HHS Public Access

Author manuscript

Clin Neuropsychol. Author manuscript; available in PMC 2016 October 01.

Published in final edited form as:

Clin Neuropsychol. 2016 October ; 30(7): 1104–1117. doi:10.1080/13854046.2016.1188989.

Parkinson's Disease and the Stroop Color Word Test: Processing Speed and Interference Algorithms

S. Sisco[1], **E. Slonena**[2], **M.S. Okun**[3], **D. Bowers**[2,3], and **C.C. Price**[2]

[1]Veterans Affairs Illiana Health Care System, Department of Veterans Affairs

[2]Department of Clinical and Health Psychology, University of Florida, Gainesville, Florida

[3]Center for Movement Disorders and Neurorestoration, Department of Neurology, University of Florida, Gainesville, Florida

Abstract

OBJECTIVE—Processing speed alters the traditional Stroop calculations of interference. Consequently, alternative algorithms for calculating Stroop interference have been introduced to control for processing speed, and have done so in a multiple sclerosis sample. This study examined how these processing speed correction algorithms change interference scores for individuals with idiopathic Parkinson's Disease (PD, n= 58) and non-PD peers (n= 68).

METHOD—Linear regressions controlling for demographics predicted group (PD vs. non-PD) differences for Jensen's, Golden's, relative, ratio, and residualized interference scores. To examine convergent and divergent validity, interference scores were correlated to standardized measures of processing speed and executive function.

RESULTS—PD - non-PD differences were found for Jensen's interference score, but not Golden's score, or the relative, ratio, and residualized interference scores. Jensens' score correlated significantly with standardized processing speed but not executive function measures. Relative, ratio and residualized scores correlated with executive function but not processing speed measures. Golden's score did not correlate with any other standardized measures.

CONCLUSIONS—The relative, ratio, and residualized scores were comparable for measuring Stroop interference in processing speed-impaired populations. Overall, the ratio interference score may be the most useful calculation method to control for processing speed in this population.

Keywords

Stroop; interference; processing speed; Parkinson's; psychometrics

Introduction

The Stroop test (Stroop, 1935; Golden & Freshwater, 2002) is a measure of verbal processing speed and response inhibition that is widely used in present day

Correspondence: Shannon Sisco at Shannon.Sisco@va.gov, Psychology Service, 1900 E. Main Street 116B Danville, IL 61832, (217) 554-5413, or Catherine Price, Ph.D., cep23@phhp.ufl.edu; 352-273-5928.

All authors report no disclosures.

neuropsychological assessment. It consists of three timed trials. The first two trials measure the speed at which participants can read color words (red, green, blue; Word Reading, *W*) and name the color of blocks of ink in red, green, and blue (Color Naming, *C*). The third "Color-Word" (CW) trial is called the 'interference trial' and requires the examinee to name the ink color of printed words (red, green, blue) in which the ink color and word are incongruent. The speed of performance on the Color-Word trial is subtracted from the speed on the Color Naming trial to calculate interference. Interference occurs when the processing of one type of information disrupts the simultaneous processing of another type of information (Sarason, Pierce, & Sarason, 2014). In the Stroop test, participants are distracted from the goal of naming the ink color by the well-rehearsed drive to read the word itself. Melara and Algom (2003) suggest the words themselves are especially hard to ignore because they are salient, surprising, and correlated with the task's target information (i.e., the ink color).

Numerous formal theories have been proposed to explain the Stroop effect (see Cohen, Dunbar, & McClelland, 1990; Kornblum, Stevens, Whipple, & Requin, 1999; Logan, 1980; Phaf, van der Heijden, & Hudson, 1990; and Zhang, Zhang, & Kornblum, 1999 for prominent examples). One of the most well-established theories (Melara & Algom, 2003) proposes that the Stroop effect results from the combination of parallel excitations within ones' processing system. There is excitation to the target word, inhibitory excitation to the distractor, and excitation of memories from previous stimuli. The greater the incongruence between the target color and the distractor word, the longer the time required for the system to reach a resolution. There may be an analogous neuroanatomical system supporting this model: the anterior cingulate cortex (ACC) has been proposed to play a role in the Stroop by monitoring performance and detecting conflict between simultaneous, competing representations (i.e., target color and distractor word), and engaging the dorsolateral prefrontal cortex (DLPFC) to resolve those conflicts (Carter & van Veen, 2007; MacDonald, Cohen, Stenger, & Carter, 2000; Heyder, Suchan, & Daum, 2004).

