1-1-2004

Attentional biases in eating disorders: a meta-analytic review of Stroop performance.

Keith S Dobson

David J A Dozois
ddozois@uwo.ca

Running Head: STROOP IN EATING DISORDERS

Attentional Biases in Eating Disorders:

A Meta-Analytic Review of Stroop Performance

Keith S. Dobson

University of Calgary

and

David J. A. Dozois

University of Western Ontario

Abstract

The Stroop task has been adapted from cognitive psychology, to be able to examine attentional biases in various forms of psychopathology, including the eating disorders. This paper reviews the research on the Stroop task in the eating disorders research area, in both descriptive and meta-analytic fashions. A total of 28 empirical studies are identified, which predominantly examine food and body/ weight stimuli in bulimic, anorexic, or dieting/ restricted food samples. It is concluded that there is evidence of an attentional bias in bulimia for a range of stimuli, but that the effect seems to be limited to body/ weight stimuli in anorexia. The evidence to date is that there is not an attentional bias in dieting samples. Limitations of the methodology employed in the extant literature include small sample sizes, unstandardized Stroop methodology, restricted gender, and a general lack of consideration of individual differences variables. Recommendations for future research area provided.

Attentional Biases in Eating Disorders:

A Meta-Analytic Review of Stroop Performance

The models that have developed in the area of the eating disorders (Fairburn & Brownell, 2002; Garner & Bemis, 1985; Vitousek & Orimoto, 1993) emphasize multiple possible causal factors in these problems. Among these models, one of the notable developments in psychopathology research has been the growth of investigations and models that emphasize cognitive processes and content (cf. Dobson & Kendall, 1993). In part spurred on by the general "cognitive revolution" in psychology, and in part by the success of cognitive-behavioral therapies, this emphasis has yielded a number of varied and rich technologies to examine psychological processes.

Among the range of cognitive factors have been implicated in the eating disorders are such issues as attitudes and beliefs about ideal body weight, body dissatisfaction, body image and perception, perfectionism have received considerable attention. For example, in a recent examination of the prediction of binge eating and purging (Byrne & McLean, 2002), it was found that overconcern with body weight and shape, coupled with the adoption of purgative behaviors, were predictive of binge eating and purging. Similarly, a recent review of the issue of body size dissatisfaction in anorexia nervosa (Skrzypek, Weheimer, & Remschmidt, 2001) found that although body size estimation is not impaired in this condition, it is the discrepancy between perceived body size and ideal size that is significantly associated with anorexia nervosa. Findings such as these highlight the role of negative attitudes and beliefs about food and body shape in the eating disorders, and suggest the possibility of disordered information processing in these conditions. For example, individuals who have eating disorders have undue concerns

about food or body image, it may be that these concerns could be demonstrated through a variety of methods that relate to different cognitive processes. One of the cognitive processes associated with eating disorders that has been the focus of considerable research, is that of selective attention. Consistent with the idea that individuals with eating disorders have negative beliefs about food and body shape, it has been suggested that these individuals are also more attentive and responsive to these relevant stimuli than non-eating disordered individuals. A body of research using a modification of the Stroop task has evolved, in which selective attentional processes have been examined. In this paper, we review the use of the Stroop task in examining eating disordered attentional processes, and examine in meta-analytic format the existing data. Conceptual and methodological issues in the extant literature, and directions for future investigations are provided.

The Stroop Task

The original or Classic Stroop task (Stroop, 1935) was developed as a means to study basic human attentional and informational processes. This task consisted of the presentation of colors printed on either neutral words or incongruent color words. Participants first named the color name of the stimuli, and then the color of the stimuli. This method was later revised to include color words (e.g., red, blue) that were printed either in the corresponding color (congruent condition) or other competing colors (incongruent condition). As these tasks are relatively easy, they form a basis for comparing the interference created by naming the color of the stimulus when the actual color is competing (e.g., the word "blue" written in green). Scores derived from the Stroop task consist of either the latency (typically in msecs) to name each stimulus, or the "interference effect", which is calculated as the average time taken to name the competing cards,

divided by the sum of the average time for the two other sets of stimuli. In some studies the number of correct color names are also recorded.

When the Stroop task was first explored (Stroop, 1935), the interference effect was discussed as an example of cognitive competition, and how such competition retards the production of correct color names. Replications of the Stroop interference effect have shown that it is reliable (Dyer, 1973; Franzen, Tishelman, Sharp, & Friedman, 1987; MacLeod, 1991). Both Stroop (1935) and MacLeod (1986, as cited in MacLeod, 1991), have reported that in normal controls the latency to name color-only stimuli is about 40% of that for naming incompatible color-word stimuli.

MacLeod's (1991) extensive review of the Stroop task identified "more than 700 Stroop-related articles in the literature" (p. 163), many of which focus on the perceptual, cognitive processing and potential neurological processes underlying the Stroop phenomenon. According to MacLeod (1991), a set of 18 reliable findings emerge in this literature which relate to the conditions which optimize or suppress interference effects, the effects of cues on Stroop performance, the effect of different localizations of stimuli, practice effects, and other performance aspects. Notably for the current review, MacLeod's (1991) authoritative review does not examine the effect of individual differences on Stroop interference, except to examine gender and age.

Despite the large amounts of research conducted with the Classic Stroop task on normal participants, it remains unclear whether the Classic Stroop effect is created by difficulties associated with the relative speed of processing color versus language (Klein, 1964), the automaticity of language (Logan, 1980; Posner & Snyder, 1975), the perceptual encoding of the

proper stimulus attribute (i.e., is a result of perceptual interference; Dyer, 1973; but see MacLeod, 1991 for evidence refuting this hypothesis), or interference related to the differential strength of the competing pathways being processed (see Cohen, Dunbar, & McClelland, 1990; Cohen, Servan-Schreiber, & McClelland, 1992; Logan, 1980; 1985; MacLeod, 1991; and Williams, Mathews, & MacLeod, 1996, for discussions of parallel distributed processing). Debate on the issue of which model best explains the Stroop interference effect continues and "the Stroop effect will continue to be a challenging phenomenon for cognitive psychologists to explain for many years to come." (MacLeod, 1991, p. 193).

Notwithstanding the controversy in cognitive psychology about the mechanism that accounts for the Stroop interference effect, the Stroop task has been utilized in psychopathology research for a considerable period of time. It has been suggested that the Stroop task is a valuable tool for examining cognitive processes in psychopathology (Segal, 1988; Williams et al., 1996), and that it affords an opportunity to compare disordered samples with normal controls in a task with relatively tight experimental control. Stroop research can be located in such diverse areas as anxiety disorders, depression, schizophrenia, eating disorders, substance abuse, and in studies that relate personality variables to cognitive interference. Recent years have witnessed a dramatic increase in the amount of Stroop research conducted in psychopathology (Williams et al., 1996).

In most of the psychopathology Stroop research, the focus is not often on general attentional interference, but is rather on the attentional bias that disordered participants exhibit relative to comparable control stimuli. As a result, the Stroop task has often been modified to use both Classic and "emotional" or disorder-appropriate stimuli (e.g., anxiety words for studies

of anxiety disordered samples, food words for studies with eating disordered samples). In effect, the model that has been tested is that appropriate emotional stimuli are more salient to individuals with a particular disorder, and that such stimuli will be selectively attended to (i.e. processed more efficiently) than for either individuals without that condition, or mismatched stimuli. The large number of studies that has been conducted using this paradigm affords a unique opportunity to make comparisons about the information-processing mechanisms among diverse clinical conditions. For example, Williams and his colleagues (1996) reviewed the Stroop task in depression and anxiety, and emphasized the mechanisms (cognitive, artifactual) underlying its interference effect. These authors demonstrated that attentional biases are found in depression and anxiety, and argued that the connectionist (parallel distributed processing) model (Cohen et al., 1990) is useful for understanding these biases.

