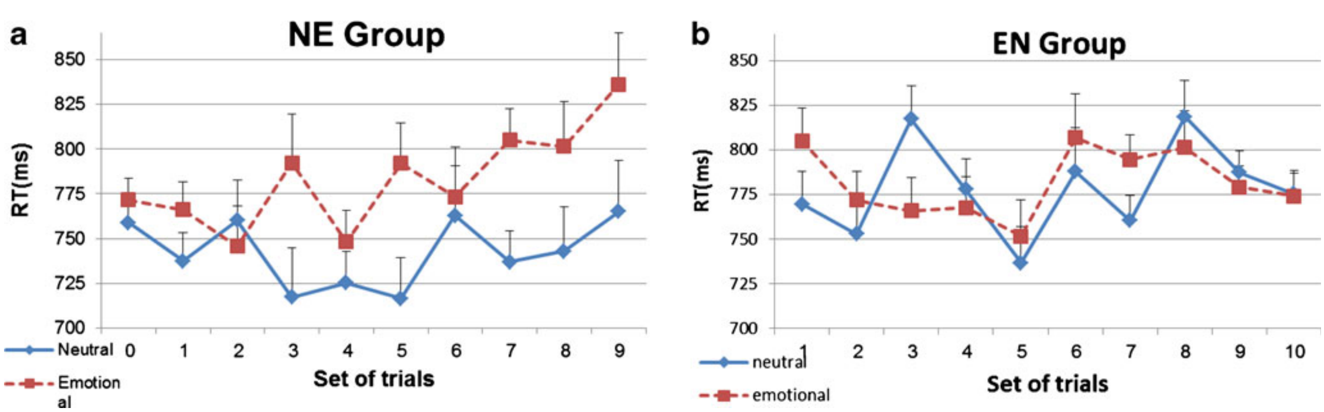

Fig. 5 Mean reaction times (RTs) plotted against consecutive subsets of five trials in two groups defined by different orders of presentation. In each panel/group, the squares denote data for emotion items, and the diamonds data for neutral items. The bars represent one standard error around each momentary mean

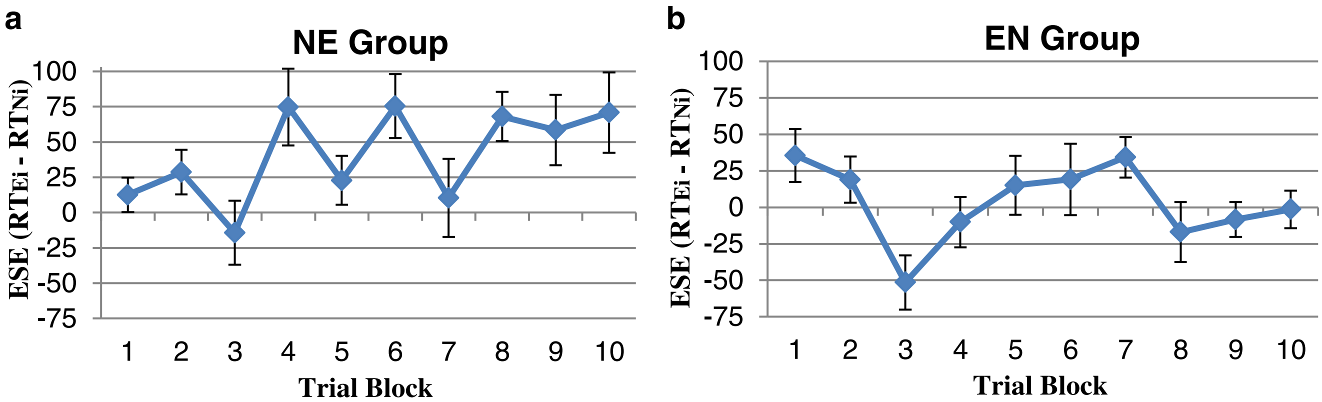

Fig. 6 Momentary emotional Stroop effect (ESE) plotted against blocks of five trials. The momentary ESE is the difference in color-naming latencies between emotion and neutral items for that trial position within the whole 50-trial block. Presented in the left panel are the data for the participants in the NE group; presented in the right panel are the comparable data for the participants in the EN group. The bars denote one standard error around the mean

Therefore, the goal of Experiment 3 was to test this threat to the validity of the habituation hypothesis. Two groups of participants responded in three blocks of trials. The first and the third block included neutral items, whereas the second block included emotion items. The critical block distinguishing the two groups was the second block. In the first group, this block composed entirely of emotion items. The same set of 20 emotion items were repeated three times, making for 60 trials in all. In the second group, the same set of emotion items was presented only once, followed by two repetitions of 20 neutral items (drawn from the first block). Sustained effects of exposure to the negative stimuli were tested by observing performance in the last block with neutral items.

The habituation account and the set size account entail different predictions. If set size or totality of exposure to negative stimuli is the decisive factor, one expects sustained effects in the first group to be stronger than in the second group. After all, the participants in the first group were exposed to 60 trials with negative stimuli compared with 20 such trials in the second group. Alternatively, if habituation is the decisive factor, the explanation that we endorse, one expects sustained effects to be larger in the second group than in the first group. The reason is that repeating the same items in the first group engendered habituation, which, in turn, should reduce sustained effects.

Experiment 3

Method

Participants

The participants were 32 Tel Aviv University undergraduates, 20 women, and 12 men, with a mean of 23 years of age. None of whom took part in the previous experiments.

Stimuli and apparatus

The apparatus and design were the same as in Experiments 1–2. The stimuli consisted of three different lists of words, each entailing 20 items. Two of the lists contained 20 neutral words each, and one list contained 20 emotional words. All the words were matched for length and for average frequency using the Word Frequency Database for Printed Hebrew (Frost & Plaut, 2001).