Traditionally, Stroop interference has been calculated using the "Difference score" originally used by Stroop (1935) and recommended by Jensen (1965), calculating the simple difference between times on the Color-Word and Color Naming trials (i.e., CW − *C*). Golden (1978) later published a formula for clinical assessment which derived a predicted Color-Word score using the Color-Word trial and the two preliminary processing speed trials (i.e., (W × C)/(W + C) − CW).

Both approaches for calculating interference, however, have been reported to be problematic when used among patient populations for whom processing speed is impaired, such as patients with multiple sclerosis (MS; Denney & Lynch, 2009) or Parkinson's disease (PD; Karayanidis, 1989; Mahurin, 2013; Muslimović et al., 2009). Poorer interference scores among these groups may be due to the inappropriate derivation of interference based on the word reading and color naming trials, which primarily capture processing speed (Denney & Lynch, 2009). The resulting interference scores may be subsequently biased and clinically uninformative.

Recent research has highlighted alternative algorithms for calculating Stroop interference scores. Reportedly, these scores can be adjusted for confounding factors such as processing speed. For patients with MS, Denney and Lynch (2009) compared three such approaches – relative, ratio, and residualized scores - to the traditional Jensen's and Golden's scores on a computerized version of the Stroop. While both Jensen's and Golden's scores differentiated participants with MS from a control comparison group (in opposite directions: Jensen's scores suggested greater interference in controls; Golden's suggested greater interference in patients), no group differences were obtained when using the three alternative interference scores, even though significant group differences existed for processing speed. Moreover, Jensen and Golden's interference calculations were highly correlated with processing speed, while the alternative scores were not. All revealed moderately but non-significantly greater levels of interference among the patients compared to controls.

The present study applied a similar comparison of interference algorithms to individuals diagnosed with idiopathic Parkinson's disease (PD), for which cognitive slowing is also a hallmark deficit (Karayanidis, 1989; Mahurin, 2013). Like MS, Parkinson's disease involves characteristic reductions in processing speed believed to be associated with subcortical and white matter pathology (Zgaljardic, Borod, Foldi, & Matis, 2003). These reductions are likely associated with disruptions to frontal-subcortical pathways and their cortical and subcortical target areas (Mahurin, 2003; Zgaljardic et al., 2003). Recent studies (Brück, Portin, Lindell, Laihinen, Bergman, Haaparanta, Solin, & Rinne, 2001; Jokinen, Karrasch, Brück, Johansson, Bergman, & Rinne, 2013) have linked dopaminergic hypofunction in the caudate among PD patients with cognitive slowing on multiple tests, including the Stroop.

Executive dysfunction is also commonly observed in early PD (Jacobs, Marder, Cote, Sano, Stern, & Mayeux, 1995; Kehagia, Barker, & Robbins, 2010; Taylor & Saint-Cyr, 1995). Fronto-striatal circuits involving the ACC and DLPFC have been implicated in cognitive control (i.e. conflict identification and resolution) on the Stroop (Carter & van Veen, 2007; MacDonald, Cohen, Stenger, & Carter, 2000; Hetder, Suchan, & Daum, 2004). Henik and colleagues (1993) report an augmented facilitory effect (i.e., naming a word printed in congruent-colored ink faster than naming a block of X's in color ink) in early-onset PD and increased interference in late-onset PD. They suggest this as evidence of basal ganglia involvement in the inhibition of automatic cognitive processes. Relatedly, increased error rates on the Stroop in PD patients off medications have also been shown to resolve once medications are administered (Djamshidian, O'Sullivan, Lees, & Averbeck, 2011).

