1

THE STROOP TASK: INDIRECTLY MEASURING CONCEPT ACTIVATION

COLIN MACLEOD AND COLIN MACLEOD

For well over a century, since the earliest days of the discipline, psychologists have sought to develop methods that permit them to uncover the current contents of mind. To oversimplify with a dichotomy, there are two intuitive ways to accomplish this goal. In the direct technique, psychologists can ask the individual what he or she is thinking, being entirely candid that this is what they are doing. Introspection relies on this explicit approach. In the indirect technique, psychologists can ask the individual to perform some seemingly unrelated task that they believe will reveal what he or she is thinking, without his or her awareness. Such psychoanalytic staples as free association and dream analysis rely on this implicit approach. Experimental psychologists and clinical psychologists have shown endless creativity in constructing many clever variants of these two approaches.

Imagine that you have come to believe that thoughts of suicide are dominating the thinking of a troubled teenager. You reason that if such

thoughts are dominant, then words related to suicide should be more highly activated than words not related to suicide. So you design a simple, quite direct test of speeded word reading. You ask that young person to read aloud—as quickly as possible—a series of words, some of which are neutral (i.e., unrelated to suicide), words such as *table* and *honesty*. But you also mix in words that are related to suicide, such as *death* and *sadness*. If the teenager's thoughts are dominated by suicide, you expect that his or her time to read these two sets of words aloud will be different. It might be slower to read the suicide-related words than the neutral words because he or she is defending against them; it might be faster to read the suicide-related words because these words and concepts are already in consciousness. Either finding would support your hypothesis and would influence how you interact with or treat this young person.

There are, of course, potential problems with this approach. The most obvious of these is that the person performing your test cannot help but notice the distinction between the suicide-related and the neutral words. After all, the person's attention is directed to and he or she is responding to the words. This could in turn influence the time to read the words for reasons other than those that led you to construct the test. The greater activation of the suicide-related words may result in those words being read more quickly. Or the person may start to explicitly search for the suicide-related words, slowing down when these words are detected. Or, guessing the rationale for the test, he or she might intentionally speed up or slow down when responding to the suicide-related words relative to the neutral words. How can you be certain that the pattern of reading times that you observe is uncontaminated by such strategic responding? It is for this very reason that the indirect approach has long been seen as offering certain advantages.

Change the procedure just described apparently only superficially and you have an indirect measure of what your troubled teenager is thinking. Print each of the words in color, such as *table* in red and *sadness* in yellow, and ask the person not to read the words but to name the colors in which they are printed as quickly as possible. Attention is now directed away from the words, despite the fact that the words are what you are interested in. Now if you find that the time to name the colors of suicide-related words is longer than the time to name the colors of neutral words, you have evidence that the suicide-related words are more activated and hence interfere with naming the colors in which they are printed more than do the neutral words. Because the person is not intentionally reading the words—indeed, he or she is explicitly told not to read them—this interference provides an indirect measure of concept activation.

It is really quite remarkable that at the height of behaviorism, with its abhorrence of all things mentalistic, this very task was created. In his dissertation, John Ridley Stroop (1935) presented participants with color words printed in color (e.g., *red* in green), instructing them either to read the word

(say "red," the more natural, direct response) or to name its print color (say "green," indirectly measuring the interference from the word on the color). He showed that, compared with a suitable control item—reading a word in normal black ink in the case of reading, or naming a color patch without a word in the case of color naming—there was substantial interference in the color-naming task but none in the reading task. Since 1935, but primarily since the mid-1960s, many hundreds of studies have been done in the subdiscipline that we now call cognitive psychology in which the Stroop task was used to explore aspects of attention, learning, and memory (see MacLeod, 1991, for a review). This powerful interference effect has helped to disentangle numerous puzzles about the operation of mind.

The classic color–word Stroop task has more recently been adapted to incorporate noncolor words in place of the traditional color words. This modification has come to be called the *emotional Stroop task* because the response to the critical noncolor words, as in the suicide illustration just described, is seen as at least partly having to do with affect. Since the 1980s, there has been a rapidly increasing number of studies in clinical psychology in which this emotional Stroop task has been used to explore aspects of anxieties and clinical disorders (see Williams, Mathews, & MacLeod, 1996). This variation on the task has been useful in diagnosing problems and even in evaluating the efficacy of treatments.

In the two chapters that follow, we describe the cognitive and clinical perspectives on the Stroop task, with emphasis on the methods used to carry out these studies. It is all too easy to let the surface simplicity of this task lead one to assume that the construction of a good Stroop experiment is straightforward. That would be a mistake. There are rich complexities to this task, as reflection on its 69-year (and counting) history should make evident. In chapters 2 and 3, our primary goal is to bring these complexities under the light and to suggest ways in which they have been and can be handled, along the way discussing what we have learned from these studies.

We cannot preface here all of the issues that we consider in more detail in our two chapters. Instead, we just raise two issues as illustrating the complexity inherent in the Stroop task. The first is at the empirical core of the task: the composition of the neutral, control condition. What should be the nature of this condition? This may seem obvious, but we show in our chapters that it is not, and that it warrants substantial thought. The second issue that we emphasize here has in fact already been raised. It is a crucial conceptual issue. Investigators routinely, probably too routinely, treat the Stroop task as an indirect measure of the processing of the to-be-ignored words. The logic goes that because the participant must name the color and not read the word, then any influence of the word on the speed of color naming is necessarily indirect. We advise caution in accepting this logic uncritically. Participants may well notice the words and even attend to them as they name their colors. At what point does this attention to the words begin to move

the task from indirect to direct? There is no easy answer to this question, but investigators using the task because of its perceived indirect advantages should always be alert for evidence that the participants may not be treating the task as indirect, even though the investigator is doing so.

We have raised these two issues not to suggest that it is impossible to construct good Stroop experiments from which useful interpretations can be drawn. On the contrary, our intention has been to suggest that, as always, good conclusions hinge on good experiments, which in turn are based on hypotheses derived from sound theories. We believe that a great deal can be learned about cognitive processes—normal and abnormal—with this tool that is so familiar to every psychologist. And we very much hope that our chapters will assist investigators in using the Stroop task to advance our understanding of these fundamental processes.

REFERENCES

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

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662. (Reprinted in 1992 in the *Journal of Experimental Psychology: General, 121,* 15–23.)

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