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

Putting some feeling into it - the conceptual and empirical relationships between the classic and emotional Stroop tasks: A commentary on Algom, Chajut and Lev (2004)

Tim Dalgleish

Medical Research Council, Cognition and Brain sciences Unit, Cambridge

Abstract

Algom et al. (2004) presented a series of definitional, conceptual and empirical arguments in support of their conclusion that the classic and emotional Stroop effects are, in their words, "unrelated phenomena" (p.336), such that the term 'emotional Stroop effect' is a misnomer in reference to the relatively greater interference in ink color-naming of emotional versus neutral words. These are strong claims. In this commentary, I critically examine each component of Algom et al.'s case and argue that, in fact, none of these components represents compelling evidence in support of their eventual conclusions.

Since its initial publication (Stroop, 1935), the Stroop task has generated a proliferation of empirical research and theoretical debate (MacLeod, 1991). In Stroop's original study, the 'classic' experimental effect that has had the greatest influence on subsequent research was that latencies to name the ink colors of incongruent color words (e.g. naming the ink color of the word *red* printed in green ink), were longer than latencies to name colors of ink patches or shapes (MacLeod, 1991; Stroop, 1935). More recently, researchers have explored a variant of the classic paradigm - the so-called emotional Stroop task. In this task, the emotional Stroop effect refers to the fact that latencies to name the ink colors of emotional words (e.g. *death*) are typically longer than latencies to name the ink colors of matched neutral words (e.g. *table*)(Williams, Mathews, & MacLeod, 1996). The emotional Stroop task has proved to be an excellent tool for elucidating the nature of cognitive processing of emotional information in various clinical and sub-clinical populations (e.g. anxious individuals) where the content of the emotion words matches the particular concerns of the participant sample (Williams et al., 1996).

Algom, Chajut and Lev (2004) presented conceptual and empirical arguments as to why the emotional Stroop and the classic Stroop are, in their words, "unrelated phenomena" (p.336). This is undoubtedly a strong claim. In this commentary, I address this claim by challenging the theoretical and empirical arguments that Algom et al. have marshalled in support of it. There were four components to Algom et al.'s case. First, they presented a conceptual review of the classic and emotional Stroop effects and proposed alternative theoretical interpretations of the emotional effect. Second, they outlined five putative empirical diagnostics aimed at resolving these alternative theoretical interpretations. Third, they presented six empirical studies that speak to these diagnostics. Fourth, they drew these various issues together in a General Discussion where they also commented on the merits of theoretical attempts to model the classic and emotional effects. In this commentary, I address in turn each of these components of the Algom et al. case.

Algom et al.'s conceptual review

Algom et al. began their conceptual review by stating: "the Stroop effect is the difference in color-naming performance between congruent (e.g. the word *red* printed in red) and incongruent (*red*

in green) stimuli" (p.324). They went on to suggest that: "the defining feature of all Stroop stimuli is the existence of a *logical* relationship, compatibility or incompatibility, between their components" (p.324). If these two statements represented a uniformly held view, then Algom et al. could probably have rested their case straight away. By their definition, the emotional Stroop task *cannot* involve a Stroop effect as it does not involve a 'logical' relationship of compatibility or incompatibility between stimulus components. However, there are a number of problems with the Algom et al. definitions both of a Stroop effect and of a Stroop stimulus that suggest that the case should remain open.

First, Algom et al.'s definition of a Stroop effect is just one of several possible alternatives and is probably not even the most widespread. For example, in his original paper, Stroop (1935) proposed that: "the difference in the time for naming the colors in which the words are printed and the same colors printed in squares (or swastikas) is the measure of the interference of conflicting word stimuli upon naming colors" (p 659). Similarly, in his well-regarded review of research on the classic Stroop effect, MacLeod (1991), referring to Stroop's original task (Experiment 2) as "the standard Stroop Color-Word Test", speculated: "How should the standard color-word test be scored?.……the most prevalent [approach] is to calculate a difference score (interference card time - pure color card time) following Stroop (1935)." (p.166). It seems then that at least one widely-accepted (and indeed the original) definition of the Stroop effect is different to that proposed by Algom et al.