The purpose of this study was to review, in both descriptive and meta-analytic format, the Stroop task in the eating disorders. There has been a recent "explosion" of research on the Stroop task in eating disordered samples, as fully 23 of the 26 studies discovered in the literature search for this review were published in or since the 1990s. Much of this work has been precipitated by cognitive formulations of eating disorders (Fairburn & Garner, 1988; Polivy & Herman, 1987), and in particular the awareness of the sensitivity of patients with eating disorders to food, as well as perceptions of body size and weight. The use of the cognitive measures has been increased, because of concern that simple questionnaires do not well measure attitudes associated with eating disorders (Bemis & Hollon, 1989); the use of the Stroop task in particular has been spurred by the development of sets of adjectives specifically designed for this population. Thus, the development of both a "Food Stroop" and a "Body Stroop" (sets of words

with food or body shape terms, respectively: Channon, Hemsley, & de Silva, 1988; Ben-Tovim, Walker, Fok, & Yap, 1989) has enabled researchers in this area to focus on specific, matched stimuli, rather than the Classic Stroop words. Rather than exploring the mechanisms thought to produce the Stroop interference effect (see Williams et al., 1996), this review assesses the Stroop findings and methodology in order to further our understanding of psychopathology and to enhance future clinical research. In the literature review that follows, a distinction will be made between the "Classic Stroop" and other "Emotional Stroop" methods, where the latter manipulate the content of the stimuli employed. Comparisons will be made between eating disordered and other groups, as well as between eating related stimuli and other, comparison stimuli, where such data exist.

The research in eating disorders has focused on three different problem areas: bulimia, anorexia nervosa, and dieting/restrained eaters. There appears to be a uniformity hypothesis in the literature, in that the rationale provided in different studies for the use of the Stroop task with different samples is analogous. However, as this distinction is consistently made in the empirical literature, we review the literature on dieting/ food restricted samples separately from those individuals who meet formal diagnostic criteria for an eating disorder, and further distinguish eating disordered samples into those who suffer from either anorexia nervosa or bulimia (American Psychiatric Association, 1994). Explicit tests of the uniformity hypothesis exist and are discussed below.

Method

The strategy employed for locating published research was to use the PsycInfo and Index Medicus indices, for research articles, chapters, and books for the period reviewed. Keywords

included any combination of "Stroop", and "eating disorder", "bulimia", "anorexia", "diet", or "restricted". Further, relevant published citations from all obtained research articles were pursued. We were able to obtain a copy of every citation that appeared to include data on the Stroop task in eating disorder. Thus, the following literature review and meta-analysis of the Stroop in psychopathology is, to the best of our knowledge, complete for the period 1935 to the end of May, 2001 (see Table 1 for the list of studies).

Insert Table 1 about here

Once the citations were assembled, both a qualitative and quantitative review of the results were undertaken. These results are reported in the sections that follow.

Results

Qualitative Comments on the Literature

As can be observed in Table 1, the predominant research method in this area has been the comparison of eating disordered samples with one or more types of control subjects. The general finding in this literature is that patients with bulimia demonstrate an interference effect with the Food Stroop task relative to normal controls (Black, Wilson, Labouvie, & Hefferman, 1997; Ben-Tovim, et al, 1989; Ben-Tovim & Walker, 1991; Cooper, Anastasiades, & Fairburn, 1992; Fairburn, Cooper, Cooper, McKenna, & Anastasiades, 1991; Jones-Chesters, Monsell, & Cooper, 1998; Perpiñá, Hemsley, Treasure, & de Silva, 1993; but see Waller & Ruddock, 1995, for a failure to replicate with a sexually abused control group). In those studies that also included and reported on the Classic Stroop task as a contrast, the typical finding is that

participants with bulimia subjects do not differ significantly from normal control participants on this task (Ben-Tovim et al., 1989; Ben-Tovim & Walker, 1991; Black, et al, 1997; Cooper et al., 1992; Perpiñá et al., 1993; but see Lovell, Williams & Hill, 1997 and Jones-Chesters, et al, 1998 for exceptions).

The results with anorexic patients largely mirror those for bulimic patients, with the exception that the effects for the Food Stroop are typically not as prominent, and in some cases prove elusive (Cooper & Fairburn, 1992; Ben-Tovim et al., 1989; Ben-Tovim & Walker, 1991; Perpiñá et al., 1993; Channon et al., 1988; Waller & Ruddock, 1995). In contrast, the analysis of differences between anorexic and normal participants using the Body Stroop usually has led to significant findings (Channon et al., 1988; Lovell, Williams & Hill, 1997; Perpiñá et al., 1993; Sackville, Schotte, Touyz, Griffiths, & Beumont, 1998). It appears that Stroop effects may be discernable with Body or Shape stimuli, but not Food stimuli in anorexia. Although this is a tentative conclusion, given the small number of studies conducted to date, if sustained it does suggest that future researchers should focus on the Body Stroop effects, or else that refinements of the Food Stroop task need to be made to find significant effects.

Some studies have included samples of dieters or restrained eaters. Cooper and Fairburn (1992), for example, had samples of bulimic, anorexic, and dieting participants. Their results showed a progressively shorter latency (i.e., less interference) with decreasing symptom severity, such that the difference between the dieting and control groups on the Food Stroop was not significant. Green and Rogers (1993), in another comparison between dieters and normal controls, showed non-significant differences for both the Food and Body Stroop tasks. Mahamedi and Heatherton (1993) examined groups of dieters and non-dieters, both before and

after ingesting food, but also failed to show group differences for either the Food or Body Stroop task. Although not a study of dieters, Smith and Miles (1986) also investigated the effect of eating on Stroop performance in randomly selected college women. In their study, one of the few in this literature to include male subjects (see also Green & McKenna, 1993, and Channon & Hayward, 1990, for exceptions), the effect of eating lunch led to nonsignificant changes in Stroop performance. Thus, it appears that the Food Stroop effect is confined to eating disordered samples.

One study that warrants brief comment in this literature is the Green and McKenna (1993) investigation of age and gender effects on the Food and Body Stroop. Their study used samples of equal number of 9, 11, and 14-year old males and females. Although male participants showed no effects, there was a significant age effect for females, such that latencies for both tasks were longer for the 14-year old group than for younger age groups. This finding was particularly clear for the Food Stroop task. The authors interpreted these findings to be "consistent with the ideas of worries about body shape, physical attractiveness, and dieting being related to the onset of adolescence" (Green & McKenna, 1993, p. 396).

Another recent trend in studies that examine the Stroop task in eating disordered samples is to examine the changes in performance associated with recovery from a disorder, or treatment. Cooper and Fairburn (1994) showed that bulimic individuals showed less interference on combined eating, shape and weight words after treatment than previously. In a detailed analysis of treatment effects, Carter, Bulik, McIntosh and Joyce (2000) failed to find a significant group by time interaction between bulimic women before and after treatment, as a function of word type (body/shape versus control words). Thus, although color-naming latencies changed as a

function of treatment, this effect was not specific to the targeted context (see also Black et al., 1997). It therefore remains an open question as to whether or not the Emotional Stroop is sensitive to treatment in eating disorder populations, and if so, whether or not this effect is specific to any particular stimulus content.

