Journal of Experimental Psychology: General
2005, Vol. 134, No. 4, 592–595

Copyright 2005 by the American Psychological Association
0096-3445/05/$12.00 DOI: 10.1037/0096-3445.134.4.592

# Vicissitudes of a Misnomer: Reply to Dalgleish (2005)

Eran Chajut
The Open University

Shlomo Lev and Daniel Algom
Tel Aviv University

The Stroop effect is psychology's classic measure gauging the selectivity of attention to individual attributes of complex stimuli. The emotional Stroop effect gauges the influence on behavior of threat and emotional stimuli. The former taps central/executive processes abstracted from particular stimulus contexts, whereas the latter taps automatic processes inextricably linked to particular stimuli in the environment. T. Dalgleish (2005) raised concerns about the data and theory that support the separateness of the 2 effects (D. Algom, E. Chajut, & S. Lev, 2004). The present reply shows that Dalgleish's objections are unwarranted and that the term *emotional Stroop effect* blurs the deep conceptual divide separating the 2 phenomena.

*Keywords:* conflict words, threat words, attention, emotion

The Stroop task and the emotional Stroop task share several notable features. They both entail the presentation of words printed in colors and the naming of the colors in which the words are printed, and, significantly, they both document the influence of the carrier words (i.e., meaning), irrelevant to the task at hand, on color naming. Dalgleish's (2005) comment is predicated on this considerable similarity between the two tasks and its implications. However, we show that fundamental differences in conceptual construal, measurement model, and sensitivity to manipulations of attention overshadow the noted similarities. A formal infrastructure underlies the Stroop phenomenon within the framework of which stimuli, derivations, and the effect itself form a tightly connected system. The emotional Stroop effect documents another phenomenon, the deep-seated sensitivity of organisms to threats in the environment and the need to address them in a radical fashion. In the classic effect, word and color are vehicles conveying a well-defined logical relationship that often engenders impairment to performance with each attribute. In the emotional effect, words are carriers of threats, affecting ongoing activity with color. Therefore, the intrusion of word on color is an incidental case of convergence on a similar outcome of two largely separate processes.

## Classic and Emotional Stroop Effects: The Incommensurability of Stimuli, Conceptual Construal, and Derivation

Measurement and attendant definitions are supreme in empirical science. Dalgleish's pertinent comments thus invite close scrutiny by way of isolating potentially different phenomena and avoiding confusion. Dalgleish pinpointed the fact that Stroop (1935) did not present congruent stimuli (color words printed in matching colors) in his original study. Stroop recorded performance with incongruent stimuli (color words printed in conflicting colors) and neutral stimuli (various shapes in color), deriving what we now call *Stroop interference*. Dalgleish further referred to Macleod (1991) to underscore the popularity of Stroop interference in gauging the Stroop phenomenon. Consequently, Dalgleish suggested that congruent stimuli are "not part of any uniformly accepted definition of the Stroop effect" (Dalgleish, 2005, p. 586), challenging our definition of the effect as the difference in performance between congruent and incongruent stimuli. Moreover, given the primacy of interference, the difference in color-naming performance between emotional and neutral words, the emotional Stroop effect might then be another species of interference properly subsumed under the Stroop effect. The allure of unification notwithstanding, these comments ignore the unique makeup of Stroop stimuli, namely, the logical relationship that the attributes of a Stroop stimulus bear to one another. This property is indispensable for the ensuing test, psychology's first and still most widely used measure of the selectivity of attention.

To appreciate the vital importance of the stimulus makeup, one should recall the feat accomplished in Stroop's (1935) classic study. Stroop, for the first time, combined color words and print colors drawn from a common set of values. The stimuli thus created are endowed with an extraordinary property: Common responses apply to their individual attributes (given an appropriate task). Thus, naming the print colors or naming (reading) the words always results in a response drawn from the same common set of color names. The overlap of the possible responses, in turn, creates two types of stimuli: congruent (word and color go by the same response) and incongruent (word and color go by different responses from the common set). A third possibility does not exist. Logic is involved in the Stroop stimulus structure.

Apart from (a) the existence of a well-defined relation (compatibility) that generates the congruent and incongruent stimuli and (b) the fact that these classes exist preexperimentally and independent of the observer's performance, the classes are logical in the further sense that one class implies the other. Mere appearance of

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

We thank Tim Dalgleish for his helpful comments on a draft of this article.

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

incongruent stimuli implicates the presence of congruent stimuli (and vice versa), regardless of whether both or just one type is presented for view. An experimenter may elect to present only incongruent stimuli (as did Stroop), yet the cognate congruent stimuli are implicated by virtue of that feat. Given a pair of dimensions with shared values, producing all possible combinations always results in congruent and incongruent stimuli. Presenting one class merely means that a subset of the items inhering in the stimulus ensemble was used. Contrary to Dalgleish's remark, congruent stimuli are part and parcel of the Stroop effect and are always present in the environment of the Stroop task, actually or potentially. Therefore, Stroop stimuli are objects composed of attributes that go by the same names and that (by virtue of this feature) entail components that bear a logical relationship to one another.