Design

Each participant performed in three consecutive blocks. The first block consisted of the 20 neutral words. The second block consisted of three repetitions of 20 emotional words (first group) or presentation of the same 20 emotion words followed by two repetitions of the 20 neutral words from the first block (second group). The final block included a new list of 20 neutral words. The words in each block were presented in one of five colors: red, blue, brown, orange, and green. Presentation within each block was different and random for each participant. The participants reported the ink color of each word orally by speaking its name into the microphone. They were given five practice trials. A short break of 30 s separated the experimental blocks.

Procedure

The participants were tested individually in a dimly lit room. Each participant was assigned to either the group entailing repetition of emotion items only or the group entailing repetition of neutral items that followed single presentation of emotion items. The participants were instructed to report the ink color of the word as quickly and accurately as possible by speaking its name into the microphone headset. The computer software recorded the responses and their timing. Stimulus exposure was response terminated. The interval between the participant's response and the appearance of the next stimulus was 500 ms. Following the experiment, the participants were debriefed and thanked for their participation.

Data analysis

The same procedure performed in the previous experiments was applied.

Results

Figure 7 provides the results. Consider the results of the "neutral repetition" group in the left-hand half of Fig. 7. In the critical second block, the participants were exposed to a set of 20 different emotion items, followed by repetitions of the neutral items from the first block of trials. A glimpse at Fig. 7 shows that participants took longer to name the ink color in the second block, which included emotion items, than in the first block of only neutral items. Planned comparisons documented an appreciable ESE of 69 ms [$F(1, 29) = 8.86$, $p = .006$]. Performance in the third block, which entailed (a new set of) neutral items, was also worse than in the first block; color naming in the third block was slower by 46 ms than that in the first block [$F(1, 29) = 5.16$, $p = .03$]. We attribute the sluggish performance with neutral items in the third block to sustained effects of exposure to the emotion items in the second block. Because none of the emotion stimuli appeared more than once, buildup of habituation was avoided. As a result, the exposure to the emotion items exerted its influence with full force in both the short term (within the second block itself = ESE) and the long term (sustained effect revealed in the subsequent third block).

Consider next the results of the "emotional repetition" group in the right-hand half of Fig. 7. In the critical second block, the participants were exposed to the same set of emotion items as the other group. However, the presentation of the emotion items was repeated three times in that block. Apart from this difference, the first and third blocks, including (different) sets of neutral items, were the same as those in the other "neutral repetition" group. The results show that performance was worse in the second block, with the emotion items, than in the first block, with neutral items. The 56-ms difference in favor of the first block amounted to an appreciable ESE [$F(1, 29) = 6.07$, $p = .02$]. However, we did not record a sustained effect of the exposure to emotion stimuli in this group. Performance in the third block, which followed emotional exposure, was on a par with that in the first block, which preceded encounter with emotion (806 and 816 ms, respectively; $F < 1$). We attribute the absence of sustained effects to the repetition of the emotion items in the second block. Habituation was generated, which, in turn, diminished the power of emotion to affect subsequent performance. Indeed, performance with respect to the last 40 emotion items within the second block itself was already as fast as that in the first and the third blocks (mean of 802 ms), documenting the absence of a sustained effect within close temporal vicinity to the emotion items.

To recap, the difference between the two groups is readily seen when comparing performance across the third and the first blocks of trials. Performance deteriorated in the "neutral repetition" group due to the lack of habituation to the emotion items in the second block of trials. Consequently, that exposure impaired subsequent performance with neutral stimuli. Because habituation set in with the "emotional repetition" group due to the repetition on items, sustained effects were absent.

In Fig. 8, we summarize the main results of Experiment 3. As can be seen in the left-hand part of Fig. 8, the ESEs were comparable in the two groups. However, sustained effects, gauged by the difference between post- and pre-emotion

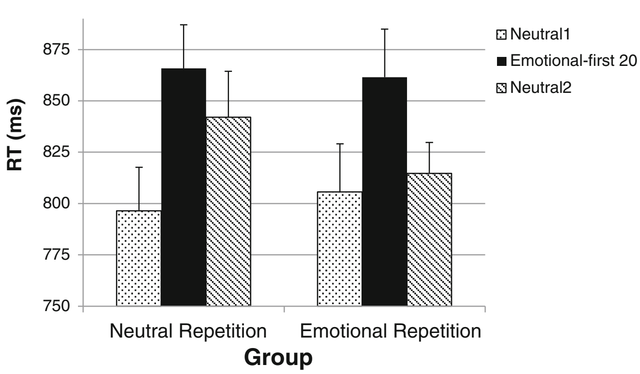

Fig. 7 Mean reaction times (RTs) to name the ink color of words in the two groups of participants. The first and third blocks of trials were the same in the two groups, each consisting of neutral words. The second block included a set of emotion words followed by repetitions of the neutral words from the first block (neutral repetition group, at the left) or included three repetition of the same set of emotion words (emotional repetition group, at the right). Error bars denote one standard error around the mean

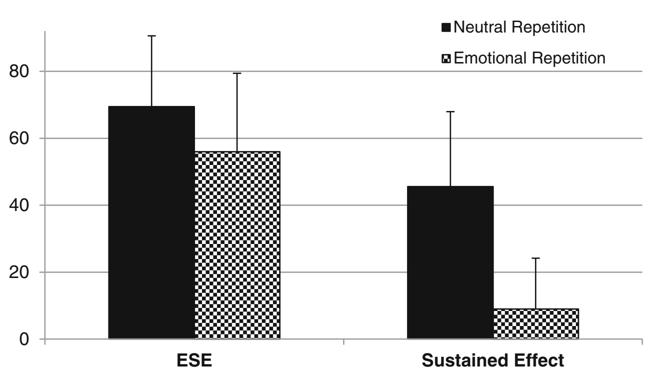

Fig. 8 Emotional Stroop effect (ESE) and the sustained effect of emotional encounter in two groups differing in repetition of emotional stimuli. The ESE is the mean RT difference in color naming between the second and first blocks of trials in each group. The sustained effect is the mean reaction time (RT) difference in color naming between the third and first blocks of trials in the each group. The error bars denote one standard error around the mean

performance with neutral items, only appeared in the group that did *not* experience repeated presentation of the emotional items. Although the ESEs are comparable in the two groups, a sustained effect is present only in the group in which the repetition of emotion items was avoided.