In summary, the literature related to eating disorders and dieting shows that effects on the Stroop Task can be discerned, particularly with the Food Stroop stimuli. The magnitude of group differences reveals that whereas bulimic participants are reliably differentiated from normal controls, anorexic participants are not as consistently different, and nonclinical participants with dieting concerns typically do not perform differently on the Body Stroop task than non-dieting controls. Whether the decreasing effects from bulimia to dieting is an effect specific to eating disorders, or is a simple effect of lower symptomatology in the dieting group cannot be deduced from the existing literature. Future research could examine the effects in different subject groups as a function of numbers of symptoms, rather than diagnosis, as a way of addressing this issue.

The differences observed between the Food and Body Stroop tasks are worthy of comment. Whereas the Body Stroop task has been able to differentiate clinical participants from control groups in a number of studies, the effects with the Food Stroop Task are inconsistent enough across studies to warrant concern about the task. Future researchers are encouraged to focus on the Body Stroop method, or else to pursue refinements to the Food Stroop in an effort to discover why effects cannot be consistently found in Anorexic samples.

In a group of disorders that is dominated by females (Maxmen & Ward, 1995; Wilson, Hefferman, & Black, 1996), it is not surprising that the majority of studies in this review have

employed female subjects. On the other hand, only 3 of the 26 studies reviewed included male subjects at all, and only one of these (Green & McKenna, 1993) specifically examined gender differences on the Food and Body Stroop tasks. Further research is encouraged in this domain, although it is worth noting that the Body Stroop task may need particular refinements for male samples, as the idealized body image, and the adjectives associated with this image, may need to be made gender-specific to find latency or interference effects.

Meta-analytic Results

Meta-analysis is a statistical technique for combining the results of several studies. In doing so, more stable estimates of effects can be ascertained, and even the reliability of effects reported in the literature can be assessed. Statistically, meta-analysis employs an effect size (ES), which is computed as the difference between the mean of a criterion group and a comparison group or condition, divided by an estimate of variability. Two main versions of effects sizes are seen in the literature. Cohen's (1977) d employs the standard deviation of the comparison group or condition as the estimate of variability, and effects sizes are thus essentially a measure of the deviance of the criterion group from the comparison group, expressed in the statistical equivalent of standardized scores. Cohen suggested using 0.20, 0.50 and 0.80 as the cut-off criteria for the identification of small, medium and large effect sizes, respectively (see also Landman & Dawes, 1982). The major alternative to d is Glass' (1976) g, which employs a weighted estimate of the population standard deviation, derived from both groups. In general, Glass' g is recommended in meta-analysis, since it uses a more stable estimate of population variability (especially important when studies employ small sample sizes), and because it is often the case that the homogeneity of variance across studies employed in meta-analysis cannot be

assumed (Kazdin & Bass, 1989). Glass' g is employed here.

A given study can potentially generate many effect sizes. In the current study, we were primarily interested, though, in several key comparisons. Thus, we examined the Stroop effects for the comparisons between all three types of eating disordered samples (bulimic, anorexic, dieting/ restricted eating) versus control samples where these existed, for each of the three main types of Stroop content, which are identified here as Food (food related words), Body/ Shape (words that reflected enlarged body size, or shape), or the Classic Stroop. Comparisons that used healthy foods or eating, or thin body shape are not reported here, as the number of such studies is very limited, and these comparisons have been of less interest in the literature.

The other set of comparisons that are reported in this review are within each group, comparing across different types of stimuli. Although in principle experimenters using the Stroop task can either examine latency scores (the time it takes to color name words of different colors) or interference scores (color naming times for emotional content, corrected for baseline or control content), the vast majority of studies in this area only report latency results. Thus, in an effort to examine the relative importance of different types of stimuli, we report here the comparative effect sizes for various Stoop stimulus contents, within each eating disordered sample (bulimic, anorexic, dieting/ eating restricted)[1]. Also, because interference scores are reported so infrequently in this literature (only 11 of 129 computed effect sizes included interference data), these results are not reported here in table format.

The report of effect sizes in the current study includes several components. The effect sizes g (Glass, 1976) and number of studies, n, are reported for all between group comparisons. The method adopted here, which was to only calculates effect sizes for those studies that

provided the mean and standard deviations for each comparison, is the most straightforward one, as it does not use any estimated data points, nor does it use tests of significance as an indirect source of effect sizes. At the same time, this method leads to the smallest number of studies possible being included in the meta-analysis, as the included studies must meet all any inclusion criteria, and provide all necessary data (e.g., Bangert-Drowns, 1992; Rosenthal, 1995). Also reported is the 95% confidence interval for each ES, the test for the homogeneity of the ES (Q; if this test is significant, it suggests that the ES is not reliably drawn from a stable population), and the fail safe n for each comparison are provided. The fail safe n is used to address the issue of the "file drawer" problem (Rosenthal, 1979), and provides an estimate of the number of unpublished studies that would have to exist to conclude that the effect size is not significant. The larger the fail safe n, the more robust a given ES is to the prospect of negative results (Rosenthal, 1979; Strube & Hartman, 1983). All calculations were conducted with the program written by Schwarzer (1989).

Insert Table 2 and Figure 1 about here

Table 2 reports the effect sizes for all between group comparisons for Food, Body/ Shape and Classic Stroop stimuli (see also Figure 1). Several results are noteworthy. First, it is clear that only the results comparing bulimic to control participants consistently yield moderate effect sizes. All three such comparisons were in the moderate range, and the fail safe n ranged from 8 to 18, indicating a fair degree of resiliency to negative results. That said, all three comparisons involving bulimic subjects had statistically significant tests of homogeneity, indicating that the

estimates of the overall effect sizes cannot be seen as reliably drawn from a single population of effect sizes. Thus, further research is warranted in this area to determine how stable these results are. Further, the fact that moderate effect sizes were noted for all three comparisons should be a cause for concern, as it suggests that bulimic individuals may have a general deficit in color naming, rather than a specific content problem. Research that examines the specificity of color naming latencies is needed, perhaps employing general negative or other emotional Stroop words, to examine the possibility of a general deficit in bulimia.

The results for participants with anorexia were not as strong as for those individuals with bulimia. Only the comparison with controls that involved body/ weight Stroop stimuli was in the moderate effect size range, whereas both the Food and Classic Stroop effects were minimal. The fact that anorexic participants display a consistent attentional bias to body/ weight stimuli, but not food or Classic Stroop stimuli, suggests a specific attentional bias that is discussed further below. Again, all three comparisons involving anorexic participants failed the test of homogeneity, suggesting that the existing estimates of effect sizes are not reliable.

Finally as regards between-group comparisons, the comparisons between dieters or restrained eating subjects and controls were consistently in the minimal range of effect sizes. The only possible exception to this negative conclusions was the comparison involving food stimuli, which had an effect size of .39, which is bordering on the moderate effect size range. However, the tests of homogeneity were not significant for both the food and body/ weight Stroop, which suggests that the estimates of effect sizes for both of these comparisons are reliable. As such, these results imply that more research using the Food Stroop with dieters or people who are restraining their eating is not likely to increase (or decrease) the observed effect

size.

Insert Table 3 about here

Table 3 presents the within group comparisons for the three different types of groups examined in the current study, comparing the results of food versus body/ shape Stroop stimuli, food versus control words, and body/ shape Stroop stimuli versus control words[2]. All comparisons generate the same conclusion, which is that the type of stimuli employed is not consistently related to attentional biases in different eating disordered populations. In particular, the test for the homogeneity of effect sizes for the comparisons between food and body/ weight stimuli were all non-significant, which indicates that these comparisons (which are all in the trivial range) are reliable. The failure of the test of homogeneity for the other comparisons, coupled with fairly large confidence intervals, does suggest that these estimates of effect sizes warrant further investigation.