The logical structure is useful not just because of its formal properties but mainly because the attendant conflict or agreement afford a singularly potent test of selectivity. Here is how. Performance on the target attribute (say, color) is compared across congruent and incongruent items. Better performance with the former—the Stroop effect—betrays the fact that people attended to an attribute (word) that is irrelevant to the task at hand, conferring psychological reality on the a priori classes of congruent and incongruent stimuli whose very definition involves two dimensions. The emotional effect lacks this infrastructure; note, incidentally, that attending to a word does not invariably imply a slowdown (or a speedup) in color naming. For instance, Burt (2002) found faster color naming with familiar than with less familiar words, and Chen and Bargh (1999) found that people were faster to push a lever away (as opposed to pulling it toward the self) when presented with a negative word (even when word evaluation was irrelevant to the task at hand). The null effects are instructive, too. The absence of the Stroop effect shows that people focused on the task-relevant attribute of color (i.e., the irrelevant words did not affect color naming). The absence of the emotional effect does not imply that people focused on color and that the irrelevant words did not affect color naming. It is possible that people attended to the emotional and neutral words (compromising exclusive focus on color) but that the two types of words affected color naming to the same extent. In summary, it is the Stroop stimulus structure that solely enables the powerful test of selective attention, psychology's "Gold Standard of attentional measures" (MacLeod, 1992).

Stroop stimuli and the Stroop effect thus comprise a tight system, which explains why, from a historical point of view, the two terms are coexistent. The latter loses its logic and power without the former (one term is virtually meaningless without the other). Conceiving the task and the ensuing effect, Stroop (1935) was remarkably appreciative when delineating the essential property of the stimuli that have come to bear his name: "The materials employed . . . are different from *any* that have been used" in that they are "*conflicting* stimuli . . . a name of one color printed in the ink of another color" (Stroop, 1935, p. 644; italics added). MacLeod (1991) underscored the centrality of congruence; he perceptively observed that Stroop was not interested in stimulus naming time (eventually, Stroop was not interested in any particular stimulus) but rather in the "interference between conflicting processes" (MacLeod, 1991, p. 164). Dalgleish's (2005, p. 585, 586) own quotes from Stroop (1935) and MacLeod (1991), marshaled as evidence against our definition of the Stroop effect, themselves contain the critical terms "conflicting . . . stimuli" and "interference card [stimuli]," respectively. It is precisely these stimuli that are missing from the emotional Stroop task. The word DEATH, printed in green, is not an incongruent or conflicting stimulus; it is a non-Stroop stimulus. The Stroop effect is not calculable with the stimuli used in studies of the emotional Stroop effect. The difference in naming the print color of the words DEATH and TABLE—the emotional Stroop effect—cannot be a Stroop effect.

The last conclusion stands regardless of how one derives the Stroop effect: as interference (SI; the difference in performance between incongruent and neutral stimuli), as facilitation (SF; the difference between neutral and congruent stimuli), or as their sum ($SE = SI + SF$; the difference between incongruent and congruent stimuli). Dalgleish was correct to point out that, often, SI has been taken to stand for the *SE*. However, regardless of which component one elects to measure (and regardless of how one dubs the effect), the *SE*, SI, and SF each entail congruent and/or incongruent stimuli. Such stimuli, we just recounted, are missing in the emotional Stroop task. Therefore, a less than consistent usage cannot alter the fact that the Stroop effect in all conceivable variation is not defined in the environment of the emotional Stroop task.

Finally, Dalgleish mentioned priming with noncolor words (Warren, 1972, 1974), discussed in MacLeod's review, to challenge the demarcation. However, MacLeod seemed to consider priming and Stroop to be separate phenomena, with the respective literatures making "increasing contact" (MacLeod, 1991, p. 173). It is moot whether Warren himself considered his findings to be Stroop effects (Stroop's seminal study is not referenced in the original Warren article, although other research by Stroop is). Perhaps most important, in a decade-long research based on the Warren studies, Burt (e.g., 1994, 2002) concluded that effects of Stroop and priming with noncolor words likely tap separate processes, with models of the former "offer[ing] little explanatory utility" for the latter (Burt, 2002, p. 1035). Therefore, the centrality of congruence to the interpretation (and generation) of the Stroop effect should not be underestimated. In the vast literature with Stroop-like stimuli, almost every feature of Stroop's (1935) original study is superseded (including word and color). The one feature that remains intact amid the gamut of stimuli, methods, and responses is the division of items into congruent and incongruent classes. As a result, the effect is enabled, and selectivity to the pertinent dimensions can be gauged. The purview of the Stroop phenomenon might not be any particular stimulus/attribute but rather the selectivity of attention to individual aspects of complex stimuli and the sensitivity of the attentional selection to various contextual factors.

