

Increasing the Clinical Utility of the Stroop Color-Word Test through Process Analysis

Jack Demick^{1*} and Casey Marks²

¹Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA

²Professional Programs at (ISC)², Clearwater, Florida, USA

Abstract

The Stroop Color-Word Test or SCWT (on Card A, S reads 100 color names, on Card B names 100 color patches, on Card C names 100 incongruous ink colors in which color names are printed) has been in existence for over 80 years. It is best known in modern psychology as a tool in the armamentarium of the clinical neuropsychologist. However, its usefulness has been limited in that researchers' singular use of achievement measures (e.g., total time per card, total number of errors per card, derived scores from total time measures, namely: speed factor, total time on Card A; color-difficulty factor, total time on B/total time on A; interference factor, total time on C - total time on B) does not finely discriminate among groups exhibiting different psycho- and neuro-pathologies. We argue here that a process-oriented approach to the SCWT - focusing on an analysis of the types of errors involved in the identification of stimuli and the maintenance of serial organization, types of nonverbal behaviors/cognitive strategies to complete the task, and metacognitive estimates of performance) - will complement its traditional achievement approach and increase its discriminative ability (cf. Werner on process vs. achievement). Toward this end, we describe studies from an ongoing research program in our laboratory. In doing so, we also attempt to rectify two longstanding concerns leveled against the SCWT - the needs to standardize its materials/administration/scoring and to develop a comprehensive set of norms, in our case one unified by a classic theory of development.

Keywords: Brain injury; Cognition; Cognitive science; Frontal lobe; Parallel processing; Process vs. achievement; Stimulus identification; Serial organization; Selective attention; Stroop Color-Word Test; Stroop effect; Heinz Werner

Increasing the Clinical Utility of the Stroop Color-Word Test through Process Analysis

The Stroop Color-Word Test or SCWT [1-4] is best known in modern psychology as an assessment tool in the armamentarium of the clinical neuropsychologist. However, the SCWT and the "Stroop effect" - the finding that naming the colors in which color words (e.g., "red," "blue," "green") are printed occurs more quickly and with fewer errors if the actual colors of the words match the colors that the words denote (e.g., "red" printed in red ink rather than in blue ink) - have been of considerable interesting empirical research for over 8 decades.¹

Reasons underlying this longstanding interest have become apparent from two major review articles published 25 years apart and an informal review of the major studies of the last 25 years. These reviews, summarized below, have suggested that SCWT investigations may be conceptualized in terms of three waves, with each wave influenced by the Zeitgeist in which the work was conducted.

In their review article covering the first 30 years of SCWT research, Jensen and Rohwer [5] suggested that there was significant interest in the SCWT because measures from this task - yielding highly stable measures of individual differences on simple, basic aspects of human performance (e.g., word reading, color naming) only tenuously related to intelligence - also show significant relations to a host of more complex psychological variables (e.g., interference is negatively correlated with

¹Some sources (e.g., Jensen and Rohwer) have dated the origins of the SCWT to significantly earlier, that is, to 1883 when the father of psychology, Wilhelm Wundt, suggested to one of his students, James McKeen Cattell, that he conduct the research for his doctoral dissertation on the time it takes for one to name objects and colors and to read the corresponding words. Cattell [12] followed this advice and published the first experimental study on the relative speeds of color-naming and color-word reading, demonstrating faster reaction times for color-word reading than for color-naming.

memory span, serial learning, and creativity and positively correlated with psychopathology). In essence, consistent with the goals of the heyday of psychology as a nascent science, the first wave of research was concerned primarily with establishing the psychometric aspects of the task and providing normative data (e.g., with no real attempt to develop comprehensive SCWT norms, making comparisons across studies difficult).

Some years later, McLeod [6] summarized a second wave of SCWT research, perhaps initiated by the cognitive revolution of the 1950s, which focused predominantly on evaluating the proposed psychological theories posited to account for the task's reliable findings. He concluded that theories placing explanatory weight on multiple dimensions - specifically, the parallel processing of relevant and irrelevant dimensions - were likely to be more successful than earlier theories attempting to locate a single snag in an isolated process (e.g., attention, general processing speed, speed of speech, automaticity of reading).

Paralleling psychology's focus over the past several decades on the underlying biological bases of behavior and experience perhaps reinforced by "The Decade of the Brain" [7], the third wave of SCWT research has increasingly focused on both indirect and direct brain correlates of SCWT performance. Examination of indirect indices

*Corresponding author: Jack Demick, Fellow, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA, Tel: (617) 495-4024; E-mail: jack_demick@harvard.edu

Received September 03, 2016; Accepted September 22, 2016; Published September 28, 2016

Citation: Demick J, Marks C (2016) Increasing the Clinical Utility of the Stroop Color-Word Test through Process Analysis. J Ment Disord Treat 2: 125. doi:10.4172/2471-271X.1000125

Copyright: © 2016 Demick J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

was exemplified, for example, by the work of West and Bell [8], who examined SCWT interference and electroencephalogram activation (providing evidence for an age-related decline of the anterior attention system) and of Langleben et al. [9], who researched the effects of the psycho-stimulant methylphenidate (MPH) on SCWT performance (demonstrating that MPH improves performance in boys both with and without ADHD with the former group, both on and off MPH, displaying slower performance across cards relative to the latter group).