Discussion

The results from the eating disorders research suggest that whereas bulimic participants consistently show attentional biases on the Stroop Task across a range of stimuli, the results are limited to the area of body/ weight stimuli for anorexic subjects, and are trivial to modest for dieting/ restricted food intake subjects. These results should provide cause for concern about using the Stroop task a measure of attentional bias in non-clinical samples. They further suggest, however, that there may be differences between patients with bulimia and anorexia as a function of whether researchers use the Food or Body Stroop task. This possibility has not yet been tested

adequately. The specificity of an attentional bias in bulimia to Food and Body/Shape words also needs to be further examined, as the current results suggest that individuals with bulimia may have a generalized deficit in attentional deployment.

The analysis of the Stroop task with anorexic samples suggests that, whereas these individuals do have an attentional bias associate with Body/ Shape stimuli, this effect does not generalize to Food stimuli. If reliable, this conclusion may help to further specify that the cognitive concern or mechanism in anorexia is specifically related to body size, as would be consistent with the diagnostic features of this disorder (American Psychiatric Association, 1994). It is possible that the Stroop effects found for Body/ Shape stimuli in anorexia reflect this attentional bias, in which anorexic individuals selectively attend to threat cues (cf. Cassiday, McNally, & Zeitlin, 1992; Foa, Feske, Murdock, Kozak, & McCarthy, 1991; Kaspi, McNally & Amir, 1995; Thrasher, Dalgleish, & Yule, 1994). Although speculative, it is feasible that individuals with anorexia are more vigilant to body/ shape than to food-related stimuli because body stimuli represent the object to be avoided, and thus represents the greater psychological threat. Greater attentional resources may be allocated toward the more immediate threat cues.

The specific processes and artifactual concerns involved in the Stroop effect have already been reviewed extensively elsewhere (see Cohen et al., 1990; Cohen, at al, 1992; Dyer, 1973; Izawa & Silver, 1988; MacLeod, 1991; and Williams et al., 1996). As these reviews point out, the currently accepted account of the Stroop effect is the Parallel Distributed Process framework, which was advanced by Cohen et al. (1990). According to this conceptualization, information is processed via several interactions of information-processing units that are connected in an associated network (see Bower, 1981, for an earlier related view). Interference is believed to

occur contingent upon (1) the relative strength of the connections which interact in the two competing pathways (one pathway for naming color, the other pathway for naming the word) and (2) the role that attention plays in the processing of these pathways (i.e., attention modulates reaction time by increasing activity in the color pathway and reducing activity in the otherwise pre-potent word pathway). It is conceivable that although eating disorders share similarly over-activated content-specific word representations (e.g., schematic structure), specific attentional deficits may be present in anorexia. Further research assessing both attentional and organizational processes is required to substantiate or refute this claim. For example, it may be that other tasks such as the Self-Referential Encoding Task, that also assess attentional processes may show similar effects, as has been observed in other disorders (Dozois & Dobson, 2001). In This regard, an expansion of research in eating disorders to other cognitive tasks is recommended.

Finally, the Stroop task has been used to evaluate treatment outcome in bulimia (Cooper & Fairburn, 1994; Lovell, et al, 1997). In general, the results indicate that therapy is associated with a decrease in attentional bias on the modified Stroop tasks. Although the number of studies examining this hypothesis is small to date, these results suggest that the Stroop task reflects malleable, and dynamic representations of self. Clinical improvement reflected on the Stroop task suggests either that cognitive structures do indeed change or that the Stroop task measures information processing (functional properties) rather than cognitive organization (structural properties). Further examination of these competing hypotheses would help to clarify the processes associated with Stroop task performance.

One of the general observations that can be made about the current review is that the

majority of research using the Stroop Task in eating disordered populations is cross-sectional. With the exception of very limited amounts of research using this task as an index of change in treatment, we are aware of no study that has yet examined attentional bias in a longitudinal or prospective research design. A recent review of the development of weight and shape concerns, however, suggests that these issues follow a developmental trajectory that might actually predate the typical onset of disordered eating (Gowers & Shore, 2001). Given that the current review found both anorexia nervosa and bulimia significantly related to body/ shape attention, this developmental process may help to explain this common Stroop Task finding. Further, if such a developmental pattern for food stimuli can be found to selectively relate to bulimic eating, this line of research may assist with the models of the etiology of eating disorders. Further, as Gowers and Shore (2001) note, if certain concerns follow a predictable developmental pathway, it may be possible to identify individuals with such concerns even before the onset of a fully developed eating disorder, and to develop preventive strategies for such at-risk individuals. Research employing the Stroop Task as a marker or precursor to other eating disorders could help to evaluate its potential as an early warning sign of later eating disorders.

One of the notable deficiencies in the Stroop Task literature in eating disorders is a lack of concurrent validation, both regards to other tasks and other clinical samples. For example, if Stroop Task performance is reliably found in both Bulimic and Anorexic individuals, as is suggested by the current review, it is reasonable to suppose that other cognitive processes should also be found using related procedures. It remains for future research to examine the extent to which Stroop Task performance is significantly correlated with other self-report questionnaires (cf., Rieder, & Ruderman, 2001) or cognitive tasks (cf., Skrzypek, et al, 2001).

In like measure to the lack of concurrent validity research on the Stroop Task in eating disorders, it is notable that very little literature has been done to evaluate the utility of emotional Stroop Task stimuli across various disorders. While it is not likely in this regard that Food and Body/Shape stimuli would demonstrate significant effects with other clinical disorders, it is possible that emotional Stroop stimuli developed in such areas as anxiety and depression may yield significant effects with eating disordered samples. As has been noted, the eating disorders not uncommonly have significant levels of anxiety and/ or depression associated with them (American Psychiatric Association, 1994). Thus, it is possible that Stroop Task effects with anxiety and/ or depression related stimuli might be found in eating disordered samples. If so, this result would help to explain unique and overlapping aspects of other clinical disorders and the eating disorders.

Methodological Issues and Research Recommendations with the Stroop Task

Several methodological issues generic to the Stroop task in psychopathology may be delineated from this literature review. One limiting factor from a meta-analytic perspective is that some researchers either do not present the means and standard deviations in their articles, which precludes the calculation of effect sizes, or their studies are correlational in nature, thus preventing appropriate comparisons with control conditions. Researchers are encouraged to report the means and standard deviations of their results.

Another common reporting tendency has been for researchers to use raw latency scores rather than interference scores as dependent measures (see Table 1). Notwithstanding the fact that reaction times are important to compute and present in research as raw indices, the most unbiased estimate of cognitive interference entails the comparison of an individual's reaction

time to target words relative to baseline scores. Without this baseline measure it is less clear what a person's latency score means, and interference becomes confounded with extraneous variables such as general cognitive deficit, fatigue, and motivation. Although there were an insufficient number of studies that reported interference effects to replicate Tables 2 and 3 with comparable reports, it is instructive to note that those calculations we did for interference effects consistently showed stronger effects than reported above for latency scores. For example, the average effect size for interference scores between bulimic and control groups on the Food Stroop was .75, based on 3 studies. The effect size for the comparison between bulimic and control participants, based on two studies, was .87. The comparable results comparing individuals with anorexia to nonpsychiatric controls, each based on a single study, however, was .73 for the Food Stroop and 1.14 for Body/ Weight Stroop. Thus, while the lack of existing literature severely limits any conclusions that can be drawn, it does appear that interference effects may generate stronger group comparisons than has been reported to date for latency scores. Researchers are encouraged to report both latency and interference in future publications.