Discussion

Experiment 3 pitted two contrasting predictions on long term effects against one another. According to one account, the decisive factor is the sheer number of emotion stimuli presented: The greater that number, the larger the impact of emotion on future behavior. This set-size account ignores the composition of the emotion items and refers only to their total number. According to the rival habituation account, the decisive factor is not the sheer number of the emotion items that one encounters but rather their repetitive nature. If the same emotion stimuli appear recurrently, they lose their emotive power. The results of Experiment 3 results support the habituation account. What matters is the presence of repeated exposure to the same negative stimulus rather than the mere number of emotion stimuli encountered. Repeated exposure to a negative stimulus immunizes the person against excessive bad feelings and poor generic performance generated by the stimulus.

In summary, Experiments 1–3 elucidated the conditions under which sustained effects of emotion do and do not appear. However, the presence of sustained effects mandates reappraisal of conventional wisdom with respect to the ESE. Given the massive exposure to emotion stimuli in the standard ESE experiment, the lack of reference to sustained effects is surprising. So is the fact that none of the existing accounts of the ESE entails reference to long term effects. The results of the present study speak to this this lacuna in theory. We offer a new perspective that does incorporate long term effects. Specifically, we suggest that the ESE comprises a case of mood induction. Experiment 4 tests this hypothesis.

In Experiment 4, we followed the standard ESE procedures (in particular, those of Exp. 2) with a single notable addition. Upon completing each block of trials (whether composed of emotion or neutral items) the participants responded to a pair of questionnaires gauging mood. We also measured the time needed to complete the questionnaires given that negative or depressed mood is associated with behavioral- or psychomotor-retardation (Clark, 1983; Gerrardshesse et al., 1994). The ESE might comprise such a retardation itself. We expected that participants completing the block with negative words display an elevated state of negative mood compared with those completing the block with neutral words. We also expected completion time of the questionnaires to take longer after responding in the block with emotion words than after responding in the block with neutral words. Finally, we expected to observe sustained effects of negative mood so that the effect is large when emotion items follow neutral items but that it vanishes with the reverse order of exposure.

Negative mood was further evaluated in an implicit fashion by analyzing the vocal responses (to the ink colors). We focused on the lowest fundamental frequency, the minimum f_0, of the vocal responses of the participant (e.g., Ellgring & Scherer, 1996; Scherer, Ladd, & Silverman, 1984; C. E. Williams & Stevens, 1969, 1972) and on loudness (e.g., Barrett & Paus, 2002; Siegman & Boyle, 1993) and hypothesized that participants speak in a softer voice and at an elevated f_0 when naming the color of an emotion item. Sustained effects were expected to emerge with those parameters of voice, too.

Experiment 4

Method

Participants

The participants were 52 Tel Aviv University undergraduates none of whom took part in the previous experiments.

Stimuli, apparatus, and design

The apparatus, stimuli, and design were the same as in Experiment 2 except for a single notable exception. Following each block of trials, the participants filled out two computerized mood questionnaires. The first was the Profile of Mood States (POMS), a tool targeting healthy individuals (McNair, Lorr, & Droppleman, 1971). The Hebrew version (translated and validated by Nets, Zeav, Arnon, & Daniel, 2005) includes 28 items referring to five mood states: *vigor*, *depression*, *fatigue*, *anger*, and *tension*. For each state, the participants responded by marking a five point scale from 0 (*not at all*) to 4 (*extremely so*) on the current emotion. For a second self-report measure of mood, the Current Feeling Scale (CFS), the participants conveyed their feelings by marking each of 5 seven-point bipolar adjective scales: happy–sad, pleasant–unpleasant, positive–negative, bad–good, and comfortable–uncomfortable (Byrne & Clore, 1970; Griffitt, 1970; Griffitt & Veitch, 1971).

For the voice data, the individual color-naming recordings from each participant in each experimental block were analyzed using the Computerized Speech Laboratory (CSL) software by KayPentax, model 4500. The program's output provided the means for loudness (in dB), for f_0, for minimum f_0, and for f_0 range.

Procedure

The procedure was the same as in Experiment 2. Each participant was randomly assigned into one of two groups differing in order of presentation. In the NE group, the participants first performed in the block with neutral items and then performed in the block with emotion items. Immediately after the completion of each block, the participants filled out the POMS and the CFS questionnaires. For the participants in the EN condition, the blocks were presented in the reverse order. In this group, too the participants twice filled out the POMS and CFS questionnaires at the end of each block of trials.

Results

The data with respect to the ESE largely replicated those of Experiment 2. A reliable effect of 21 ms in the NE group [$t(25) = 2.15$, $p = .04$] evaporated with the participants in the EN group who performed in the reverse order [$F(1, 50) = 4.21$, $p = .045$, for the interaction of order of blocks and stimulus valence]. Our main interest in this experiment focused, however, on mood induction. The mean ratings on the POMS and the CFS questionnaires are given in Fig. 9. The data were analyzed in a repeated measures ANOVA with block valence (E, N) and questionnaire (POMS, CFS) as within-subjects variables and group order (EN, NE) as a between-subjects variable. Visual inspection of Fig. 9 reveals the presence of a large effect on negative mood of exposure to negative words [$F(1, 50) = 20.92$, $p < .0001$], which is nonetheless qualified by the interaction with order of presentation [$F(1, 50) = 5.13$, $p = .03$]. This interaction reflects the sustained action of negative mood once it is induced.