Examination of direct brain correlates of SCWT performance was subsequently made possible by the emergence of a wide range of neuro imaging techniques especially MRI techniques assessing information about the basic structure of the brain and abnormalities in brain structure, and fMRI techniques providing information about brain activity when individuals perform various types of motor or cognitive tasks. Stated succinctly, imaging studies [10,11] have revealed that SCWT performance is primarily associated, in normal individuals, with frontal and parietal cortex activity and, in particular, with the anterior cingulate cortex and dorsolateral prefrontal cortex, implicating working memory, processing speed, conceptual ability, semantic knowledge, and executive functioning (e.g., the ability to inhibit responses, to establish, maintain, and shift sets on tasks, to plan ahead, and to use feedback to modify responses). These findings [12,13] have been consistent with poor response inhibition (increased SCWT interference) in patients with focal frontal lesions and presumptive frontal dysfunction (including those diagnosed with alcoholism, dementia, HIV encephalopathy, Huntington's disease, Parkinson's disease, prenatal alcohol exposure, schizophrenia, and traumatic brain injury).

SCWT Forms and Measures

Stroop's original task consisted of three cards administered sequentially. On Card A (word reading), subjects were asked to read color names as quickly and accurately as possible; on Card B (color naming), they were asked to name the colors of color patches as quickly and accurately as possible; and on Card C (incongruous colors-words), they were asked to name, as quickly and as accurately as possible, the color of the ink in which incongruous color words were written (e.g., "blue" printed in red ink). The words and color patches were arranged in a 10 x 10 matrix of evenly spaced rows and columns with an attempt made to avoid any regularity in sequence of color (horizontally or vertically). Because Stroop did not specify the exact size of the cards, the size or shapes of the color patches, their spacing, or why he utilized five colors (red, blue, green, purple, and brown, the last of which he changed at some unspecified point from yellow for increased visibility), a version similar to Stroop's original version was produced by the C. H. Stoelting Company, manufacturers of psychological laboratory equipment, which is no longer in print and commercially available. However, a variation of this latter version (e.g., employing only red, blue, and green, adding a practice line of 10 items on each card) has consistently been employed in numerous SCWT studies conducted at Clark University and is often referred to in the literature as the Clark University Stroop Color-Word Test (copies of which are currently available from the first author).

Stroop's standard scoring of the task consisted of the following. The primary measure from the task was the interference factor (Stroop effect), which was determined by subtracting the total time on Card B (color patches) from the total time on Card C (incongruous color names and inks). Secondary measures included the: speed factor (total time on Card A); and color-difficulty factor (total time on Card B/total

time on Card A). These measures demonstrated that subjects typically exhibit faster performance on Card A than on Card B than on Card C. However, the best method for determining SCWT interference has more recently become a controversial issue in the literature with a variety of alternatives proposed [14]. Further, for reasons not entirely clear, numerous contemporary researchers have chosen to employ only total time scores and/or their derivatives (above), neither obtaining nor employing error scores in their analyses.

Partly related to the difficulty in obtaining the original SCWT, the Stoelting version, and/or the Clark University SCWT and to attempts to expand and understand the basic paradigm, numerous variations of the SCWT have appeared in the literature [15]. Some of the more common theoretical and methodological variations have included the use of: (a) curve-shaped words; (b) reverse instructions; (c) spatial stimuli; (d) numerical stimuli; and (e) emotional stimuli. For the most part, such variations using alternative stimulus presentations and modalities have evidenced findings similar to the original Stroop effect with the most useful and subsequently widespread variation employing emotional stimuli (participants are asked to identify the color of the ink in which neutral words, e.g., door, and emotional words, e.g., violence, are printed), particularly in attempts to understand a variety of psychopathologies. For example, using a modified emotional SCWT, Chung and Jeglic [16] have reported that past suicide attempters are generally slower than non-attempters in responding to the word suicide and that, among females, past attempter's exhibit more delayed response when their most recent attempt was made within the past 12 months.

Achievement-Oriented Analyses

What is important in the original SCWT and its myriad of variants is that studies have generally employed only achievement measures (e.g., scores predominantly based on total time measures and, to a lesser extent, measures of total number of errors) in their analyses. This analytic strategy is potentially problematic in several regards.

Relevant here, the well-known developmental psychologist, Werner [17], in a landmark article, made the important distinction between process versus achievement. Specifically, he argued that it is wrong to assume that any achievement, that is, the final solution to a problem, is an objective measure of some underlying unitary mechanism. Rather, he suggested that the final solution to a problem may be arrived at through diverse processes, which may reflect different activities of various structures in the central nervous system. Thus, although two individuals may obtain the same IQ scores on a standard intelligence test, the underlying process that they utilize toward obtaining their final answers most probably reflect different patterns of cognitive assets and liabilities. He [18] later underscored this distinction in his discussion that any cognitive act involves microgenesis, that is, an unfolding process over time. These Wernerian notions have for some time now served quite profitably as the basis for Kaplan's [19] process approach to neuropsychology and suggest that a process analysis of the SCWT, heretofore unattempted, may similarly contribute to the increased utility, clinical and otherwise, of the SCWT.