The word lists employed in the research reviewed present additional methodological issues salient to understanding the Stroop in psychopathology. Many investigators have generated their own word lists, of various lengths, and with various properties. Few of these word lists have been standardized (although some investigators do control for word length), even though issues such as the emotional salience of the word lists has been shown to affect performance on the Stroop task (Martin, Williams, & Clark, 1991; Riemann & McNally, 1995). The Classic Stroop word lists, matched lists, number of words, and number of categories used to

explore cognitive interference in psychopathology have also varied across studies. Some studies employ word lists written on cards, while others employ computerized presentation of stimuli. Although variations to the Classic Stroop task have not typically led to significant differences (see MacLeod, 1991), the impact of diverse methodologies on the investigation of the modified Stroop task in eating disorders research remains to be determined.

It may prove useful for researchers to use more specific (e.g., theoretically relevant) stimulus sets in their designs. One possibility is to have participants choose words that are relevant to their condition prior to administering the Stroop task. This strategy has been used in a few studies on the Stroop in depression (see Segal, Gemar, Truchon, Guirguis, & Horowitz, 1995) and could easily be adapted for research in eating disorders. Of course, the danger in such an approach is that generalizability becomes limited. This recommendation must also be tempered by the fact that the effect sizes for the content comparisons were smaller than for the comparisons between groups. Another option is for researchers to gather normative data on words that are particularly salient to, and which maximally distinguish between, different diagnostic groups. These words could also be controlled for their emotional salience, imagability, and familiarity. Consideration of other nonlinguistic strategies (e.g., pictorial images; Bradley, Mogg, & Millar, 2000; Lavy & van den Hout, 1993; Mogg, Millar, & Bradley, 2000) may also contribute to the cognitive literature.

Sample selection issues are also important in Stroop research. Although groups are usually established either on the basis of diagnosis or predetermined cut-off scores on psychometric instruments, the sample make-up is often not thoroughly described. If clinicians were ever to use the Stroop task in clinical practice to evaluate treatment or as a predictive index,

the hiatus between research and practice needs to be minimized. One important way to make Stroop research more relevant to practice would be for researchers to more clearly define their samples. In the area of eating disorders, for example, it may be that diagnostic category is not as important as the number or severity of symptoms. Given that the general pattern observed in this study was for bulimic individuals to show stronger effects than for anorexic participants, which were in turn more strong than for dieters/ restricted eating subjects, it may be that different attentional effects are the result of symptom severity; if so, this difference provides another illustration of the need for thorough sample description.

Another sample selection issue pertains to sample size. The sample sizes were low to moderate in a number of the studies reviewed. Thus, the power to detect true differences between groups may also be low, introducing Type II error into the design (see Kazdin, 1994; Kazdin, Siegel, & Bass, 1990).

Gender emerges as a major limiting sample factor in the literature on the Stroop task and eating disorders. The majority of the studies employed have only used female participants, and only a few investigations have incorporated gender as a criterion variable in their designs. Although MacLeod (1991) argued that gender was not critical on the Classic Stroop task, it may be an important variable to examine in Stroop tasks in the eating disorders area, in order to determine the extent to which males and females differ in the organization and processing of schema- and disorder-specific content. Sensitivity to possible gender differences is needed in future investigations of the Stroop.

Researchers may also wish to capitalize on (or develop) other information-processing measures adapted from experimental cognitive psychology to assess cognitive differences in

psychopathology (e.g., the Self-Referent Encoding Task; see Dozois & Dobson, 2001; Dobson & Shaw, 1987). For example, bulimic patients may be especially reactive to the Stroop task, but less so to other procedures. Clinical research will likely continue to be advanced with the use of other experimental cognitive methodologies.

Conclusion

Prior to the adoption of methodologies derived from experimental cognitive psychology, clinical investigators were unable to directly measure the cognitive processes or structures involved in maladaptive functioning. Instead, indirect inferences were often drawn, from self-report data, about the logical underlying cognitive operations associated with psychopathology (Ingram & Kendall, 1986). As Segal (1988) aptly stated, "the strategy of relying on [patient] self-reports to validate a construct whose operation is intended to explain these self-reports becomes increasingly circular unless additional referents can be provided to demonstrate schematic processing" (p. 147). The Stroop methodology has provided one of these additional referents, and has permitted researchers to analyze cognitive processing in a way that is less transparent than self-report (Segal, 1988).

The current review has demonstrated that performance on the emotional Stroop Task is related to eating disorders. In particular, it appears that although selective attention to either Food or Body/ Shape stimuli cannot be reliably detected in dieting or eating restrictive samples, sensitivity to Body/ Shape stimuli can be seen in anorexic samples, and selective attention to both types of stimuli can be seen in Bulimic subjects. These results may indicate stimulus specificity in different disorders, or may reflect increasing overall levels of disturbance, and thus broader areas of Stroop Task performance, in these various groups. Although the results

obtained in the current review help to explicate the relation of the emotional Stroop Task in eating disordered groups, there are a number of research directions that have yet to be explored. Principle among these are the longitudinal development of attentional biases in the eating disorders, the covariation of the emotional Stroop Task with other measures of cognitive processing, and the specificity of emotional Stroop stimuli across various clinical conditions.

In summary, the modified Stroop task has demonstrated utility in psychopathology. The areas of theoretical and practical controversy highlighted in this review signify fertile ground for continued investigation and innovation. We anticipate, and offer above numerous promising areas for the Stroop task in both research and practice, and echo MacLeod's (1991) enthusiasm about "The progress that will be examined in the subsequent review of the Stroop literature some time early in the next millennium" (p. 193).

Author Notes

Dr. Dobson is at the Department of Psychology, University of Calgary, Calgary, Alberta, Canada T2N 1N4, whereas Dr. Dozois is at the Department of Psychology, University of Western Ontario, London, Ontario, Canada N6A 5C2. Please address requests for reprints to the first author. The support of the Alberta Heritage Foundation for Medical Research (Dr. Dobson) and the Ontario Mental Health Foundation (Dr. Dozois) are noted.

References

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

Bangert-Drowns, R. L. (1992). Review of developments in meta-analytic method. In A. E. Kazdin (Ed.). Methodological issues and strategies in clinical research (pp. 439-467). Washington, DC: American Psychological Association.

Bemis, K., & Hollon, S. D. (1989). The investigation of schematic content and processing in eating disorders. Cognitive Therapy and Research,

Ben-Tovim, D. I., Walker, M. K., Fok, D., & Yap, E. (1989). An adaptation of the Stroop test for measuring shape and food concerns in eating disorders: A quantitative measure of psychopathology?. International Journal of Eating Disorders, 8, 681-687.

Ben-Tovim, D. I., & Walker, M. K. (1991). Further evidence for the Stroop test as a quantitative measure of psychopathology in eating disorders. International Journal of Eating Disorders, 10, 609-613.

Black, C. M. D., Wilson, G. T., Labouvie, E., & Hefferman, K. (1997). Selective processing of eating disorder relevant stimuli: Does the Stroop test provide an objective measure of Bulimia Nervosa? International Journal of Eating Disorders, 22, 329- 333.

Bower, G. H. (1981). Mood and memory. American Psychologist, 36, 129-148.

Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. Cognition and Emotion, 14, 789_808.

Byrne, S. M., & McLean, N. J. (2002). The cognitive-behavioral model of bulimia nervosa: A direct evaluation. International Journal of Eating Disorders, 31, 17- 31.

Carter, F. A., Bulik, C. M., McIntosh, V. V., & Joyce, P. R. (2000). Changes on the Stroop test following treatment: Relation to word type, treatment condition, and treatment outcome among women with Bulimia Nervosa. International Journal of Eating Disorders, 28, 349- 355.

Cassiday, K. L., McNally, R. J., & Zeitlin, S. B. (1992). Cognitive processing of trauma cues in rape victims with post-traumatic stress disorder. Cognitive Therapy and Research, 16, 283-295.

Channon, S., & Hayward, A. (1990). The effect of short-term fasting on processing of food cues in normal subjects. International Journal of Eating Disorders, 9, 447-452.