Consider the results with respect to the POMS scale in the left-hand panel of Fig. 9. In the NE group, the participants experienced negative mood following performance in the block with emotion items to a greater extent than they did following performance in the block with neutral items. The participants in the EN group also tended to experience negative mood in excess after completing the negative items, but not significantly so. The parity reflects the persistence of negative mood after initial exposure to negative items. Essentially the same results obtained with the CFS (in the right-hand panel of Fig. 9), although the sustained effect was less pronounced in the EN group.

Mood state also affected the time needed to complete the two questionnaires at the end of each block of trials. An overall ANOVA revealed a main effect of first versus second presentation of the questionnaire [a 312-ms difference favoring the second testing; $F(1, 150) = 130.38$, $p < .0001$]. This trivial result reflects the fact that all participants filled out the scales the second time faster than they did first time. Over and above this effect of practice, it took participants longer to complete the scales following exposure to negative words than following exposure to neutral words (by 47 ms, on average), a difference that was qualified by an interaction with group [$F(1, 50) = 4.20$, $p = .046$].

The results of the analysis of voice patterns supported the presence of negative mood. Overall, the participants spoke in a softer voice in response to emotion items (mean of 51.03 dB) than they did in response to neutral items [mean of 51.47 dB; $F(1, 50) = 4.47$, $p = .04$]. This result was qualified by an interaction with order [$F(1, 50) = 17.88$, $p < .0001$; see Fig. 10], so that the difference in loudness favoring neutral items was limited to the NE group (means of 48.1 and 50.5 dB, respectively, for emotional and neutral items in that group; Tukey HSD, $p < .001$). Apparent in Fig. 10, too, is an overall difference in loudness between the two groups [$F(1, 50) = 2.65$, $p = .11$]. An ad hoc explanation would implicate the larger number of men randomly assigned to the EN group.

For the parallel results on the minimum f_0 parameter, our participants exhibited an increase in minimum f_0 in the emotion block as compared with the neutral block [frequencies of 71.56 and 71.12 Hz, respectively; $F(1, 50) = 6.75$, $p = .01$]. For minimum f_0, too, the interaction with order [$F(1, 50) = 5.36$, $p = .002$; see Fig. 11] indicates that only in the NE group was an effect on minimum f_0 produced in the emotional block (means

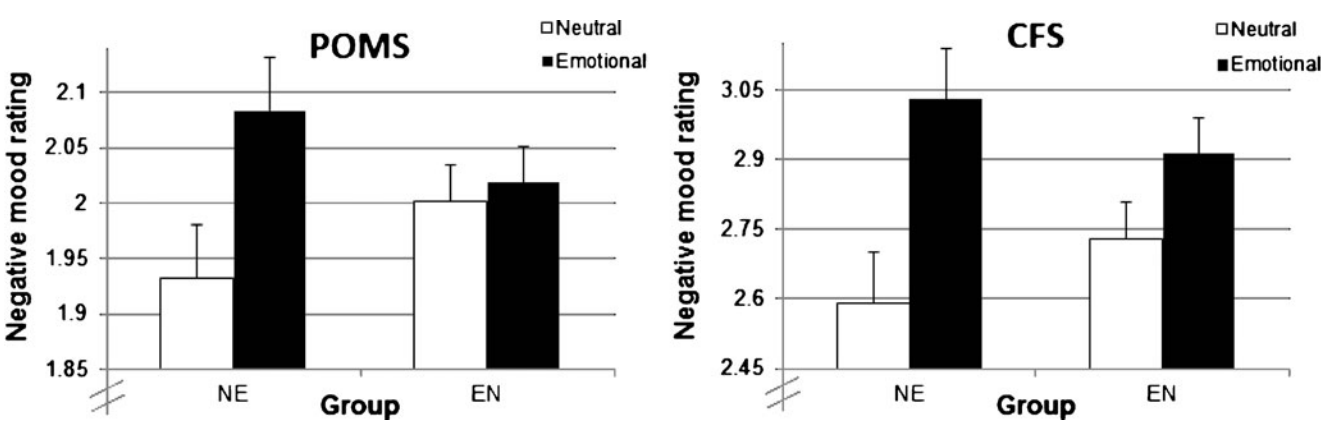

Fig. 9 Mean ratings on the two mood questionnaires, in two groups defined by the order of blocks. The black columns stand for ratings following a block with emotion items, whereas the white columns stand for ratings following a block with neutral items. Error bars represent one standard error around the mean

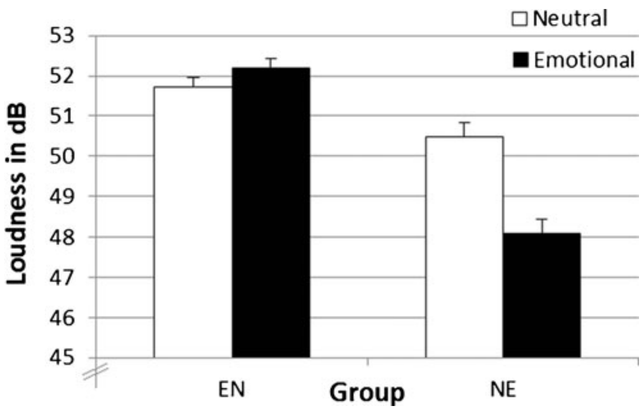

Fig. 10 Loudness (in dB) for naming ink colors across emotion and neutral items, in two groups of participants defined by block order. Error bars denote one standard error around the mean