## Classic and Emotional Stroop Tasks: The Incommensurability of Design and Measurement

In the classic task, the same word appears in congruent and incongruent combinations; the difference in color naming performance is the Stroop effect for that word. The listwide Stroop effect is actually the sum of such item-specific effects (Jacoby, Lindsay, & Hessels, 2003). The fact that the task-irrelevant word is identical across congruent and incongruent trials ensures that any difference

between them (the Stroop effect) cannot be due to differences in the words or lexical factors associated with word recognition (Larsen, Mercer, & Balota, 2005). In the emotional task, by contrast, the words in the emotional and control lists are never the same. It is impossible in principle to match the stimuli on all conceivable variables. This structural feature also separates the emotional Stroop task from a range of well-known attention tasks.

A common feature of the popular tasks of attention developed by researchers such as Stroop, Garner, Posner, Eriksen, and Treisman is that each places the same stimulus under two different conditions (as congruent or incongruent item, with or without irrelevant variation, in cued/noncued location, as target or flanker, or as target or distractor, respectively). As a result, one can derive item-specific effects in each of these tasks, allowing for stimulus- or location-independent tests of selectivity and focusing. The immutable stimuli of the emotional Stroop task do not enable such a test. Above and beyond the difference in experimental control, this common feature (or lack thereof) permits a penetrating look at the nature of the respective processes. Selectivity and focusing is the purview of attention tested in those tasks; to measure selectivity, it must be abstracted from the particular stimulus context. The emotional Stroop effect, by contrast, is a reaction to a particular stimulus. One person is afraid of cats but not of dogs, whereas another person is afraid of dogs but not of cats. Each reacts to his or her particular stimulus (by, say, slowing down or speeding up ongoing activity). The fears are not interchangeable. Interchangeability, one might recall, is the hallmark of many tasks of selective attention. Therefore, the processes of attention tapped by the Stroop task (and the other tasks noted) are not associated with particular stimuli (although, of course, they are tested with particular stimuli), whereas processes of automatic vigilance and attention to threat are inextricably bound with particular stimuli.

Conceiving the orientation of attention task, Posner referred to attention as covert eye movements (distinguishing it from the overt eye movements of perception); more recently (e.g., Posner & Raichle, 1994), Posner construed attention as a form of executive function. According to Logan and Zbrodoff (1999), the distinctive function of attention is creating logical propositions for subsequent perceptual action and selection. Note that in all the formulations, attention taps central processes and not the arbitrary stimulus platform that these processes engage. In contradistinction, the processes engendered by threat are mainly automatic (preattentive) reactions to the specific stimuli, taking a toll on ongoing activity with other stimuli. Attention is affected, but in a way that differs from that tapped by the Stroop effect. The latter is process bound, whereas the automatic reactions entailed in the former are stimulus bound.

## Empirical Dissociation: Manipulation of Attention Modifies the Classic Effect but Not the Emotional Effect

The seesaw relationship between color and word is a signature of the attentional processes sustaining the Stroop effect. Small to zero amounts of interference (and good performance) with one come at the cost of large interference (hence, poorer performance) with the other. Moreover, the direction of the interference is reversible by changes in the relative salience of the words and the colors. When the words are more salient than the colors (the condition prevailing in the majority of published Stroop studies), then the typical Stroop effect ensues. However, when the words and colors used are matched in salience, then the *asymmetry* in interference (often, the effect itself) disappears. And when the colors are made more salient than the words, a reverse Stroop effect is present, by which the colors interfere with the words more than vice versa (Stroop, 1935; see Melara & Algom, 2003, for a review). Because dimensional interrelation is not the source of the emotional effect, color–word complementarity is absent from the emotional Stroop effect. An emotional slowdown is present for both color naming and word reading (as well as for lexical decision). And this sweeping slowdown is largely unaffected by changes in salience. We examined the emotional Stroop effect under both matched and mismatched salience. A slight tendency toward larger disruption of word reading in the former condition did not alter the overall picture by which emotional words impaired color naming and word reading appreciably under both matched and mismatched salience.

The Stroop effect is extremely sensitive to the presence and magnitude of task-irrelevant variation (as are other measures of attention). When such variation is minimized, the effect (reflecting the failure of selective attention) naturally vanishes. In sharp contrast, the emotional Stroop effect is immune to changes in task-irrelevant variation. In the typical setup, the carrier words vary from trial to trial, so there exists considerable task-irrelevant variation. We eliminated task-irrelevant variation by presenting a single emotional word that changed in color from trial to trial in one block and a single neutral word that similarly changed in color in another block. Color naming was poorer in the former condition, recording for the first time single-stimulus emotional Stroop effects. Remarkably, these effects were on a par with those derived under the typical setup in which the task-irrelevant words do vary from trial to trial.