Channon, S., Hemsley, D., & deSilva, P. (1988). Selective processing of food words in anorexia nervosa. British Journal of Clinical Psychology, 27, 259-260.

Cohen, J. D. (1977). Statistical power analysis for the behavioral sciences (2nd ed.). New York: Academic Press.

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

Cohen, J. D., Servan-Schreiber, D., & McClelland, J. L. (1992). A parallel distributed processing approach to automaticity. American Journal of Psychology, 105, 239-269.

Cooper, M. J., Anastasiades, P., & Fairburn, C. G. (1992). Selective processing of eating-, shape-, and weight-related words in persons with Bulimia Nervosa. Journal of Abnormal Psychology, 101, 352- 355.

Cooper, M. J., & Fairburn, C. G. (1992). Selective processing of eating, weight and shape

related words in patients with eating disorders and dieters. British Journal of Clinical Psychology, 31, 363-365.

Cooper, M. J., & Fairburn, C. G. (1993). Demographic and clinical correlates of selective information processing in patients with bulemia nervosa. International Journal of Eating Disorders, 13, 109-116.

Cooper, M. J., & Fairburn, C. G. (1994). Changes in selective information processing with three psychological treatments for bulimia nervosa. British Journal of Clinical Psychology, 33, 353-356.

Cooper, M., & Todd, G. (1997). Selective processing of three types of stimuli in eating disorders. British Journal of Clinical Psychology, 36, 279- 281.

Dobson, K. S., & Kendall, P. C. (Eds.) (1993). Psychopathology and Cognition. New York: Academic Press.

Dobson, K. S., & Shaw, B. F. (1987). Specificity and stability of self-referent encoding in clinical depression. Journal of Abnormal Psychology, 96, 34-40.

Dozois, D. J. A., & Dobson, K.S. (2001). Information processing and cognitive organization in unipolar depression: Specificity and comorbidity issues. Journal of Abnormal Psychology, 110, 236- 246.

Dyer, F. N. (1973). The Stroop phenomenon and its use in the study of perceptual, cognitive, and response processes. Memory and Cognition, 1, 106-120.

Fairburn, C. G., & Brownell, K. D. (Eds.) (2002). Eating Disorders and Obesity: A Comprehensive Handbook (2nd Edition). New York: Guilford Press.

Fairburn, C. G., Cooper, P. J., Cooper, M. J., McKenna, F. P., & Anastasiades, P. (1991).

Selective information processing in bulemia nervosa. International Journal of Eating Disorders, 10, 415-422.

Fairburn, C. G., & Garner, D. M. (1988). Diagnostic criteria for anorexia nervosa and bulimia nervosa: The importance of attitudes to shape and weight. In D. M. Garner and P. E. Garfinkel (Eds.) Diagnostic issues in anorexia nervosa and bulimia nervosa. (pp. 36- 55). New York: Brunner/ Mazel.

Foa, E. B., Feske, U., Murdock, T. B., Kozak, M. J., & McCarthy, P. R. (1991). Processing of threat-related information in rape victims. Journal of Abnormal Psychology, 100, 156-162.

Franzen, M. D., Tishelman, A. C., Sharp, B. H., & Friedman, A. G. (1987). An investigation of the test-retest reliability of the Stroop color-word test across two intervals. Archives of Clinical Neuropsychology, 2, 265-272.

Garner, D. M., & Bemis, K. M. (1985). Cognitive therapy for anorexia nervosa. In D. M. Garner and P. E. Garfinkel (Eds.) Handbook of psychotherapy for anorexia nervosa and bulimia. (pp. 107- 146). New York: Guilford Press.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 10, 3 -8.

Gowers, S. G., & Shore, A. (2001). Development of weight and shape concerns in the aetiology of eating disorders. British Journal of Psychiatry, 179, 236- 242.

Green, M., Corr, P., & De Silva, L. (1999). Impaired color naming of body shape-related words in Anorexia Nervosa: Affective valence or associative priming? Cognitive Therapy and Research, 23, 413- 422.

Green, M. W., & McKenna, F. P. (1993). Developmental onset of eating related color-naming interference. International Journal of Eating Disorders, 13, 391-397.

Green, M. W., & Rogers, P. J. (1993). Selective attention to food and body shape words in dieters and restrained nondieters. International Journal of Eating Disorders, 14, 515-517.

Ingram, R. E., & Kendall, P. C. (1986). Cognitive clinical psychology: Implications of an information processing perspective. In R. E. Ingram (Ed.), Information processing approaches to clinical psychology (pp. 3-21). London: Academic Press.

Izawa, C., & Silver, N. C. (1988). Response variations to Stroop color-word stimuli. Genetic, Social, and General Psychology Monographs, 114, 211-255.

Jansen, A., Huygens, K., & Tenney, N. (1998). No evidence for a selective processing of subliminally presented body words in restrained eaters. International Journal of Eating Disorders, 24, 435- 438.

Jones-Chesters, M. H., Monsell, S., & Cooper, P. J. (1998). The disorder-salient Stroop effect as a measure of psychopathology in eating disorders. International Journal of Eating Disorders, 24, 65- 82.

Kaspi, S. P., McNally, R. J., & Amir, N. (1995). Cognitive processing of emotional information in post-traumatic stress disorder. Cognitive Therapy and Research, 19, 433-444.

Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of Psychotherapy and behavior change (4th ed., pp. 19 - 71). New York: John Wiley & Sons.

Kazdin, A. E., & Bass, D. (1989). Power to detect differences between alternative treatments in comparative psychotherapy outcome research. Journal of Consulting and Clinical

Psychology, 57, 138- 147.

Kazdin, A. E., Siegel, T. C., & Bass, D. (1990). Drawing on clinical practice to inform research on child and adolescent psychotherapy: Survey of practitioners. Professional Psychology: Research and Practice, 21, 189-198.

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

Lavy, E., & van den Hout, M. (1993). Selective attention evidenced by pictorial and linguistic Stroop tasks. Behavior Therapy, 24, 645-657.

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

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. Canadian Journal of Psychology, 39, 367-386.

Lovell, D. M., Williams, J. M. G., & Hill, A. B. (1997). Selective processing of shape-related words in women with eating disorders, and those who have recovered. British Journal of Clinical Psychology, 36, 421- 432.

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

Mahamedi, F., & Heatherton, T. F. (1993). Effects of high calorie preloads on selective processing of food and body shape stimuli among dieters and nondieters. International Journal of Eating Disorders, 13, 305-314.

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

Maxmen, J. S., & Ward, N. G. (1995). Essential psychopathology and its treatment (2nd ed.). New York: W. W. Norton & Company.

Mogg, K., Millar, N., & Bradley, B. P. (2000). Biases in eye movements to threatening facial expressions in generalized anxiety disorder and depressive disorder. Journal of Abnormal Psychology, 109, 695- 704.

Ogden, J., & Greville, L. (1993). Cognitive changes to preloading in restrained and unrestrained eaters as measured by the Stroop task. International Journal of Eating Disorders, 14, 185-195.

Perpiñá, C., Hemsley, D., Treasure, J., & deSilva, P. (1993). Is the selective information processing of food and body words specific to patients with eating disorders? International Journal of Eating Disorders, 14, 359-366.

Polivy, J., & Herman, C. P. (1987). Diagnosis and treatment of normal eating. Journal of Consulting and Clinical Psychology, 55, 635- 644.

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

Rieder, S., & Ruderman, A. (2001). Cognitive factors associated with binge and purge eating disorders: The interaction of body dissatisfaction and body image importance. Cognitive Therapy and Research, 25, 801- 812.

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

Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results.