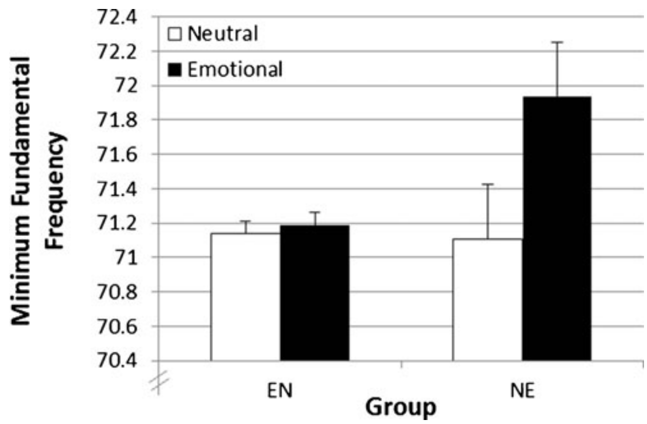

Fig. 11 Minimum f_0 (in Hz) across emotion and neutral items, in two groups of participants defined by block order. Error bars denote one standard error around the mean

of 71.94 and 71.11 Hz, respectively, for the emotion and neutral blocks; Tukey HSD, $p = .006$). Once again, there was an initial difference between the two groups, $F(1, 50)$, $p = .065$.

We have also examined the pattern of correlations across the individual participants with respect to the ESE. If emotion drives the ESE directly, the degree of interference should correlate with the affective response. We found that the ESE correlated with responses on the POMS questionnaire (the difference between the second and the first completion; $r = .24$, $p = .047$). A similar correlation, marginally significant, was found between the ESE and the responses on the CFS questionnaire ($r = .18$, $p = .097$). The overall correlation between the combined responses on the two questionnairs and the ESE amounted to $r = .22$ ($p = .06$). Notably, there was a correlation between the time to complete the POMS and the ESE ($r = .32$, $p = .01$, the first time and $r = .25$, $p = .037$, the second time). Concerning the implicit variable of voice, f_0 correlated with responses on the CFS ($r = .28$, $p = .022$) and correlated similarly with the combined response on the CFS and POMS ($r = .26$, $p = .031$).[1]

Discussion

The results of Experiment 4 provide support for the hypothesis that exposure to a few dozen negative words within the ESE paradigm induces a commensurate state of mood. Upon exposure to a block of negative items, the participants experienced negative mood, which differed reliably from the state of mood that emerged upon completion of a block with neutral words. The negative mood thus induced was expressed both in self-report questionnaires and in an implicit way feature of the participants' vocal reactions. Exposure to negative words wrought about a sustained effect of negative mood, too. Once this state set in, it was lingering on well beyond the first exposure. The exposure to negative words was also conductive to sluggish performance on such a trivial task as the filling out of a questionnaire. Some of these effects were visible on the individual level, too, evinced by the correlation between the affective response and the size of the ESE, The ESE itself may document, in part, psychomotor retardation associated with negative or depressive mood states.

Finally, the finding of mood induction within the framework of the ESE is notable from a methodological point of view, too. This process occurred with single-word presentations in a context that did not encourage reading. Nevertheless, a negative mood did set in, demonstrating the power of emotion signals to intrude on ongoing performance and mood.

General discussion

Sustained effects of exposure to negative stimuli are well known in everyday experience. In contrast, few ESE studies in the laboratory have tested or documented sustained effects of emotion—despite the fact that the participants are exposed to scores of negative stimuli in this paradigm. One would expect the effect of such an exposure to linger on, so that performance in a block of neutral items that followed one with emotion items would be unusually sluggish. Again, this order-of-blocks effect is missing from large portions of the existing ESE research. We attribute the absence in the laboratory of an order-of-block or sustained effect to habituation. Recall that the ESE is calculated as the difference in the *mean* performance between blocks with emotion and neutral stimuli, so that it is based on *all* items. However, fine-grain analysis of individual items (Exp. 1) revealed that this global ESE is actually based on the first few trials, with subsequent presentations (mostly, repetitions of the same small set of words) producing (largely) matched performance across emotion and

[1] All correlations were in the predicted direction (sign-wise) and were fairly noticeably numerically, but several did not reach conventional levels of statistical significance. We report here those that did.

nonemotion items. Consequently, we argue that experimental habituation, characteristic of ESE studies, is the reason for the lack of an order effect in much of the pertinent literature. Consistent with this hypothesis, when we removed habituation by presenting each item only once (Exps. 2–3), an order-of-block effect did emerge, evincing sustained effects of emotion in the laboratory, too. Under habituation-free conditions, an ESE obtains only when emotion stimuli *follow* neutral ones, but does not obtain in the reverse sequence of events.

Existing accounts of the ESE, whether attention or freezing, readily explain the slowdown observed with emotion items upon their presentation or within a small time window of a few seconds (at most). However, the accounts are silent on long-term effects, so that they do not conveniently account for the present results. By limiting purview to concurrent processing (or to that of a small sequence of items; cf. McKenna & Sharma, 2004; see also Frings et al., 2010), existing theorizing is unnecessarily restrictive. This outlook ignores the trivial fact of massive exposure to negative items within the ESE experiment. It is unlikely that the participant can escape with impunity long term effects of such an exposure. It is at this juncture that we propose the novel perspective of mood induction. If the ESE ineluctably induces negative mood, long-term effects are expected and are readily accounted for.

Consequently, Experiment 4 tested the prediction that negative mood is induced within the ESE procedure. The results attested to the presence of an elevated negative mood following performance with emotion items compared with neutral items. Moreover, this mood was lingering on to affect a subsequent set of neutral items. In a similar fashion, it took participants longer to fill out the same mood questionnaires following exposure/performance with negative stimuli than with neutral stimuli. Finally, negative mood was documented in an implicit fashion by analyzing the voice of the participants responding to the ink color of negative and neutral items. We note that mood induction has sometimes been mentioned in the ESE literature (e.g., Gilboa-Schechtman et al., 2000; Richards et al., 1992). However, in virtually all instances mood was induced *prior* to the ESE task. The variable of interest was the match or mismatch between the induced mood and the ESE stimuli. In contrast, we suggest that the ESE task itself comprises a mood induction procedure. The present hypothesis can also explain the paucity of ESE with positive items (Richards et al., 1992; J. M. G. Williams et al., 1996).