Second, as noted, Algom et al.'s definition of the Stroop effect involves subtraction of latencies in the congruent condition (e.g. the word *red* printed in red) from latencies in the incongruent condition. Although this particular definition may indeed be useful, when considered in isolation from other definitions (as in the Algom et al. article) it is potentially problematic. The difficulty here is that congruency of stimulus dimensions in the Stroop task often leads to *facilitated* color-naming, compared to standard control conditions such as rows of Xs (e.g., Glaser & Glaser, 1982). The Algom et al. definition of the Stroop effect therefore involves potentially subtracting a facilitation condition from an interference condition, with no genuine baseline such as color-naming patches of ink or rows of XXXX's. Consequently, employing Algom et al.'s definition would not allow independent empirical study of putative Stroop interference and facilitation effects and would therefore potentially be an obstacle to research progress rather than an advance. Algom et al. do not seem to have appreciated this

general point. Indeed, they stated that control conditions such as rows of Xs are "gratuitous in the classic task" (p.325). Interestingly, they added as a rider that "in the classic task incidentally, congruent stimuli are responded faster than XXXX stimuli" (p.325). In fact, this is not always the case (Sichel & Chandler, 1969) and such data provide a clear justification for why the definition of the Stroop effects needs to allow separate examination of putative interference and facilitation effects. Interestingly, the congruent case (e.g. naming the ink color of the word *red* printed in red) that forms part of the Algom et al. definition was not really investigated until some 30 years after Stroop's original publication (Dalrympole-Alford & Budayr, 1966), perhaps underlining the fact that it is not part of any uniformly-accepted definition of the Stroop effect.

The third difficulty concerns Algom et al.'s definition of a "Stroop stimulus"; namely, a stimulus for which the 'logical' relationship of congruency/incongruency between stimulus components holds. Again, this notion of a Stroop stimulus is not a uniformly-held view. For example, MacLeod (1991) in his aforementioned review included studies with priming conditions that consisted entirely of stimuli for which the Algom et al. 'logical' relationship is not present. Indeed, MacLeod (1991) included in his list of major empirical results regarding the Stroop effect the finding that "A color-unrelated word can be made to cause greater interference (or facilitation, or both) with color naming if its meaning is activated by a related word or phrase before the color-naming trial"(p. 203). Although one can see why Algom et al. might have wanted to underline the fact that certain Stroop stimuli are qualitatively different in that they involve a particular 'logical' relationship between the stimulus components (as such distinctions may well be important), it is not clear that these stimuli alone should be worthy of the Stroop label. Similarly, it simply does not follow that tasks involving these stimuli are "unrelated" to tasks involving stimuli where this particular 'logical' relationship does not hold.

Algom et al.'s proposed alternative conceptions of the emotional Stroop effect

In the next section of their paper, Algom et al. laid out two theoretical conceptualizations for the emotional delay (slowing of color-naming for emotional versus neutral words) on the emotional Stroop task. Algom et al. presented their first account as akin to that operating in the classic Stroop effect, such that the task-irrelevant emotional words are automatically processed, thus interfering with

color naming: "[emotional words] are processed in an obligatory fashion and intrude on other concurrent activities" (p.325).

According to Algom et al.'s second ("alternative") account , however: "the slowdown observed in naming the color of emotional words is unrelated to the Stroop effect" (p.325). Nevertheless, this alternative account relies on the fact that: " A dedicated system…captures threatening stimuli in an automatic fashion, prioritizes processing, thereby precipitating the temporary interruption of ongoing activity (e.g. color naming)" (p.325).

These two accounts seem to overlap considerably. Both seem to rely on the notion that emotional stimuli automatically recruit processing resources, thus leading to interference in other cognitive activities. The clearest apparent difference between the two accounts, based on this section of the Algom et al. paper, seems to be that the first account is in some way an 'attentional' effect whereby emotional words differentially recruit attentional resources away from color-naming, whereas the second account is an 'inhibition' account whereby the presence of a threatening stimulus leads to inhibition of all cognitive activity. However, this putative difference seems to slip through one's fingers when Algom et al. revisit their two accounts in the General Discussion section of their paper. Here, in reference to their second account, they stated "the generic slowdown observed does not necessarily imply that the underlying mechanism is that of inhibition. Resources can be allocated in particular ways such that threat-irrelevant stimuli receive fewer resources. This depletion in turn precipitates the disruption of pertinent behaviors" (p. 335). By my reading, once you drop the inhibition component of Algom et al.'s alternative account, this second account becomes virtually indistinguishable from the first.