Psychological Bulletin, 86, 638- 641.

Rosenthal, R. (1995). Writing meta-analytic reviews. Psychological Bulletin, 118, 183-191.

Rubino, I. A., Baggi, L., San Martino, L., Cuzzolaro, M., Pezzarossa, B., & Martignoni, M. (1998). Patterns of adaptation to conflict in bulimia and temporo-mandibular joint disorder. Perceptual and Motor Skills, 86, 979- 984.

Sackville, T., Schotte, D. E., Touyz, S. W., Griffiths, R., & Beumont, P. J. V. (1998). Conscious and preconscious processing of food, body weight and shape, and emotion-related words in women with Anorexia Nervosa. International Journal of Eating Disorders, 23, 77- 82.

Schwarzer, R. (1989). Meta-analysis. (Unpublished program and manual). Freie Universitat Berlin, Germany.

Segal, Z. V. (1988). Appraisal of the self-schema construct in cognitive models of depression. Psychological Bulletin, 103, 147-162.

Segal, Z. V., Gemar, M., Truchon, C., Guirguis, M., & Horowitz, L. M. (1995). A priming methodology for studying self-representation in major depressive disorder. Journal of Abnormal Psychology, 104, 205-213.

Skrzypek, S., Weheimer, P. M., & Remschmidt, H. (2001). Body image assessment using body size estimation in recent studies in anorexia nervosa: A brief review. European Child and Adolescent Psychiatry, 10, 215- 221.

Smith, A. P., & Miles, C. (1986). Effects of lunch on selective and sustained attention. Neuropsychobiology, 16, 117-120.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of

Experimental Psychology, 18, 643-662.

Strube, M. J., & Hartmann, D. P. (1983). Meta-analysis: Techniques, applications, and functions. Journal of Consulting and Clinical Psychology, 51, 14- 27.

Thrasher, S. M., Dalgleish, T., & Yule, W. (1994). Information processing in post-traumatic stress disorder. Behavior Research and Therapy, 32, 247-254.

Vitousek, K., & Orimoto, L. (1993). Cognitive-behavioral models of Anorexia Nervosa, Bulimia Nervosa, and obesity. In K. S. Dobson and P. C. Kendall (Eds.) Psychopathology and Cognition. (pp. 191- 243). New York: Academic Press.

Walker, M. K., Ben-Tovim, D. I., Jones, s., & Bachok, N. (1992). Repeated administration of the adapted Stroop test: Feasibility for longitudinal study of psychopathology in eating disorders. International Journal of Eating Disorders, 12, 103-105.

Waller, G., & Ruddock, A. (1995). Information-processing correlates of reported sexual abuse in eating-disordered and comparison women. Child Abuse and Neglect, 19, 745-759.

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

Wilson, G. T., Heffernan, K., & Black, C. M. D. (1996). Eating disorders. In E. J. Mash & R. A. Barkley (Eds.). Child psychopathology (pp. 541-571). New York: Guilford Press.

Footnotes

[1] We acknowledge that this reporting structure does not provide the results of every possible combination of effect sizes. Indeed, we calculated a total of 129 specific effect sizes from the reviewed studies, which fell into 54 possible contrast types (many with no available data). Readers who are interested in other specific comparisons may contact the first author.

[2] It should be noted that these comparisons do not involve the Classic Stroop stimuli as comparisons, but are rather based on those studies that generated comparison conditions for the emotional Stroop words used, or reported results for "neutral" content stimuli.

Table 1.

Studies that have examined the Emotional Stroop Task in Eating Disorders.

Study	**Participants**	**Conditions/ Groups**	**Gender**	**Independent Variable(s)**	**Stroop Content**	**Dependent Variable(s)**
Smith & Miles (1986)	College students	Before/after lunch	30F/ 18M	Time	Food Body	Latency Interference
Channon et al. (1988)	Clinical	Anorexia Normal controls	20F/ 0M 20F/ 0M	Group Stroop Content	Food Stroop Body Stroop	Latency
Ben-Tovim et al. (1989)	Clinical	Anorexia Bulimia Controls	17F/ 0M 19F/ 0M 38F/ 0M	Group Stroop Content	Classic Stroop Food Stroop	Latency
Channon & Hayward (1990)	Normals	Fasting Nonfasting	8F/ 8M 8F/ 8M	Fasting Condition Stroop Content	Food Stroop Body Stroop	Latency
Ben-Tovim & Walker (1991)	Clinical	Anorexia Bulimia High drive for thinness Low drive for thinness	27F/ 0M 29F/ 0M 37F/ 0M 22F/ 0M	Group Stroop Content	Food Stroop Classic Stroop	Latency
Fairburn et al. (1991)	Clinical	Bulimia Normal controls	24F/ 0M 50F/ 0M	Group Stroop Content	Classic Stroop Food Stroop	Latency
Cooper & Fairburn (1992)	Clinical	Anorexia Bulimia Symptomatic dieters	12F/ 0M 12F/ 0M 12F/ 0M	Group	Food Stroop	Latency
Cooper et al. (1992)	Clinical	Bulimia Normal controls	36F/ 0M 18F/ 0M	Group Stroop Content	Classic Stroop Food Stroop	Latency Interference

Walker et al. (1992)	Normals	Time	60F/ 0M	Stroop Content	Food Stroop Body Stroop	Latency
Cooper & Fairburn (1993)	Clinical	Bulimia	75F/ 0M	Stroop Content	Food Stroop Classic Stroop	Latency
Green & McKeena (1993)	Normals	9-year-olds 11-year-olds 14-year-olds	20F/ 20M 20F/ 20M 20F/ 20M	Age Gender Stroop Content	Food Stroop Body Stroop	Latency
Green & Rogers (1993)	Normals	Dieters Restrained eaters Normal controls	13F/ 0M 15F/ 0M 27F/ 0M	Group Stroop Content	Food Stroop Body Stroop	Latency
Mahamedi & Heatherton (1993)	Normals	Dieters, Nondieters	47F/ 0M (total)	Group	Food Stroop	Latency
Ogden & Greville (1993)	Normals	Dieters	28F/ 0M	Stroop Content Caloric Load	Food Stroop Body Stroop	Latency
Perpiñá et al. (1993)	Clinical	Anorexia Bulimia Normal controls	18F/ 0M 14F/ 0M 32F/ 0M	Group Stroop Content	Food Stroop Body Stroop	Latency
Cooper & Fairburn (1994)	Clinical	Bulimia	58F/ 0M	Pre/post treatment Stroop Content	Classic Stroop Food Stroop	Interference
Waller & Ruddock (1995)	Clinical	Anorexia, Bulimia Normal controls	20F/ 0M 30F/ 0M 30F/ 0M	Group Abuse history Stroop Content	Classic Stroop Food Stroop Sexual abuse	Latency

Black, et al, 1997	Clinical	Bulimia Restrained eaters Unrestrained eaters	16F/ 0M 16F/ 0M 13F/ 0M	Group Stroop Content	Food Stroop Body Stroop Classic Stroop	Latency
Cooper & Todd, 1997	Clinical	Bulimia Anorexia Control	12F/ 0M 12F/ 0M 18F.0M	Group Stroop Content	Food Stroop Weight/ Shape Stroop	Latency
Lovell, et al, 1997	Clinical	Current Bulimia Recovered Bulimia Current Anorexia Recovered Anorexia Control	24F/ 0M 11F/ 0M 31F/ 0M 23F/ 0M 33F/ 0M	Group Stroop Content	Food Stroop Weight Stroop	Latency
Jones-Chesters, et al, 1998	Clinical	Bulimia Control	16F/ 0M 16F/ 0M	Group Presentation format Stroop Content	Food Stroop Weight Stroop Emotional Stroop	Latency
Rubino, et al, 1998	Clinical	Bulimia TMJ Patients	45F/ 0M 45F/ 0M	Group	N/A	Color-word test types and clusters
Sackville, et al, 1998	Clinical Normal	Anorexia High restraint Low restraint	20F/ 0M 20F/ 0M 33F/ 0M	Group Stroop Content	Food Stroop Shape Stroop Emotional Stroop	Latency
Jansen, et al, 1998	Normal	High restraint Low restraint	13F/ 0M 15F/ 0M	Group Presentation format Stroop Content	Body Stroop Neutral Content	Latency

Green, Dorr, & DeSilva, 1999	Clinical	Anorexia Control	34F/ 0M 39F/ 0M	Group Stroop Content	Body Stroop Control Content	Latency
Carter, et al, 2000	Clinical	Bulimia	99F/ 0M	Pre/post treatment Stroop Content	Food/Body Control Content	Latency

Note. F = females; M = males; N/A = not available.