Two examples of problems associated with using only achievement-oriented analyses of the SCWT have been apparent in recent research. First, beginning in the first wave of SCWT research through the present, investigators have provided normative data on SCWT performance, although not in a particularly comprehensive manner. For example, as part of a larger investigation on the role of cognitive style in adult automobile driving behavior, Demick and Harkins [20]

administered the SCWT to 231 U.S. (Massachusetts) participants in four age groups: young adults (20-39 years); middle-aged adults (40-59 years); early older adults (60-74 years); and advanced older adults (75+ years). In line with the field's relatively new focus on diversity in human experience and action, Moering, Schinka, Mortimer, and Borenstein Graves [21] administered the SCWT to 236 U.S. African American adults, divided into the two age groups of 60-71 years and 72-84 years, living in Florida.

Similar investigations have recently been conducted outside of the United States. For instance, as part of a larger prospective study on the determinants of cognitive aging, van der Elst, van Boxtel, van Breukelen, and Jolles [22] administered the SCWT to 1,788 participants between the ages of 21 and 81 years living in The Netherlands. As a follow-up to this study 3 years later, van Boxtel, ten Tusscher, Metssemakers, Willems, and Jolles [22] again tested 838 individuals aged 52 years and older from the original study on the SCWT and other tasks. Seo et al. [23] administered the SCWT to 564 older adults ranging in age from 60 to 90 years living in Korea. Finally, as part of a large-scale project to generate culture-specific norms for a variety of neuropsychological tasks, Peña-Casanova et al. [24] administered the SCWT to 691 participants in Spain.

Clearly a massive amount of data has been collected in these and other investigations. However, although these studies have made successful attempts to employ only healthy, community-dwelling participants, it remains unclear as to whether the similarities and differences across samples in SCWT normative scores based only on achievement measures are a function of characteristics of the task itself (e.g., form, administration, scoring); characteristics of researchers' analytic strategies (e.g., some have provided raw scores while others have reported age-, sex-, and/or education-adjusted scores; some have administered the SCWT on multiple occasions even in light of ongoing reports of a pronounced practice effect beginning with Stroop [2]); and/or characteristics of the sociocultural contexts in which the data have been collected (see below).

A second more important issue concerns the usefulness of employing SCWT achievement measures to discriminate among various patient groups. For example, how clinically useful is it to know that a given patient's SCWT interference score is greater than that of a normal subject, particularly when those with a range of psychopathologies and neuropathology's have been shown to exhibit more interference than clinical controls (cf. Strauss et al. [13])?

For example, in a meta-analysis of the sensitivity and specificity of the SCWT in children, Homack and Riccio [25] have reported that, although children and adolescents with an ADHD consistently exhibit poor performance relative to those without clinical diagnoses, the SCWT does not discriminate ADHD groups from other clinical groups consistently across studies. Further, Savitz and Jansen [26] suggested that the issue becomes even more complex within diagnostic groups. They have found that, relative to a control group of normal boys, those diagnosed with an ADHD exhibit poorer performance on Cards A (word reading) and C (incongruous color words), suggesting ADHD-specific deficits in both reading and executive functioning; however, when boys with both an ADHD and a learning disorder in reading are excluded from the analysis, the difference between groups occurs only on Card C, indicating that poor reading skills may produce false negatives on the SCWT.

In line with this, national comorbidity surveys [27] have indicated that 48% of the American adult population experiences mental illness

in their lifetimes with 27% experiencing the simultaneous presence of at least two and often more chronic disorders with slightly higher rates for children and adolescents, all of which are projected to increase in the future. Thus, failure to take into account the high prevalence of the comorbidity of psychiatric illnesses has the potential to confound clinical research in general and normative clinical research on tasks such as the SCWT in particular.

Complementing SCWT Achievement Measures with Process Measures

The above review has indicated that all previous studies including the SCWT have employed only achievement measures. There was one early exception, however. Employing serial scoring (the recording of total time for S to complete every two rows, i.e., 20 responses, on Card C), Smith and Klein [28] identified three patterns of responders: dissociatives (whose five scores rise and fall discontinuously reflecting the faltering of attention); cumulatives (who show increasingly slower reading times over the five scores); and stabilizers (whose scores maintain an even course over the five scores, leading them to exhibit lesser interference than the other two groups). Further, the researchers reported that the three groups differ on other psychological tasks (e.g., on the Group Embedded Figures Test - see Demick, [29] dissociatives exhibit greater variability in performance than the other groups possibly related to increased confabulation, i.e., perception of hidden stimuli in places where none actually exist, while cumulatives are generally slower than the other two groups). Although employing a variation of the achievement measure of total time score, such an analysis should be lauded for making inroads into the problem of underlying process and for contributing to the convergent validity of the SCWT. Nonetheless, there is a need to go further.