Furthermore, later in their General Discussion, inspired by the ideas of Arne Ohman (e.g. Ohman, Flykt, & Esteves, 2001), Algom et al. suggested that the emotional words in the emotional Stroop task are processed by a 'preattentive attentional system' whereas classic Stroop stimuli are processed by a 'postattentive attentional system'. The former system is fast, being automatic and parallel, while the latter is slow, being deliberate and serial. However, the database on the classic and emotional Stroop tasks shows that such a straightforward differentiation of the two phenomena is not really valid. For example, both effects survive masking of the word (color or emotional) stimuli (e.g.

Daza, Ortells, & Fox, 2002; Mogg, Bradley, Williams, & Mathews, 1993), suggesting, by Ohman's criteria, the operation of preattentive effects for both paradigms. Furthermore, both effects seem to involve postattentive processes. For instance, separating the distracter words from the colors on both tasks leads to changes in the degree of interference (Fox, 1993; MacLeod, 1991).

In summary, Algom et al. have failed to present two clearly contrasting accounts of the emotional Stroop effect and indeed they have shifted their theoretical positions at different points in their article.

It is perhaps useful at this point to offer a somewhat different analysis of the putative critical differences between the classic and emotional Stroop tasks. Both tasks seem to involve automatic reading of the irrelevant words (MacLeod et al., 1991). This automatic reading confers some interference with ink color naming which is why the presence of *any* word generally leads to a slow-down in such color naming, relative to naming ink patches. However, in addition, the semantic content of the word leads to additional interference (e.g. Klein, 1963). In the classic Stroop, this additional interference is because the color words are semantically related to goals linked to the *ongoing task demands* (i.e., to color-naming) and, on the critical trials, it is because the color words are incongruent with the correct response required by the task. In the emotional Stroop, in contrast, the emotional words are related to goals reflecting *chronic concerns* about threat or the like (but are unrelated to the task demands to color-name). Furthermore, it seems likely that the recruitment of processing resources by the semantic analysis of threat-related words on the emotional Stroop task is more automatized (reflecting as it does chronic, survival-related processing) than the analogous situation in the classic task involving the processing of color words, which are only temporarily relevant to ongoing concerns due to the particular set of task demands (Logan, 1980, 1988).

Applying this type of analysis, one would predict that the presence of emotional words would lead to interference on a range of resource-dependent cognitive tasks because the emotional words (by virtue of their relevance to current concerns) would recruit processing resources in a relatively automatic fashion away from those tasks. One such task might be ink color-naming, i.e. the emotional Stroop paradigm. In contrast, one would not predict that the presence of color words would have any differential effects on cognitive task performance, except on tasks that involved the processing of

color-information. Given this, one empirical challenge is to identify the task parameters that set boundaries on any interference effects of emotional words. For example, in contrast to the emotional Stroop data, classifying numbers as odd or even, while trying to avoid unmasked distracter words, seems to be unaffected by whether the distracters are emotional or neutral (Fox, 1996).

Although in this section I have criticised the particulars of Algom et al.'s conceptual overhaul of the emotional Stroop, it is important to state that the process of elucidating the conceptual differences between the classic and emotional Stroop effect is without question a worthy exercise. However, with respect to Algom et al.'s overall mission, it is equally important to note that demonstrating that any two paradigms might require somewhat different theoretical conceptualizations is not sufficient grounds to argue that they are unrelated - the conceptual similarities between the two paradigms may remain considerable and outweigh any differences.

<center>Algom et al.'s five diagnostics</center>

In the next section of their paper, Algom et al. proposed five "diagnostic" to distinguish their two theoretical accounts (discussed in the previous section). Even if one suspends skepticism temporarily about whether the two accounts are sufficiently different, it is not completely clear that all of these diagnostics would be able to distinguish between them.

The first diagnostic refers to the fact that the classic Stroop effect is stronger when the to-be-ignored color word varies, relative to when it is held constant. Algom et al. proposed that examining whether this is also true for the emotional Stroop effect could distinguish their two theoretical positions. However, if the processing of threatening information is "obligatory" (p.325) as proposed in Algom et al.'s first account or is a result of "a dedicated system [that] captures threatening stimuli in an automatic fashion…." (p.325) as proposed in their second account, it may well be the case that varying the content of the to-be-ignored emotional words, versus using a constant word, would have little effect on the degree of interference due to the relative automaticity of any attentional (account 1) or inhibition/defense (account 2) effects. In fact, there is no clear theoretical prediction here from either account. In other words, diagnostic one is not a diagnostic.