A third major dissociation concerns sequential effects. Such effects (facilitation or interference on trial $n + 1$ as a function of the foil and response on trial $n$) have been investigated with respect to the Stroop phenomenon. However, the majority of published Stroop studies do not examine sequential effects, perhaps reflecting their relatively minor role with respect to the core Stroop phenomenon. In contrast, the major component of the emotional Stroop effect is itself a sequential effect. Threats are sustained in time, if for a fraction of a second. An emotional word thus mainly affects the response to the subsequent word (regardless of that word's valence). The consequences for experimental design are noteworthy: In a block consisting of emotional words, each word (except the first) is preceded by an emotional word, whereas in a block with neutral words, none is preceded by an emotional word. Because the sustained effects evaporate over trials in a block with mixed presentations, the emotional Stroop effect should routinely obtain when the emotional and neutral words appear in separate blocks. These predictions have been confirmed: Larger effects are found with separate presentation, and often, the effect is absent with mixed presentation (Algom, Chajut, & Lev, 2004; Holle, Neely, & Heimberg, 1997; McKenna & Sharma, 2004; Richards, French, Johnson, Naparstek, & Williams, 1992).

## What's in a Name?

Quite a lot. At the rhetorical level, despite the perfunctory qualifiers, the term *Stroop effect* is used to refer to the emotional

effect without alerting the reader to the differences in calculations and interpretability. At the level of design, in an apparent deference to the Stroop label, color words are sometimes included in studies of the emotional effect without substantive justification. The virtual absence of reading in studies of the emotional effect is another likely outcome of the allegiance forged with the Stroop effect. Incidentally, color naming is not a privileged task in the classic effect (Stroop was careful to test word reading in all of his experiments). At the level of theory, there are attempts at importing models developed for the Stroop effect to the domain of the emotional Stroop effect without due attention to the response overlap and stimulus conflict lying at the heart of the former.

In conclusion, let us elucidate the question posed at the outset: Why do words intrude on naming their print colors? Two fundamentally different processes have been identified. In the first process, an underlying semantic conflict (or agreement) is capitalized on to create stimuli with a complete overlap of responses to their individual attributes. Such stimuli pose a peculiar problem for the cognitive system because the to-be-ignored dimension (word meaning) maps directly onto the response set (color names) of the to-be-selected dimension. The cognitive system experiences difficulty in rejecting the output from word reading (color name) as a task response. In the second process, special words (a threatening word, a taboo word, one's own name) can slow color naming, but this is not due to the difficulty of rejecting those word as valid task responses. Thus, the word *cancer* can slow down color naming, but not because people are at risk of responding "cancer." The word likely engenders a surprise or startle response, and such responses generally slow processing. Emotional words do not pose a problem of dimensional selectivity in a color-naming task because such words are not valid responses.

McKenna and Sharma (2004) recently concluded that "the emotional Stroop effect is not strictly speaking a Stroop effect" (p. 382) and entertained the possibility that it is "a misnomer" (p. 390). They suggested another label ("emotional intrusion"), but they acknowledged that the current term is now well established. We agree, and we issue a caveat to watch out for sundry confusions wrought by the term *emotional Stroop effect*.

## 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.

Burt, J. S. (1994). Identity primes produce facilitation in a color naming task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47*(A), 957–1000.

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

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

Dalgleish, T. (2005). Putting some feeling into it—The conceptual and empirical relationship between the classic and emotional Stroop tasks: Comment on Algom, Chajut, and Lev (2004). *Journal of Experimental Psychology: General, 134,* 585–591.

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

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

Larsen, R. J., Mercer, K., & Balota, D. (2005). *Lexical characteristics and word recognition parameters for emotion words: Effects of word negativity.* Manuscript submitted for publication.

Logan, G. D., & Zbrodoff, N. J. (1999). Selection for cognition: Cognitive constraints on visual spatial attention. *Visual Cognition, 6,* 55–81.

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

MacLeod, C. M. (1992). The Stroop task: The "gold standard" of attentional measures. *Journal of Experimental Psychology: General, 121,* 12–14.

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.

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

Posner, M. I., & Raichle, M. E. (Eds.). (1994). *Images of mind.* New York: Freeman.

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.

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

Warren, R. E. (1972). Stimulus encoding and memory. *Journal of Experimental Psychology, 94,* 90–100.

Warren, R. E. (1974). Association, directionality, and stimulus encoding. *Journal of Experimental Psychology, 102,* 151–158.



Received March 21, 2005
Revision received July 14, 2005
Accepted July 19, 2005 ■