The seeds for a more comprehensive analysis of SCWT performance were sown by some of my early colleagues at Clark University, who couched their interpretations of the Stroop effect in broader psychological terms than previous investigators favoring S-R theory. Conceptualizing performance on Card C as reflecting a general capacity to maintain a course of action in the face of intrusion by other stimuli [30,31], the ability to resist interference is related to a basic principle of Wernerian theory [18], namely, that cognitive development (in fact, all aspects of development) involves increasing differentiation and hierarchic integration (here, with Stroop Card C requiring the ability to differentiate color words from their ink colors, subordinating the former to the latter).²

Interpreting error categories in terms of two sets of processes underlying SCWT performance -namely, identification (of the appropriate aspect of the stimulus item such as color) and (difficulty in the process of maintaining) serial organization - Clark University researchers (below) observed in Stroop performance that overall achievement or end state is dependent on different underlying processes or means related to age (cf. Werner, [18], on means-ends relationships in development whereby, e.g., similar means may lead to different ends and multiple means may lead to the same end).

For example, administering the SCWT to children and youth in four age groups (6, 9, 12, 16 years), Rand, Wapner, Werner, and McFarland [32] observed that whereas inserted linguistic words or

²For extensions of Werner's (1957) organismic-developmental theory, the interested reader is referred to Wapner and Demick's (1998) holistic, developmental systems-oriented approach to person-in-environment functioning across the life span and to Demick's (2014, 2016) more recent reframing of the latter into holistic/systems-developmental theory.

phrases (e.g., S inserts words such as “yes,” “and,” “that’s” to maintain serial organization) decrease with age, inserted non-linguistic utterances (e.g., S inserts vocalizations such as “uh” or “um” between items) increase with age. Thus, the function of inserted utterances, regardless of their form, changes from a developmentally less advanced strategy in children (with inserted linguistic utterances related to errors in maintaining serial organization, in turn leading to an increase in total time) into a developmentally advanced process for adolescents (with nonlinguistic utterances subsequently developed as a cognitive strategy to maintain serial organization, in turn leading to a decrease in total time). Further, Comalli, Wapner, and Werner [33] observed that, younger and older Ss (ranging in age from 7 to 80 years) often achieve their SCWT achievement scores through somewhat different means, although the functions of these means may be similar. For instance, whereas children attempt to overcome interference by literal pointing at the items, older adults often use “verbal pointing” by preceding each item with “that’s a ...”

Following these studies, a more comprehensive SCWT process analysis was developed. Categories of errors for both the processes of identification and maintaining serial organization were expanded, nonverbal behaviors (potentially reflecting errors and/or cognitive strategies) were identified, and metacognitive judgments about achievement measures (e.g., total time, total errors, use of cognitive strategies/mnemonics) were included (see Table 1 for our elaborated scoring system). We then videotaped subjects ranging in age from 3 to 93 years during SCWT performance (or a variant) and subjected the videotapes to both achievement and process analyses. While aspects

of these data have been presented at professional meetings over the years with our ultimate goal of providing comprehensive SCWT performance norms in the future, we will discuss aspects of these data in summary form here.

Data on Development

As indicated above, Rand et al. [32] began their data collection with 6-year-olds and Comalli et al. [33] on 7-year-olds. This was based on the assumption that, prior to at least the age of 6 years, reading ability is not sufficiently established to serve as a potent factor of interference on SCWT Card C. However, toward assessing this assumption, we employed a downward extension of the SCWT not requiring reading, namely, Santostefano’s [34] Fruit Distraction Test (FDT) from which we employed three cards: one that required Ssto name 100 color patches as quickly as possible, a second card requiring them to name the colors of 100 appropriately colored fruits (e.g., “yellow” bananas, “red” apples), and a third that asked them to name as quickly as possible the correct color of inappropriately colored fruits (e.g., blue bananas requiring the S to respond with “yellow”).

Examining the performance of preschoolers on the FDT, we have found that:(a) relative to 3- and 4-year-olds, 5-year-olds perform significantly faster on all three cards; (b) make fewer errors on all cards; and (c) use more efficient (e.g., nodding) as well as emphatic (e.g., foot stomping, shouting) to maintain serial organization. Taken in combination with findings at later ages, these results have suggested that there is a developmental progression in the processes and strategies underlying sequential organization. Specifically, when compared with

A. Deviant Responses to Items	(Reflecting difficulty in the process of identification of appropriate aspect of stimulus item) <ol style="list-style-type: none"> 1. Inappropriate color responses (e.g., S says “red” in response to the stimulus word “green”) 2. Contaminated responses (e.g., S says “brue,” a contamination of the word “blue” with the word “red”) 3. Inarticulate utterances (e.g., S whispers, stutters, mispronounces) 4. Partly wrong responses (e.g., S begins to make a wrong response, then corrects it) 5. Wholly wrong corrected responses (e.g., S gives a wrong response, then corrects it) 6. Reading the word rather than naming the color 7. Inserted color words in addition to naming color
B. Deviant Responses to Sequences	(reflecting difficulty in the process of serial organization) <ol style="list-style-type: none"> 1. Insertion of color words (e.g., S repeats previous responses to restore correct identification and/or to insulate responses from one another in the sequence) 2. Omissions (e.g., S neglects to respond to an item or a line of items) 3. Inserted linguistic words or phrases (e.g., S inserts words such as “yes,” “and,” “that’s”) 4. Inserted nonlinguistic utterances (e.g., S inserts vocalizations such as “uh” or “um” between items) 5. Jumbled order of responses (e.g., S loses place and starts the line over) 6. Repeats line 7. Corrects former item after starting response to next item 8. Others
C. Nonverbal Strategies	<ol style="list-style-type: none"> 1. Pointing at each word 2. Nodding 3. Lip movements 4. Squinting 5. Rocking body 6. Chunking by two’s 7. Chunking by three’s 8. Indexing line with finger 9. Facial movements 10. Enunciation of words 11. Shouting 12. Hard breathing 13. Holding head 14. Holding hands behind back 15. Holding hands at sides 16. Covering items or parts of items 17. Others
D. Metacognition	<ol style="list-style-type: none"> 1. Estimates of total time for each card 2. Estimates of total number of errors per card 3. Perceived cognitive strategies for process of identification 4. Perceived cognitive strategies for process of serial organization