The second diagnostic notes that a hallmark of the classic Stroop effect is the asymmetry of interference to color and word whereby, on the classic task, color-naming is interfered with by

incongruous words but color word reading is not similarly affect by the presence of incongruous print colors. This asymmetry is also revealed in the related phenomenon of Garner interference in which variation in the content of the to-be-ignored stimulus (as discussed with respect to diagnostic one) has a greater effect on color-naming than on word reading, relative to the baseline condition of invariant stimulus content. Algom et al. posed the question of whether such asymmetry in Garner interference levels would also be present in the emotional task.

Before presenting some conceptual arguments regarding this second diagnostic, I must confess to some confusion at to what exactly Algom et al. were seeking to test here in their empirical studies. I assume that the reading component of interference asymmetry on the emotional task, as assessed by Garner interference, involves comparison of a filtering task - the effect of *varying* ink color on reading a varying list of emotional words (relative to neutral) - in comparison to a baseline task - varying the words while keeping the ink color *constant*. This reading component is then compared to a color-naming component where the baseline task involves keeping the words *constant* (while varying the ink colors) and the filtering task involves *varying* the words (while varying the ink colors). There are two reasons for making this assumption. First, it is the most analogous situation to the classic Stroop effect. Second, it fits with Algom et al.'s statements in their experimental preview section (headed "Empirical study: Color-naming and reading in the emotional Stroop task") and in the Method section of Experiment 1 where their experimental design is most clearly articulated. For example: "For reading emotional words, in the baseline condition the participant read emotional words, changing from trial to trial, printed in an invariant color " (p.327).

The reason for my confusion is that other aspects of Algom et al.'s methodological description suggest a different design altogether. For example, Algom et al. also stated in the Method section of Experiment 1: "there were 50 stimuli (5 words x 5 ink colors x 2 repetitions) in each of the filtering tasks and 10 stimuli (1 word x 5 ink colours x 2 repetitions) in each of the baseline tasks" (p.328). This seems to imply that for all of the baseline tasks word identity was not in fact varied but held *constant* and that color was always varied.

Elsewhere, in the Results section of Experiment one, the implication seems to be that neither word nor color is varying in the baseline tasks: "In the baseline tasks, the participants read neutral

words *printed in an invariant color* faster by 36ms than they did emotional words *printed in an invariant color*…This slow down is an emotional Stroop effect with *a single color* for word reading" (p.329; my italics).

Clearly, my conceptual arguments (below) concerning the effects of varying ink color on word reading, in relation to this second diagnostic (and later diagnostics too) are only relevant if this is what Algom et al. were actually interested in. To this end, I have made the assumption based on their pre-experimental statements that this was indeed their experimental focus. What Algom et al. actually then examined in their experiments, however, remains somewhat unclear.

Given this assumption on my part, there are two conceptual points to make about diagnostic two. The first is that it makes little sense to look at the interference of print colors on the reading of emotional (or neutral) words as one would not expect any interference in this situation as the colors bear no relationship to either ongoing task demands (i.e. emotional/neutral word reading), or to chronic personal concerns. The second point is that, whether or not there is asymmetry in Garner interference does not distinguish between the two theoretical accounts that the authors have put forward as the effect relies on the same principles that constitute the first diagnostic - namely, task-irrelevant variation (see the discussion of diagnostic one). To spell this out, diagnostic two involves comparison of two conditions. The first condition - whether varying the ink colors of emotional words affects their reading times relative to neutral words - is misconceived as there is no reason to predict that ink color should influence emotional word reading and, critically therefore, the two theoretical accounts of Algom et al. do not make differential predictions on this issue. The second condition is essentially diagnostic one and, as I have argued, the two theoretical accounts do not make a differential prediction here either. Based on these two conceptual arguments, I would contend that diagnostic two is also not a diagnostic.

The third diagnostic that Algom et al. have proposed is related to their second diagnostic of interference asymmetry. This third diagnostic refers to the fact that alterations in the relative salience of words and print colors affect the extent and direction of any interference asymmetry in the classic Stroop effect. As with diagnostic two, this analysis makes little sense in the case of the emotional Stroop task where there is no theoretical rationale as to why print colors should interfere with

emotional word reading. For example, making the words less salient should have little influence on the differential speed of naming emotional versus neutral words as one would never expect print color to interfere more with emotional versus neutral words, irrespective of how salient the words are.

Again, most pertinently perhaps, whether or not alterations in salience can influence the magnitude of any interference asymmetry on the emotional Stroop task does not distinguish between the two theoretical accounts that Algom et al. have outlined.