Table 2

Effect Sizes for Between- Group Comparisons of the Stroop Task in Eating Disorders

Comparison	Effect Size	n of studies	95% Confidence Interval	Homogeneity Q	Fail safe n
Food Stroop					
Bulimic vs. Control	.59	9	.28 □ .90	23.98**	18
Anorexic vs. Control	.20	5	-.30 □ .69	16.21**	0
Dieting/ restricted vs. Control	.39	6	.14 □ .64	4.50	6
Body/ Weight Stroop					
Bulimic vs. Control	.57	6	.08 □ 1.06	24.67**	11
Anorexic vs. Control	.45	6	-.08 □ .99	28.23**	8
Dieting/ restricted vs. Control	.09	5	-.27 □ .46	5.95	0
Classic Stroop					
Bulimic vs. Control	.48	6	.07 □ .89	17.82**	8
Anorexic vs. Control	-.10	2	-1.32 □ 1.12	9.44**	0
Dieting/ restricted vs. Control	-.32	2	-1.19 □ .55	3.85*	1

Notes:

- Effect Sizes (ES) are based on Glass' (1977) g statistic.

*- P < .05 ** p < .01

- Fail safe n calculations are based on the number of nonsignificant studies needed to reduce the average effect size to # .20.

Table 3

Effect Sizes for Within- Group Comparisons of the Stroop Task in Eating Disorders

Comparison	Effect Size	n of studies	95% Confidence Interval	Homogeneity Q	Fail safe n
Bulimic					
Food vs. Body/ Shape	-.08	6	-.34 □ .18	3.38	0
Food vs. Control	-.13	8	-.59 □ .32	28.31**	0
Body/ Shape vs. Control	.16	6	-.46 □ .80	30.35**	0
Anorexic					
Food vs. Body/ Shape	-.08	5	-.35 □ .19	2.28	0
Food vs. Control	-.25	4	-1.10 □ .60	21.79**	1
Body/ Shape vs. Control	-.12	5	-.85 □ .61	26.83**	0
Dieting/ Restricted Eating					
Food vs. Body/ Shape	-.06	4	-.37 □ .26	1.47	0
Food vs. Control	-.09	3	-.72 □ .54	6.61*	0
Body/ Shape vs. Control	-.03	4	-.56 □ .54	9.56*	0

Notes:

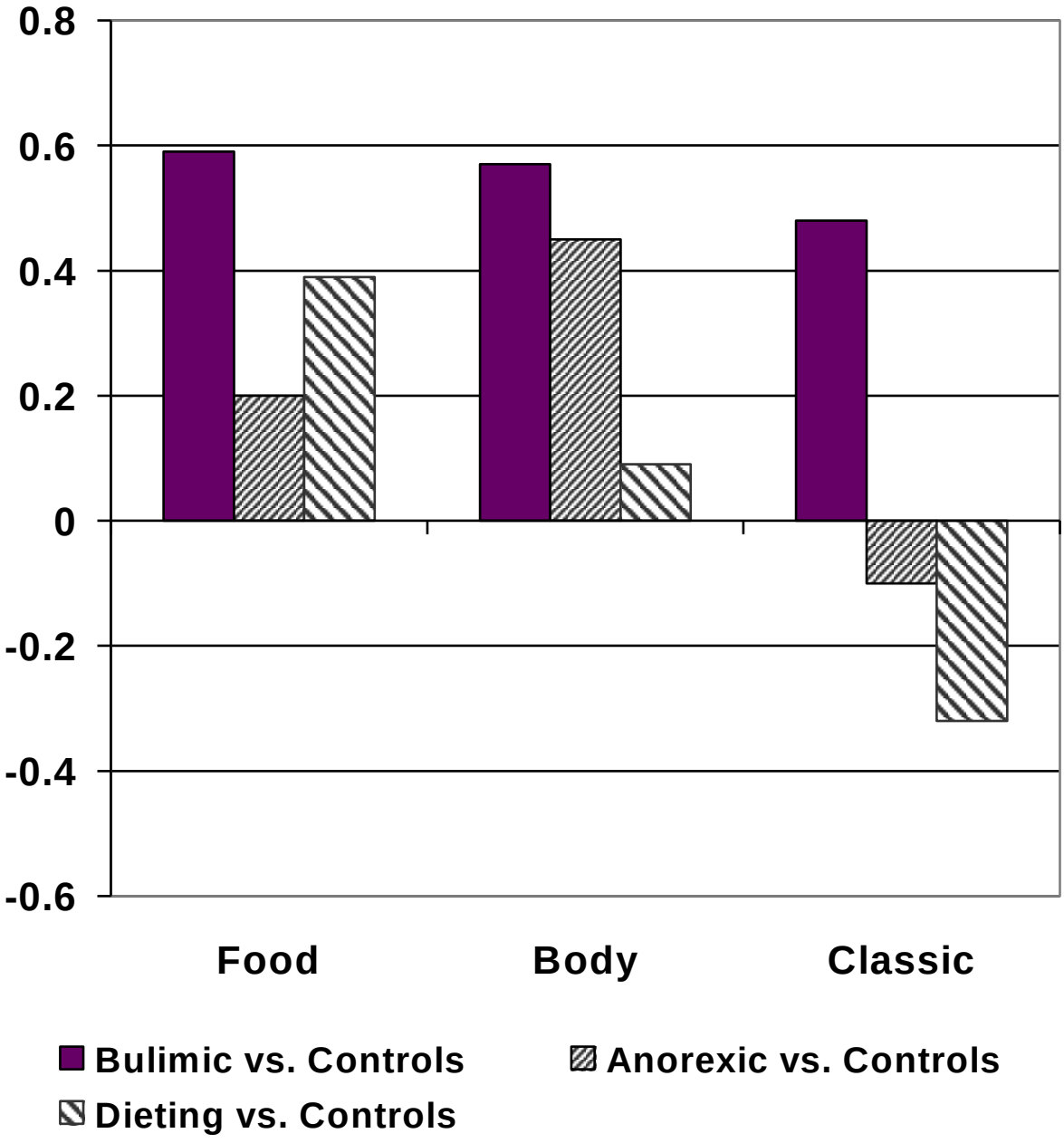
- Effect Sizes (ES) are based on Glass' (1977) g statistic.

*- P < .05 ** p < .01

- Fail safe n calculations are based on the number of nonsignificant studies needed to reduce the average effect size to # .20.

Figure Captions

Figure 1. Average Effect Sizes for Between- Group Comparisons of the Stroop Task in Eating Disorders, for Food, Body/ Shape, and Classic Stroop stimuli.


0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
Food
Body
Classic
Bulimic vs. Controls
Anorexic vs. Controls
Dieting vs. Controls