Table 1: Verbal Errors and Nonverbal Strategies: SCWT Process Analysis.

those obtained from adults, they have suggested that younger children may need to use more pronounced/emphatic or externalized/overt strategies prior to being able to use more internalized strategies.

Further, observation of preschoolers' performance on the FDT - particularly that of the 5-year-olds and some of the 4-year-olds with respect to their ability to maintain serial organization through strategy usage - led us to examine, in a separate study, whether Card C of the SCWT might be employed as a measure of reading readiness. Storey and Demick [35] hypothesized that, relative to children who are reading ready - with reading readiness generally defined as the point at which an individual is ready to learn to read with readiness skills thought to include appreciation of stories and books, age-appropriate oral language development and vocabulary, phonemic awareness, understanding of basic print concepts (e.g., pages written in English are read from left to right beginning at the top of the page), understanding that letters represent the sounds of language, visual discrimination (particularly of shapes), and the ability to identify at least some letters of the alphabet (National Institute for Literacy, [26]) - relative to children who are reading ready, children who are not reading ready will perform better on SCWT Card C because they will not be as distracted by the color words themselves.

Toward this end, we examined the potential usefulness of SCWT interference (10 min administration) as an alternative measure of reading readiness to the traditional Metropolitan Readiness Tests, 6th Edition or MRT (1.5 hr minimum administration). Thirty-four preschoolers (divided by age, sex, and race) completed both tasks with a modified SCWT (Cards A and B). For the sample as a whole, there were moderate to strong correlations between children's SCWT interference scores and the two major MRT scores (raw composite scores, percentile scores).

These findings, in line with expectation, spoke to the use of the SCWT interference factor as an alternative measure of reading readiness. However, the data also suggested that the SCWT may be particularly useful as a screening tool - for example, in an initial multidimensional school screening assessing more general school readiness on a variety of cognitive and psychosocial variables or in a health clinic that screens for the possibility of learning disorders - rather than as a task providing a more in-depth analysis of the basic processes underlying reading (e.g., visual discrimination, phonological awareness, comprehension) that might be useful in designing more specific reading curricula. Thus, this study once again suggests that underlying process analysis broadly defined is a crucial component for complementing information obtained through achievement measures.

With respect to the larger issue at hand, namely, underlying process of the SCWT, we have also seen modification of strategies at later developmental stages, namely, in adults 70 years of age and older. For example, older adults typically emphasize boundaries between items, use sharper more articulate pronunciation, and employ non-linguistic assertions (e.g., "er," "ah") rather than the inserted linguistic phrases characteristic of younger subjects. Very striking in older adults' verbal strategies are shifts or changes in chunking organization, that is, typical chunking of items by two's or three's are lost under the pressure of Card C.

Further, through such analyses, we are beginning to get a sense that the cognitive functioning of older individuals may not be characterized solely as exhibiting signs of deterioration - a common finding in the literature - but rather as exhibiting signs of a reorganization of cognitive strategies. That is, older adults' use of less efficient strategies

(e.g., random counting on fingers, elaborate tongue movements and/or breathing patterns) appear to contribute to slower performance. In addition, subjects' verbalized strategies on completion of the task (e.g., the common response of "I'm not good at concentrating") have suggested exploration into illogical thought processes (particularly with respect to their metacognition where they typically tend to overestimate vastly, e.g., their total times and number of errors on all cards) rather than actual cognitive deficits as contributing factors to older adults' less than optimal cognitive functioning. Further support again a deterioration hypothesis of aging comes from our achievement-oriented findings that, relative to young adults, the total time score of older adults are impaired only on Card C.

Cross-cultural Similarities and Differences

In collaboration with colleagues in Japan [36], we have also assessed cross-cultural similarities and differences in processes underlying SCWT identification and serial organization toward directly determining the universality of the Stroop effect and processes underlying identification and serial organization. Because Japanese texts can be written in two ways with both writing styles existing side by side today - in Western style, that is, in horizontal rows from the top to the bottom of the page, or in traditional Japanese style, that is, in vertical columns from the right to the left side of the page—two different sets of test cards, one for reading stimulus items vertically (top to bottom) and a second for reading horizontally (left to right) were employed in both sociocultural contexts (U.S., Japan). Order of cards—horizontal versus vertical—was counter balanced and all Ss were undergraduates with an equal number of males and females in each sample.