The fourth diagnostic concerns the complementary relationship between word and color on the classic task such that "good performance with the words comes in tandem with poor performance with the colors and vice versa" (Algom et al, 2004; p.327). This is essentially a variation of the points in diagnostics two and three and is a function of the logical relationship between stimulus dimensions on the classic task. However, again, as there is no reason to suppose that print color should interfere with emotional (or neutral) word reading, this complementarity seems of little relevance to the emotional Stroop task. Furthermore, it is again difficult to see how it can distinguish between the two accounts that the authors have put forward.

What Algom et al. also seem to be saying here, however, is that their first theoretical account would predict that emotional words would lead to more color-naming interference than neutral words (the emotional Stroop effect) but no difference in *word-reading* times for emotional versus neutral words. In contrast, their second general 'inhibition'/'defense' account would predict an emotional word delay for *both* color-naming and reading. However, it is not clear that this is an issue of complementarity as there remains no reason to suppose that increasing the salience of the ink colors should affect differential reading speeds for emotional versus neutral words in any way at all. Rather, this claim seems to constitute a separate (sixth) diagnostic altogether; that is, whether the presence of emotional words interferes with word reading or not. This issue is of course both important and interesting. However, it cannot distinguish between Algom et al.'s two theoretical positions as it is of course possible that the presence of emotional words produces an emotional Stroop effect for one reason (perhaps even in line with Algom et al.'s first theoretical account), yet delays reading for another reason (perhaps in line with Algom et al.'s second theoretical account). This issue is revisited

in the final section of this commentary when I discuss individual differences on the emotional Stroop task**.**

Finally, for their fifth diagnostic, Algom et al. suggested that the extent to which any emotional delay extends in time to affect the processing of subsequent stimuli (regardless of their valence) can distinguish between their two accounts. They argued that their first account would not predict this but that their second account, which prescribes a general slowing in activity, would make such a prediction. It is not at all clear that these differential predictions follow from these two accounts. It is perfectly plausible that subsequent trials could be affected regardless to which account one subscribes; for example, because processing resources have been recruited differentially away from the dimension of print color (account one) or because there is a generic slowing/"freezing" of all ongoing activity (account two). Indeed, one could even apply this argument to the classic Stroop task by submitting that 'attending' to the incongruent color word interferes with concurrent color-naming but also means that on the next trial attentional resources are not optimally focussed on the dimension of ink color such that color-naming is impaired on that trial also. This is not a fanciful idea because there is a large literature on the phenomena of negative priming that relies on trial-to-trial carry over effects of Stroop processing, albeit of a different form. Interestingly, there exist unpublished data that speak to this issue (Mathews, personal communication). In this unpublished study, both incongruent classic color Stroop stimuli trials and emotional Stroop stimuli trials were always followed by a trial involving color-naming a row of Xs (thus differing from the McKenna & Sharma, 2004, study that focused only on the emotional Stroop). The critical comparisons were color-naming times for the row of Xs following the classic or emotional trials versus following a neutral word control trial. The data showed that *both* the emotional Stroop stimuli and the incongruent classic Stroop stimuli had carry-over effects in terms of slowing color-naming on the subsequent trial, relative to the control condition. In sum, the fifth diagnostic is also not a diagnostic.

Algom et al.'s six experiments

In the next section of their article Algom et al. described six experiments to address their five diagnostics, along with the additional 'sixth' diagnostic of whether there is an emotional delay with reading (and, in fact, lexical decision) as well as with color-naming. Algom et al. interpreted their data as supportive of the following conclusions: (a) the emotional Stroop effect is not sensitive to task-irrelevant variation; (b) The emotional effect does not display word-color reciprocity in interference; (c) the emotional effect is not affected by salience changes of word or color; (d) emotional words slow reading in the same way as they do color naming; (e) the emotional effect vanishes when emotional- and neutral-word trials are presented mixed together in the same block suggesting that it relies on carry-over effects to subsequent trials (in fact, the experiment that purportedly supports conclusion 'e' [Experiment 6] does not include an emotional Stroop condition, focusing instead entirely on word reading, and so this conclusion is not really justified); and, (f) the emotional delay is present when the task of reading is replaced by that of lexical decision.

I have already noted two things about the empirical data: first, they are largely non-diagnostic of the two theoretical positions that Algom et al. outlined in the introductory sections of their article. Second, despite this, the data are interesting in that they tell us something about the experimental parameters that differentiate the classic and emotional Stroop tasks. However, this last point should not be overstated. Providing evidence that the classic and emotional tasks are susceptible to different experimental manipulations is not the same as showing that they are unrelated phenomena. Rather, it simply helps to clarify any differences between the phenomena.