We are aware that mood induction is vulnerable to *demand characteristics* (e.g., Buchwald, Strack, & Coyne, 1981). For example, participants often attempt to fake fatigue and depression-like reactions in such situations. However, speeded responses requiring maximal performance (Clark, 1983; McKenna & Lewis, 1994) or implicit measures such as features of speeded speech are much less susceptible to faking. In the present study, these measures were further correlated with those of self-report, conferring validation support on the results as a whole. Sustained effects observed with both vocal parameters and performance also suggest that our results are not due to demand characteristics.

The present results establish the role of mood and affect in the generation of the ESE and in the maintenance of long-term effects. Nevertheless, one should not rule out a role for further variables, in particular for that of attention. The results of the study by Blanchette and Richards (2013) implicated attention to threat (in high-anxious individuals) as the root cause of the ESE. Contrary to the present results, these authors did not find a correlation between affect and the ESE. We note, though, that potentially significant differences in method and design separate the two studies. In particular, the conditioned negativity of the items in the Blanchette and Richards study was probably short-lived, and possibly affected by repetition. In Experiment 4, we eschewed repetition of items in order to exclude habituation. The issue deserves further scrutiny.

Nevertheless, we agree with Blanchette and Richards (2013) on the (partial) contribution of attention. Having been exposed to the negative items in the emotion block, the participants are alerted to alluring danger. Consequently, they might learn to pay more attention to the carrier words in a subsequent block with neutral items. This extra attention will, in turn, engender slower responses to color. Explanations based exclusively on processes of attention meet some difficulty though in accounting for (a) the psychomotor retardation observed when filling out the mood questionnaires, (b) the mood scores themselves, as well as (c) the change in vocal responses. Possibly, mood augmented by attention provides the best explanation at this point.

What is the mechanism by which negative mood interferes with performance? The resource allocation model by Ellis and Ashbrook (1988, 1989), originally developed to explain mood induced behavioral deficit, has been influential in various cognitive domains. According to the model, the momentary emotional state acts to regulate the amount of capacity allocated to a given task. The reason is that the large amount of mental resources drawn by the need to control emotion (an intra-person process) takes a toll on performance with the task at hand. This mechanism can be described in terms of attention allocation (think of the attention model of the ESE), but it is different from the common attention explanation. In particular, the resource allocation mechanism mandates the allocation of cognitive resources to deal with the *individual's emotional state* rather than with any task-irrelevant aspect of the external stimulus. As a result, the effect of emotion is *general*, so that responses under emotion are slower on *any* task. This prediction is consistent with the findings by Algom et al. (2004) who reported that responses to emotional items are sluggish regardless of the task at hand (whether color naming or reading). We contend that the long exposure to emotional content included within the ESE mandates the allocation of cognitive resources to process the associated emotional experience so that performance is affected on subsequent tasks, too (regardless of the valence of the pertinent items).

On the level of brain circuitry, we speculate that the one supporting responses under arousal and anxiety is similarly active in negative mood. Recall that anxiety and tension was one of the mood qualities measured in the profile of mood states probed in Experiment 4. It joined other manipulations producing (not merely measuring) overall distress. Neural signals observed in anxiety are generated by sustained activity of the insular cortex and the ventral basal forebrain (including the bed nucleus of the stria terminalis). These play a modulatory role in stress and arousal management (Davis, 1988; Paulus & Stein, 2006; Somerville et al., 2013). The role of the insula in physiological regulation under anxiety is widely supported (Hoehn-Saric, Schlund, & Wong, 2004; Paulus, Feinstein, Castillo, Simmons, & Stein, 2005; Paulus & Stein, 2006). We tentatively suggest that this pathway is also the dominant neural circuit associated with the sustained effects of negative mood recorded in this study.

Given the prominent role of mood in generating the ESE, a natural question refers to its scope. Do the mood induction and habituation effects extend to positive words? Can mood induction explain the interference reported when ornithologists name birds (Dalgleish, 1995)? We do not think so. The present theory applies only to negative mood as a main source for generating the ESE. The effects associated with positive mood are generally much weaker than those associated with negative mood (Westerman, Spies, Stahl, & Hesse, 1996; see also Kenealy, 1986). Of more consequence, positive mood encourages speedy responding rather than retardation For example, in the original study by Velten (1968) the participants in positive mood displayed faster writing speed and decision time and generated word associates more than did those in negative or neutral mode (see, again, the meta-analysis by Westerman et al., 1996, who reported the same dissociation between negative and positive mood on an unusually large range of measures). In a commensurate manner, the ESE is typically smaller if it is at all recorded with positive stimuli.

The present results and conclusions invite another look at explanations based on executive control and selective attention Failure of fully selective attention to the ink color can derive not only from inadvertent word reading (the classic Stroop effect) or (in addition) from the menacing content of the carrier word (the ESE) but also from the generic factor of negative mood. The participant fails to focus on the ink color not merely due to the emotional content of the word but (also) due to the negative mood induced by the atmosphere created by the words as a group. Research is needed to disentangle these separate factors of attention under emotion (cf. Mama, Ben-Haim, & Algom, 2013). Further studies are also needed in order to assess if the present findings generalize to clinical populations and additional cognitive tasks (cf. Amir et al., 2009).

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. doi:10.1037/0096-3445.133.3.323

Algom, D., Zakay, D., Monar, O., & Chajut, E. (2009). Wheel chairs and arm chairs: A novel experimental design for the emotional Stroop effect. *Cognition and Emotion, 23,* 1552–1564.

Amir, N., Najmi, S., & Morrison, A. S. (2009). Attenuation of attention bias in obsessive–compulsive disorder. *Behaviour Research and Therapy, 47,* 153–157. doi:10.1016/j.brat.2008.10.020

Banse, R., & Scherer, K. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*(3), 614–636.

Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2007). Threat-related attentional bias in anxious and non-anxious individuals: A meta-analytic study. *Psychological Bulletin, 133,* 1–24. doi:10.1037/0033-2909.133.1.1

Barrett, J., & Paus, T. (2002). Affect-induced changes in speech production. *Experimental Brain Research, 146,* 531–537.

Becker, E. S., Rinck, M., Margraf, J., & Roth, W. T. (2001). The emotional Stroop effect in anxiety disorders: General emotionality or disorder specificity? *Anxiety Disorders, 15,* 147–159.

Ben-David, B. M., Calderon, N., & Algom, D. (2005). A signal detection analysis of the emotional Stroop effect. In J. S. Monahan, S. M. Sheffert, & J. T. Townsend (Eds.), *Fechner Day 2005* (pp. 35–38). Traverse City MI: International Society for Psychophysics.

Ben-David, B. M., Chajut, E., & Algom, D. (2012). The pale shades of emotion: A signal detection theory analysis of the emotional Stroop task. *Psychology, 3,* 537–541.

Blanchette, I., & Richards, A. (2013). Is emotional Stroop interference linked to affective responses? Evidence from skin conductance and facial electromyography. *Emotion, 13,* 129–138.

Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1996). Picture media and emotion: Effects of a sustained affective context. *Psychophysiology, 33,* 662–670.

Buchwald, A. M., Strack, S., & Coyne, J. C. (1981). Demand characteristics and the Velten mood induction procedure. *Journal of Counseling and Clinical Psychology, 49,* 478–479.

Byrne, D., & Clore, G. L. (1970). A reinforcement model of evaluative responses. *Personality: An International Journal, 31,* 103–128.

Chajut, E., Mama, Y., Levy, L., & Algom, D. (2010a). Avoiding the approach trap: A response bias theory of the emotional Stroop effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1567–1572.

Chajut, E., Schupak, A., & Algom, D. (2010b). Emotional dilution of the Stroop effect: A new tool for assessing attention under emotion. *Emotion, 10,* 944–948.

Clark, D. (1983). On the induction of depressed mood in the laboratory—Evaluation and comparison of the velten and musical procedures. *Advances in Behaviour Research and Therapy, 5,* 27–49.

Coleman, R. E. (1975). Manipulation of self.esteem as a determinant of mood of elated and depressed women. *Journal of Abnormal Psychology, 84,* 693–700.

Dalgleish, T. (1995). Performance on the emotional Stroop task in groups of anxious, expert, and control subjects: A comparison of computer and card presentation formats. *Cognition and Emotion, 9,* 341–362.

Davis, M. (1988). Are different parts of the extended amygdala involved in fear versus anxiety? *Biological Psychiatry, 15,* 1239–1247.

Ellgring, H., & Scherer, K. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior, 20,* 83–110. doi:10.1007/BF02253071

Ellis, H. C., & Ashbrook, P. W. (1988). Resource allocation model of the effects of depressed mood states on memory. In K. Fiedler & J.

Forgas (Eds.), *Affect, cognition, and social behavior: New evidence and integrative attempts* (pp. 25–43). Toronto, ON: Hogrefe.

Ellis, H. C., & Ashbrook, P. W. (1989). The "state" of mood and memory research: A selective review. *Journal of Social Behavior and Personality, 4,* 1–21.

Fox, E., Russo, R., Bowles, R., & Dutton, K. (2001). Do threatening stimuli draw or hold visual attention in subclinical anxiety? *Journal of Experimental Psychology: General, 130,* 681–700. doi:10.1037/0096-3445.130.4.681

Frings, C., Englert, J., Wentura, D., & Bermeitinger, C. (2010). Decomposing the emotional Stroop effect. *Quarterly Journal of Experimental Psychology, 63,* 42–49. doi:10.1080/17470210903156594

Frost, R., & Plaut, D. (2001). The Word-Frequency Database for Printed Hebrew. Retrieved from http://word-freq.mscc.huji.ac.il/index.html

Gerrards-Hesse, A., Spies, K., & Hesse, F. (1994). Experimental inductions of emotional states and their effectiveness—A review. *British Journal of Psychology, 85,* 55–78.

Gilboa-Schechtman, E., Revelle, W., & Gotlib, I. (2000). Stroop interference following mood induction: Emotionality, mood congruence, and concern relevance. *Cognitive Therapy and Research, 24,* 491–502.

Gomez, P., Zimmermann, P. G., Guttormsen-Schar, S., & Danuser, B. (2009). Valence lasts longer than arousal: Persistence of induced moods as assessed by psychophysiological measures. *Journal of Psychophysiology, 23,* 7–17.

Gouaux, C., & Gouaux, S. M. (1971). The influence of induced affective states on the effectiveness of social and non-social reinforcers in an instrumental learning task. *Psychonomic Science, 22,* 341–343.

Griffitt, W. (1970). Environmental effects on interpersonal affective behavior: Ambient effective temperature and attraction. *Journal of Personality and Social Psychology, 15,* 240–244.

Griffitt, W., & Veitch, R. (1971). Hot and crowded: Influences of population density and temperature on interpersonal affective behavior. *Journal of Personality and Social Psychology, 17,* 92–98.

Higgins, E., Rholes, W., & Jones, C. (1977). Category accessibility and impression-formation. *Journal of Experimental Social Psychology, 13,* 141–154.

Hoehn-Saric, R., Schlund, M. W., & Wong, S. H. (2004). Effects of citalopram on worry and brain activation in patients with generalized anxiety disorder. *Psychiatry Research, 131,* 11–21.

Holle, C., Neely, J. H., & Heimberg, R. G. (1997). The effects of blocked versus random presentation and semantic relatedness of stimulus words on response to a modified Stroop task among social phobics. *Cognitive Therapy and Research, 21,* 681–697.