With respect to achievement measures, our major findings included the following. First, data from both Japanese and American subjects revealed the shortest time to read items on Card A, somewhat longer to name items on Card B, and the longest time to name items on Card C (classic Stroop effect). Second, the interaction of spatial arrangement (horizontal vs. vertical) and sequence (horizontal-vertical vs. vertical-horizontal) was significant only for the U.S. sample. Third, total number of errors occurred, in both cultures, most frequently on Card C, less frequently on Card B, and least frequently on Card A.

For process-oriented measures, the following findings were obtained. First, total number of identification errors and total number of serial organization errors - similar to the previous finding on total number of errors - occurred, in both cultures, more frequently on Card C than they did on Card B than they did on Card A. Second, for total number of serial organization errors, there was an interaction between card and sex: among the Japanese, a greater number of serial organization errors in females than in males was observed, whereas among Americans these errors were more common in males than in females.

Third, analysis of specific identification errors revealed that for the Japanese but not for the Americans: (a) inappropriate color responses occurred more on Card B than on the other two cards; (b) inarticulate utterances occurred more on Card C than on the other two cards; and (c) both partly wrong responses and wholly wrong responses occurred most on Card C, less on Card B, and least on Card A. Fourth, analysis of specific serial organization errors again revealed that, again for the Japanese but not for the Americans, there was: (a) for insertion of color words, a main effect of cards with most instances occurring on Card C, fewer on Card B, and fewest on Card A, and an interaction between cards and spatial arrangement with a greater frequency of such errors occurring on Card C more often under the vertical condition than

under the horizontal condition; (b) for inserted linguistic words or phrases, a spatial arrangement effect with more instances occurring under the horizontal (vs. vertical) condition, a card effect with more occurring on Card C than on the other cards, and a card by spatial arrangement interaction with a spatial arrangement effect occurring only on Card C; and (c) for inserted nonlinguistic words or phrases, a card effect with most instances occurring on Card C, fewer on Card B, and fewest on Card A.

Significant cultural differences also occurred in nonverbal behaviors (cognitive strategies) on Card C. Nonverbal behaviors with body movement (e.g., pointing, rocking body, indexing line with finger, hard breathing) were observed more in American (vs. Japanese subjects), while static nonverbal behaviors (e.g., holding hands at one's side) were observed more often in Japanese (vs. American) subjects. Moreover, chunking by two's or three's was observed more often in American subjects whereas lip and facial movements were observed more often in the Japanese.

This study has demonstrated differential patterns of scores related to the distinction between underlying process (e.g., verbal errors, nonverbal strategies) and final achievement (e.g., total time), indicating that stressful organismic states - as evidenced by performance on Card C of the SCWT—are revealed most comprehensively through examination of both verbal performance and nonverbal behavior. Further, the findings have suggested that the SCWT has the potential to become an important tool for investigating cross-cultural differences in cognitive processes. For example, relative to U.S. college students (who, e.g., employ overt dramatic strategies to maintain serial organization), Japanese college students (who are relatively immobile during task performance) may either internalize nonverbal strategies earlier, employ other strategies not as yet identified, and/or manifest different strategies as least partly related to genetics (e.g., Lewis, Ramsay, Kawakami, [37], have documented that, during and following routine inoculation, Caucasian infants exhibit a more intense initial affective response and a longer latency to quiet than Japanese infants, who show a greater cortisol response).

Psycho- and Neuro-Pathology

We are also currently in the process of filming individual psychiatric outpatients with a range of different diagnoses (e.g., bipolar disorder, major depressive disorder, autism spectrum disorder, schizophrenia, traumatic brain injury) during SCWT performance. Once normative data from control individuals across the life span are coded, analyzed, and documented, we will ultimately be able to compare the performance of psychiatric/neuropsychiatric patients and normal controls on both achievement- and process-oriented SCWT measures toward determining the developmental status of a given individual. This is in line with Werner's [18] comparative orientation, which suggests that there are formal similarities in the behavior and experience of different individuals with respect to developmental status (e.g., children-older adults, those under the influence of drugs-normals, those affected by fatigue-normals, individuals in agrarian-industrialized societies). Relevant here, our analyses may demonstrate similarities between individuals with various psycho- and neuro-pathologies and those with less or more advanced developmental status (e.g., those with schizophrenia may exhibit strategy usage similar to younger children, while those with personality disorders may employ strategies more akin to those used by adolescents). In this way, we hope to strengthen the discriminative validity of the SCWT, making it much more useful as a tool for differential diagnosis based on the complementarity of process-vs. achievement-oriented analyses than has heretofore been the case.

Summary and Conclusions

The SCWT, which has been in existence for over 80 years, has been and will continue to be of considerable psychological interest to clinicians and researchers alike.

However, its usefulness has been somewhat limited in that, up until present, its singular use of achievement-oriented measures (e.g., total time per card, number of errors per card) does not finely discriminate among differing groups of individuals.