Having noted the potential importance of such experiments, it is also important to point out that the experiments taken together have a number of significant methodological limitations that temper enthusiasm for the findings. First, Algom et al. used their experimental data to build an argument about cognitive processing in the face of threat. However, they did not have sufficient control conditions in their studies to demonstrate that their effects were unequivocally threat-related. For instance, they may have achieved analogous results if they had used positive emotion words, indicating that the effects were perhaps due to emotionality, rather than threat, as has been found in a number of individual differences studies (e.g. Dalgleish, 1995, Experiment 2; Martin et al., 1991; Mogg & Marden, 1990), although this possibility seems less likely given previous results in the

literature with *unselected* samples (e.g. McKenna & Sharma, 2004). Or, they may have achieved

similar effects with words related to personal concerns or expertise, indicating that the effects were

perhaps due to self-relevance rather than emotionality or threat (for instance, Dalgleish, 1995, showed

analogous effects to the emotional Stroop effect using bird names with groups of ornithologists). Or,

they may have found analogous results with words coding loss, indicating perhaps that the effects

were related to negative emotion dimensions more generally, rather than specific to threat specifically.

These types of control condition are routinely used in emotional Stroop research to disentangle various

theoretical interpretations of this nature (see Williams et al., 1996).

      Secondly, some of the Algom et al. findings could be due to other artifactual differences

between their emotional (threat) words and the neutral words. For example, the emotional words used

by Algom et al. frequently referred to potential situations that could arise (e.g. crisis, disaster),

whereas the neutral words were essentially referring to objects or concepts. Alternatively, although

Algom et al. sought to rule out category effects in their data on the basis that their neutral words, in a

given experiment, were sourced from a single category (living environment, political science) to make

them analogous to the emotional words, this is not really sufficient to rule out category effects. What

is required is to show that the semantic relationship between the words is comparable across the

emotional words and the neutral words. Related to this is the need to show that the two sets of words

are equally representative members of their respective categories. In the Algom et al. studies, it is

perfectly plausible that the emotional words were more tightly semantically inter-related and so, when

presented in unmixed blocks, there was greater inter-stimulus priming leading to more 'interference'

relative to the neutral material - this would therefore be an effect unrelated to emotion. This would

also provide an account for why the emotional word effects disappear in the mixed block condition

(Experiment 6).

      Thirdly, although this is not strictly a methodological issue, these methodological concerns

need to be considered in the light of the extant literature on the emotional Stroop task -something

Algom et al. did not always do. For example, the emotional Stroop effect is not uniformly present in

healthy participant samples (e.g. Dalgleish, 1995 - Experiment two; Martin, Williams & Clark, 1991;

Thrasher, Dalgleish, & Yule, 1994). A similar issue pertains to the emotional word reading effects. In

contrast to the robust effect in Algom et al.'s data, other studies (e.g. Geller & Shaver, 1976; Martin et al., 1991, Experiments 1 and 2) have not found that reading times are slower for emotional versus neutral words in healthy subjects. Finally, Algom et al. argued, on the basis of their Experiment 6 with mixed and unmixed blocks, that the absence of an emotional effect in word reading in the mixed blocks is support for their second theoretical account that the emotional Stroop effect is due to some form of general 'inhibition' that affects the subsequent trials and the effect is therefore 'washed out' when emotional and neutral trials are mixed together. I have noted several problems with this interpretation above. However, in addition, there are numerous studies in the literature in which emotional and neutral words are presented in mixed lists and where the clinical groups show an emotional Stroop effect (e.g. Dalgleish, 1995; Foa, Ilai, McCarthy, Shoyer & Murdoch, 1993; Lavy, van den Hout & Arntz, 1993; McNally, Riemann, Louro, Lukach & Kim, 1992; Moradi, Taghavi, Neshat-Doost, Yule, & Dalgleish, 1999). If Algom et al.'s theoretical analysis is to hold, this should not be possible.

Overall, the fact that the Algom et al. data may not be completely representative of the picture painted by the extant literature in various different ways, emphasizes how important it is to demonstrate that their effects are not due to artifacts in methodology and/or materials that are specific to their studies before their results can be confidently interpreted at the level of theory (see Vasey, Dalgleish, & Silverman, 2003).