Kenealy, P. M. (1986). The Velten mood induction procedure: A methodological review. *Motivation and Emotion, 10,* 315–335.

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. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition, 32,* 474–488.

Mama, Y., Ben-Haim, M. S., & Algom, D. (2013). When emotion does and does not impair performance: A Garner theory of the emotional Stroop effect. *Cognition and Emotion, 27,* 589–602. doi:10.1080/02699931.2012.726212

Matheny, K. B., & Blue, F. R. (1977). The effects of self-induced mood states on behavior and physiological arousal. *Journal of Clinical Psychology, 3,* 936–940.

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

McKenna, F. P., & Lewis, C. (1994). A speech rate measure of laboratory induced affect—The role of demand characteristics revisited. *British Journal of Clinical Psychology, 33,* 345–351.

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. doi:10.1037/0278-7393.21.6.1595

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.

McNair, D. M., Lorr, M., & Droppleman, L. F. (1971). *EdITS manual for the Profile of Mood States*. San Diego, CA: Educational and Industrial Testing Service.

McNally, R. J., Riemann, B. C., & Kim, E. (1990). Selective processing of threat cues in panic disorder. *Behaviour Research and Therapy, 28,* 407–412.

Natale, M. (1977). Effects of induced elation–depression on speech in the initial interview. *Journal of Consulting and Clinical Psychology, 45,* 45–52.

Natale, M. (1978). Effect of induced elation and depression on internal locus of control. *Journal of Psychology, 100,* 315–321.

Nets, Y., Zeav, A., Arnon, M., & Daniel, S. (2005). Translating a single-word items scale with multiple subcomponents—A Hebrew translation of the Profile of Mood States. *Israeli Journal of Psychiatry Related Science, 42,* 263–270.

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130,* 466–478. doi:10.1037/0096-3445.130.3.466

Parrott, W., & Sabini, J. (1990). Mood and memory under natural conditions—Evidence for mood incongruent recall. *Journal of Personality and Social Psychology, 59,* 321–336. doi:10.1037/0022-3514.59.2.321

Paulus, M. P., Feinstein, J. S., Castillo, G., Simmons, A. N., & Stein, M. B. (2005). Dose-dependent decrease of activation in bilateral amygdala and insula by Lorazepam during emotion processing. *Archives of General Psychiatry, 62,* 282–288.

Paulus, M. P., & Stein, M. B. (2006). An insular view on anxiety. *Biological Psychiatry, 60,* 383–387.

Pereira, M. G., Volchan, E., Leal de Souza, G. G., Oliveira, L., Campagnoli, R. R., Pinheiro, W. M., & Pessoa, L. (2006). Sustained and transient modulation of performance induced by emotional picture viewing. *Emotion, 6,* 622–634.

Phaf, R. H., & Kan, K. J. (2007). The automaticity of emotional Stroop: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry, 38,* 184–199. doi:10.1016/j.jbtep.2006.10.008

Rankin, C. H., Abrams, T., Barry, R. J., Bhatnagar, S., Clayton, D. F., Colombo, J., & Thompson, R. F. (2009). Habituation revisited: An updated and revised description of the behavioral characteristics of habituation. *Neurobiology of Learning and Memory, 92,* 135–138.

Richards, A., French, C. C., Johnson, W., Naparstek, J., & Williams, J. (1992). Effects of mood manipulation and anxiety on performance of an emotional Stroop task. *British Journal of Psychology, 83,* 479–491.

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America, 76,* 1346–1356.

Segerstrom, S. C. (2001). Optimism and attentional bias for negative and positive stimuli. *Personality and Social Psychology Bulletin, 27,* 1334–1343.

Siegman, A., & Boyle, S. (1993). Voices of fear and anxiety and sadness and depression—The effects of speech rate and loudness on fear and anxiety and sadness and depression. *Journal of Abnormal Psychology, 102,* 430–437. doi:10.1037/0021-843X.102.3.430

Smith, J. C., Bradley, M. M., & Lang, P. J. (2005). State anxiety an affective physiology: Effects of sustained exposure to affective pictures. *Biological Psychology, 69,* 247–260.

Sobin, C., & Alpert, M. (1999). Emotion in speech: The acoustic attributes of fear, anger, sadness, and joy. *Journal of Psycholinguistic Research, 28,* 347–365.

Somerville, L. H., Wagner, D. D., Wig, G. S., Moran, J. M., Whalen, P. J., & Kelley, W. M. (2013). Interactions between transient and sustained neural signals support the generation and regulation of anxious emotion. *Cerebral Cortex, 23,* 49–60. doi:10.1093/cercor/bhr373

Teasdale, J. D., Taylor, R., & Fogarty, S. J. (1980). Effects of induced elation-depression on the accessibility of memories of happy and unhappy experiences. *Behaviour Research and Therapy, 18,* 339–346.

Thompson, R. F. (2009). Habituation: A history. *Neurobiology of Learning and Memory, 92,* 127–134.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour Research and Therapy, 6,* 473–482.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A metaanalysis. *European Journal of Social Psychology, 26,* 557–580.

Williams, C. E., & Stevens, K. N. (1969). On determining the emotional state of pilots during flight: An exploratory study. *Aerospace Medicine, 40,* 1369–1372.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustical Society of America, 52,* 1238–1250.

Williams, J. M. G., Mathews, A., & MacLeod, C. (1996). The emotional Stroop task and psychopathology. *Psychological Bulletin, 120,* 3–24. doi:10.1037/0033-2909.120.1.3

Witthöft, M., Rist, F., & Bailer, J. (2008). Enhanced early emotional intrusion effects and proportional habituation of threat response for symptom and illness words in college students with elevated health anxiety. *Cognitive Therapy and Research, 32,* 818–842.