For example, although many [38] have recommended its use for the diagnosis of brain disorders, more recent research has documented that, relative to normal controls, those with focal frontal lesions and presumptive frontal dysfunction - including those with a range of psycho- and neuro-pathologies such as those with alcoholism, dementia, HIV encephalopathy, Huntington's disease, Parkinson's disease, prenatal alcohol exposure, schizophrenia, and traumatic brain injury - suffer from poor response inhibition, often measured by increased SCWT interference scores differing little from one another. Thus, by relying only on achievement-oriented measures, the SCWT does not do as good a job at discriminating among various conditions as it might were it to consider Werner's [18] distinction between process and achievement. That is, the discriminative validity of the SCWT might be increased were clinicians and researchers to complement traditional achievement measures with more process-oriented ones.

Toward this end, we have presented an overview and summary of our ongoing SCWT analyses, which have employed both traditional achievement measures and newer process-oriented measures. Drawing on elaborated Wernerian viewpoints [39], we have proposed a category system for the analysis of types of SCWT errors - specifically, those related to the underlying processes of identification and of serial organization - and nonverbal behaviors that serve as cognitive strategies for the identification and serial organization of SCWT items. Based on our data, we have also included in this system a variety of measures assessing subjects' metacognition (e.g., estimates of total time, number of errors, and cognitive strategies), which have been shown to impact some participants' performance, specifically older adults, in negative ways. Collectively, this more comprehensive scoring system will increase the SCWT's ability to discriminate among different patient conditions. Further, in line with our elaborated Wernerian viewpoint, we have attempted to demonstrate the ways in which this comprehensive system has applicability to a variety of other problem areas of interest such as cognitive development and cross-cultural psychology more generally.

Finally, the present work also addresses two longstanding concerns about the SCWT. That is, some have taken issue with the fact that there is no standard version of the SCWT with respect to its materials, administration, and scoring. However, we present data here on the Clark University Stroop Color-Word Test, which has been employed in a number of investigations for some time now, and recommend its continued usage on a much larger scale. Although the videotaping of subjects' SCWT performance might have at one point in time been considered daunting and cumbersome, the widespread proliferation of smartphones with filming capabilities has somewhat muted this criticism. Further, although a single study in the literature [40-42] has proposed the video-recorded SCWT as a new model of experimentally-induced anxiety, our participants typically acclimate easily to performing the task while being filmed regardless of whether the technology is older or newer.

Second, others have criticized existent SCWT work of which there is much on the basis that there has been no real attempt by any one individual to develop comprehensive SCWT norms. While we agree with the sentiment that this lofty goal has not been attempted by any one researcher alone, here we have attempted to draw the reader's attention to the numerous studies - past, present, and future - that have been generated against the backdrop of Werner's [18] comparative-developmental theory. Collectively, these studies have the potential to be integrated into a comprehensive set of SCWT norms with potential applicability to those in the normal population as well as to those who exhibit a variety of conditions (e.g., psychopathology, neuropathology) and to those who reside in different sociocultural contexts (broadly defined). In these ways, we also hope through this work to demonstrate to others the potential usefulness of grand theories that cut across differing aspects of persons and of environments toward maintaining and further developing a unified psychological science.