<center>Algom et al.'s General discussion section</center>

In their General Discussion section, Algom et al. build on their earlier conceptual and empirical examinations of the emotional Stroop effect and make a number of important claims about these paradigms in particular, and about scientific endeavor more generally. I have already considered in detail their claim that "The Stroop and the emotional Stroop effects are unrelated phenomena" (p. 336) and will not rehearse these arguments again. Rather, I would like to consider here two further issues that Algom et al. raise in this section of their article.

First, Algom et al. outlined their view that there are problems applying a popular connectionist model of the classic Stroop paradigm (Cohen, Dunbar, & McClelland, 1990) to the emotional Stroop

effect. On the back of this they then stated, in reference to the work of Williams et al. (1996): "we certainly disagree with the general quest 'to have models that can apply to both the original color-conflict situation and the emotional Stroop' (p.22)" (p. 336). It is understandable that if Algom et al. believe that the two phenomena are simply unrelated they might have little enthusiasm for models that speak to both. However, as I have suggested throughout this commentary, it is not clear that the two phenomena are unrelated and therefore such modeling endeavors are valid and justified until this issue is convincingly resolved.

Perhaps this point can be better illustrated by considering another domain - cognitive neuroscience. Here, studies have shown that the emotional Stroop and classic Stroop effects have some neural substrates *in common*, reflecting the related aspects of the two tasks, and other neural substrates that are *distinct*. For example, (Compton et al., 2003), in their fMRI study of the classic and emotional Stroop tasks, concluded: "dorsolateral frontal lobe activity was increased by both negative and incongruent color words, indicating a common system for maintaining an attentional set in the presence of salient distracters. In posterior regions of the brain, activity depended on the nature of the information to be ignored". (p. 81). However, perhaps a more convincing demonstration is the existence of connectionist models, derived directly from the work of Cohen et al. (1990) and utilizing their strength of processing pathways approach, that do in fact model both the classic and emotional Stroop tasks within the same networks (e.g. Matthews & Harley, 1996 and unpublished work by Battye - Battye, 2003; Battye, Barnard & Page, 2004).

On a more general note, although there is a clear place for theoretical and computational models that focus on particular cognitive or behavioral phenomena (such as the model of Cohen et al., 1991), a valid and worthy pursuit in psychology is surely to develop macro-theoretical models that can speak to a wider range of psychological effects. At the end of the day, it is such macro-theories that are likely to prove more valid when it comes to mapping functional level, psychological, models onto neuroscience models of human mentation to provide integrated theories of mind-brain (Barnard, May, Duke, & Duce, 2000; Dalgleish, 2004a, 2004b).

Secondly, throughout their article Algom et al. generally glossed over individual differences effects on the emotional Stroop task. This is somewhat surprising as it is in the study of such

differences that the emotional Stroop has most often been used. Only rarely (e.g. McKenna & Sharma, 2004) have emotional Stroop researchers been interested in the effect in the sorts of unselected samples used by Algom et al. (see Williams et al., 1996). Reading through the introductory and empirical sections of Algom et al.'s paper, the emotional Stroop researcher could be forgiven for concluding that it is frankly unnecessary to look at the effects of emotional words on color-naming - instead, one could simply ask participants to read the words aloud to produce analogous results. This would be a useful empirical outcome of the Algom et al. data series - if it were true. However, when Algom et al. did address the issue of individual differences in their General Discussion, it transpires that they did carry out an individual differences study involving patients with a diagnosis of PTSD. Here they found that the emotional Stroop effect was *greater* than the emotional delay with word reading, in contrast to all of the other results in their six experiments. On this basis, to their credit, Algom et al. stated that "it is arguable that the comparable performance across reading and color-naming is limited to healthy, unselected groups of participants" (p. 336). What Algom et al. did not do, however, is draw out the theoretical implications of these data. Essentially, such a result presents significant problems for any account that suggests that the emotional Stroop effect is simply the result of a generic "defense" or "inhibition" mechanism that acts similarly on all ongoing activity (Algom et al.'s second theoretical account) because the effects seem to be different in clinical groups for color-naming and word reading (not to mention that other cognitive tasks are unaffected by the presence of emotional words e.g. number categorisation, Fox, 1996). Some other theoretical account is therefore clearly required to map out why some tasks are unaffected by the presence of threat while others are not.

References

Barnard, P., May, J., Duke, D., & Duce, D. (2000). Systems, interactions and macrotheory. Transactions on computer human interaction, 7, 222-262.