References

1. Ligon EM (1932) A genetic study of color naming and word reading. *American Journal of Psychology* 44: 103-121.
2. Stroop JR (1935a) The basis of Ligon's theory. *American Journal of Psychology* 47: 499-504.
3. Stroop JR (1935b) Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18: 643-662.
4. Stroop JR (1938) Factors affecting speed in serial verbal reactions. *Psychological Monographs* 50: 38-48.
5. Jensen AR, Rohwer WD (1966) The Stroop color-word test: A review. *Acta Psychologica* 25: 36-93.
6. MacLeod CM (1991) Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin* 109: 163-203.
7. Jones EG, Mendell LM (1999) Assessing the decade of the brain. *Science* 284: 739.
8. West R, Bell MA (1997) Stroop color-word interference and electroencephalogram activation: Evidence for age-related decline of the anterior attention system. *Neuropsychology* 11: 421-427.
9. Langleben DD, Monterosso J, Elman I, Ash B, Krikorian G, et al. (2006) Effect of methylphenidate on Stroop color-word task performance in children with attention deficit hyperactivity disorder. *Psychiatry Research*, 141: 315-320.
10. Adelman NE, Menon V, Blasey CM, White CD, Warsofsky IS, et al. (2002) A developmental fMRI study of the Stroop color-word task. *NeuroImage* 16: 61-75.
11. Leung H, Skudlarski P, Gatenby JC, Peterson BS, Gore JC (2000) An event-related functional MRI study of the Stroop color word interference task. *Cerebral Cortex* 10: 552-560.
12. Cattell JMK (1886) The time it takes to see and name objects. *Mind* 11: 63-65.
13. Strauss E, Sherman EMS, Spreen O (2006) A compendium of neuropsychological tests: Administration, norms, and commentary. (3rd edn) New York, Oxford University Press.
14. Lansbergen MM, Kenemans JL, van Engeland H (2007) Stroop interference and attention-deficit/hyperactivity disorder: A review and meta-analysis. *Neuropsychology* 21: 251-262.
15. Strauss GP, Allen DN, Jorgensen ML, Cramer SL (2005) Test-retest reliability of standard and emotional Stroop tasks: An investigation of color-word and picture-word versions. *Assessment* 12: 330-337.
16. Chung Y, Jeglic EL (2016) Use of the modified emotional Stroop task to detect suicidality in a college population. *Suicide and Life-Threatening Behavior: The official Journal of the American Association of Suicidality* 46: 55-66.
17. Werner H (1937) Process and achievement: A basic problem in education and developmental psychology. *Harvard Educational Review* 1: 353-368.
18. Werner H (1957) Comparative psychology of mental development. New York: International Universities Press.
19. Kaplan E (1988) The process approach to neuropsychological assessment. *Aphasiology* 2: 309-312.
20. Demick J, Harkins D (1999) Cognitive style and driving skills in adulthood: Implications for licensing of older adults. *International Association of Traffic Science and Safety Research* 23: 1-16.
21. Moering RG, Schinka JA, Mortimer JA, Bornstein Graves A (2004) Normative data for elderly African Americans for the Stroop color and word test. *Archives of Clinical Neuropsychology* 19: 61-71.
22. van der Elst W, van Boxtel MPJ, van Breukelen GJP, Jolles J (2006) The Stroop color-word test: Influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* 13: 62-79.
23. Seo EH, Lee DY, Choo IH, Kim SG, Kim KW, et al. (2008) Normative study of the Stroop color and word test in an educationally diverse elderly population. *International Journal of Geriatric Psychiatry* 23: 1020-1027.
24. Peña-Casanova J, Quiñones-Úbeda S, Gramut-Fombuena N, Quintana M, Aguilar M, et al. (2009) Spanish Multicenter normative studies (NEURONORMA project): Norms for the Stroop color-word interference test and the tower of London-Drexel. *Archives of Clinical Neuropsychology* 24: 413-429.
25. Homack S, Riccio CA (2004) A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of Clinical Neuropsychology* 19: 725-743.
26. National Institute for Literacy (2003) A child becomes a reader: Birth through preschool. Jessup, MD: EdPubs.
27. Kessler RC, Merikangas KR (2004) The National Comorbidity Survey Replication (NCS-R): Background and aims. *International Journal of Methods in Psychiatric Research* 13: 60-68.
28. Smith GJW, Klein GS (1953) Cognitive controls in serial behavior patterns. *Journal of Personality* 22: 188-213.
29. Demick J (2014) Group Embedded Figures Test (GEFT) manual (Rev. ed.) (including computerized test). Menlo Park, CA: Mind Garden.
30. Lashley KS (1951) The problem of serial order in behavior. In: Jeffress LA (ed.) *Cerebral mechanisms in behavior: The Hixon symposium*. New York: Random House.
31. Miller GA, Galanter E, Pribram KH (1960) *Plans and the structure of behavior*. New York: Henry Holt and Company.
32. Rand G, Wapner S, Werner H, McFarland JH (1963) Age differences in performance on the Stroop color-word test. *Journal of Personality* 31: 534-558.
33. Comalli PE, Wapner S, Werner H (1962) Interference effects of Stroop color-word test in childhood, adulthood, and aging. *Journal of Genetic Psychology* 100: 47-53.
34. Santostefano S (1978) *A biodevelopmental approach to clinical child psychology: Cognitive controls and cognitive control therapy*. New York: Wiley.
35. Demick, J (2016) Changing the face of parent development: Holistic/systems developmental therapy for parents. In: Kirkcaldy BD (ed.) *Personal enrichment in our lives: Psychotherapy in parenthood and beyond*. Torino, Italy: Edizioni Minerva Medica.
36. Toshima T, Demick J, Miyatani M, Ishii S, Wapner S (1996) Cross-cultural differences in processes underlying sequential cognitive activity. *Japanese Psychological Research* 38: 90-96.
37. Lewis M, Ramsay DS, Kawakami K (1993) Differences between Japanese infants and Caucasian American infants in behavior and cortisol response to inoculation. *Child Development* 64: 1722-1731.
38. Golden CJ (1976) Identification of brain disorders by the Stroop color and word test. *Journal of Clinical Psychology* 32: 654-658.
39. Wapner S, Demick J (1998) Developmental analysis: A holistic, developmental, systems-oriented perspective. In: Damon E, Lerner RM (eds.) *Handbook of child psychology: Vol. 1. Theoretical models of human development*, (5th edn) New York: Wiley.
40. Leite JR, Seabra MDV, Sartori VA, Andreatini R (1999) The video-recorded Stroop color-word test as a new model of experimentally-induced anxiety. *Progress in Neuro-Psychopharmacological and Biological Psychiatry* 23: 809-822.
41. Storey A, Demick J (2016) Stroop color-word test as a measure of reading readiness. Manuscript in preparation.

42. vanBoxtel MPJ, ten Tusscher MPM, Metsemakers JFM, Willems B, Jolles J (2001) Visual determinants of reduced performance on the Stroop color-

word test in normal aging individuals. *Journal of Clinical and Experimental Neuropsychology* 23: 620-627.