Battye, T.G. (2003). Connectionist modelling of attention and anxiety. Unpublished Doctoral thesis: University of Cambridge.

Battye, T.G., Barnard, P. & Page, M. (2004). Connectionist models of the emotional Stroop effect. Manuscript submitted for publication.

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

Compton, R. J., Banich, M. T., Mohanty, A., Milham, M. P., Herrington, J., Miller, G. A., Scalf, P. E., Webb, A., & Heller, W. (2003). Paying attention to emotion: An fMRI investigation of cognitive and emotional Stroop tasks. Cognitive, Affective, and Behavioural Neuroscience, 3, 81-96.

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.

Dalgleish, T. (2004a). Cognitive approaches to Posttraumatic Stress Disorder (PTSD): The evolution of multi-representational theorizing. Psychological Bulletin, 130, 228-260.

Dalgleish, T. (2004b). The emotional brain. Nature Reviews Neuroscience, 5, 582-589.

Dalrympole-Alford, E. C., & Budayr, B. (1966). Examination of some aspects of the Stroop Color-Word Test. Perceptual and Motor Skills, 23, 1211-1214.

Daza, M. T., Ortells, J. J., & Fox, E. (2002). Perception without awareness: Further evidence from a Stroop priming task. Perception and Psychophysics, 64, 1316-1324.

Foa, E.B., Ilai, D., McCarthy, P.R., Shoyer, B. & Murdock, T. (1993). Information processing in obsessive-compulsive disorder. Cognitive Therapy and Research, 17, 173-189.

Fox, E. (1993). Attentional bias in anxiety: selective or not? Behavior Research and Therapy, 31, 487-493.

Fox, E. (1996). Selective processing of threatening words in anxiety: The role of awareness. Cognition and Emotion, 10, 449-480.

Geller, V., & Shaver, P. R. (1976). Cognitive consequences of self-awareness. Journal of Experimental Social Psychology, 12, 99-108.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. Journal of Experimental Psychology: Human Perception and Performance, 8, 875-894.

Klein, G. S. (1963). Semantic power measured through the interference of words with colour-naming. American Journal of Psychology, 77, 576-588.

Lavy, E.H., van den Hout, M., & Arntz, A. (1993). Attentional bias and spider phobia: Conceptual and clinical issues. Behaviour Research and Therapy, 31, 17-24.

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

Logan, G. D. (1988). Toward an instance theory of automatisation. Psychological Review, 95, 492-527.

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

Martin, M., Williams, R. M., & Clark, D. M. (1991). Does anxiety lead to selective processing of threat-related information? Behaviour Research and Therapy, 29(2), 147-160.

Matthews, G. & Harley, T.A. (1996). Connectionist models of emotional distress and attentional bias. Cognition and Emotion, 10, 561-600.

McKenna, F. P. & Sharma, D. (2004). Reversing the emotional Stroop effect reveals that it is not what it seems: The role of fast and slow components. Journal of Experimental Psychology: Learning, Memory and Cognition, 30, 382-392.

McNally, R.J., Riemann, B.C., Louro, C.E., Lukach, B.M. & Kim, E. (1992). Cognitive processing of emotional information in panic disorder. Behaviour Research and Therapy, 30, 143-149.

Mogg, K., Bradley, B. P., Williams, R., & Mathews, A. (1993). Subliminal processing of emotional information in anxiety and depression. Journal of Abnormal Psychology, 102(2), 304-311.

Mogg, K. & Marden, B. (1990). Processing of emotional information in anxious participants. British Journal of Clinical Psychology, 29, 227-229.

Moradi, A., Taghavi, R., Neshat-Doost, H., Yule, W., & Dalgleish, T. (1999). The performance of children and adolescents with PTSD on the Stroop colour naming task. Psychological Medicine, 29, 415-419.

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

Sichel, J. L., & Chandler, K. A. (1969). The Color-Word Interference Test: The Effects of varied Color-Word Combinations upon Verbal Response Latency. Journal of Psychology, 72, 219-231.

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

Thrasher, S. M., Dalgleish, T., & Yule, W. (1994). Information processing in Post-Traumatic Stress Disorder. Behaviour Research and Therapy, 32, 247-254.

Vasey, M. W., Dalgleish, T., & Silverman, W. (2003). Research on information processing factors in child and adolescent psychopathology: A critical commentary. Journal of Clinical Child and Adolescent Psychology, 32, 81-93.

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

Author note