Impaired Stroop interference performance, however, may be related to the hallmark cognitive slowing in PD. PD patients' overall cognitive status has been shown to vary with the degree of cognitive slowing, and compared to other neurodegenerative disease populations, the extent of their cognitive slowing is disproportionate to general cognitive performance (Pate & Margolin, 1994; Price, Tanner, et al., 2016). Executive deficits in PD have been suggested to be due at least in part to deficits in processing speed resources (Cooper, Sagar, Tidswell, Jordan, 1994; Grossman, Zurif, Prather, Kalmanson, Stern, & Hurtig, 2002; Lee, Grossman, Morris, Stern, and Hurtig, 2003) versus deficits in internal cognitive control, particularly on the Stroop (Brown & Marsden, 1991; Woodward, Bub, & Hunter, 2002). Similarly, Verhaeghen (2011) suggested that neurological changes associated

with cognitive slowing – dopamine loss and pathology within the white matter and subcortical gray matter – may be substantially responsible for observed age-related reductions in executive control on the Stroop and other measures.

Within a large clinical movement disorder program, we have observed that processing speed impairments in the two preliminary trials of the Stroop test are common, and render the traditionally-derived (i.e., Jensen's, Golden's) interference scores difficult to interpret. Like Denney and colleagues (2009), we contrasted traditional interference measures (Jensen's, Golden's) to alternative interference measures (relative, ratio, residualized) in a sample of patients with early-stage PD and healthy non-PD peers. Individuals with idiopathic PD were expected to demonstrate lower scores relative to non-PD peers on the processing speed trials of the Stroop test (i.e., Word Reading, Color Naming) with the traditional interference scores showing greater interference for non-PD peers using Jensen's score, and greater interference for patients using Golden's score. The relative, ratio, and residualized scores, in contrast, were expected to demonstrate no significant group differences between PD and non-PD groups. . Secondly, we examined the convergent validity of the relative, ratio, and residualized interference scores with other measures of executive functioning, and also examined divergent validity of the scores with other measures of processing speed.

Methods

Participants

This study was approved by the University of Florida Institutional Review Board and followed the Principles of the Declaration of Helsinki. For this retrospective investigation, we reviewed data from individuals with PD ($N = 92$) and non-PD age matched peers ($N = 102$) who had signed consent forms allowing their data to be used for research purposes. The cohort of idiopathic PD outpatients and research participants were drawn from an outpatient, university-affiliated movement disorders clinic (MDC) and an ongoing NIH research investigation. A movement disorder neurology specialist within the Center for Movement Disorders and Neurorestoration (CMDNR) at the University of Florida (UF) completed all diagnostics for idiopathic PD, with PD diagnosis guided by the United Kingdom PD Society Brain Research Criteria (Gibb & Lees, 1988). From the UF CMDNR research database and affiliated NINDS funded investigations, a total of 194 participants in the MDC database were reviewed to obtain the 136 (58 PD, 68 non-PD) participants who met inclusion criteria. Study inclusion criteria for PD required Hoehn and Yahr scale range of 1–3, being "on" medication at time of testing, and no-dementia per a Dementia Rating Scale score ≥ 130 raw and MMSE ≥ 26, no history of deep brain stimulation (DBS), and no other neurological disorder history. Color-blindness was not formally assessed, but participants reporting it were excluded from the sample. Non-PD "healthy" peers were recruited from newspaper advertisements and community memory screenings, and were involved in separate federally funded research investigations.

Measures

Golden's (1978) version of the Stroop Test was used in all three studies of which the present dataset is comprised. This version provides paper stimuli for each trial: in the first, columns

of the words “red,” “blue,” and “green” are printed in black ink (Word Reading; *W*); in the second, columns of “xxxx” are printed in red, blue, or green ink (Color Naming; *C*); in the third, the words “red,” “blue,” and “green” are printed in a colored ink (red, blue or green) that does not match the word (Color-Word; CW). Participants read each page aloud as quickly as possible for 45 seconds and receive a score for each trial representing the number of items correctly read aloud. Interference scores are calculated from performance on these three trials.

Five approaches to calculating interference (two “traditional,” three “alternative”) were employed and compared in the present study, replicating the approaches examined by Denney and colleagues (2009). The formulae for the approaches are illustrated in Table 2. Jensen’s ($C - CW$; 1965) and Golden’s ($[(W \times C)/(W + C)] - CW$; 1978) scores, the traditionally-used calculations for interference, were compared with the “relative” ($[(C - CW)/C] \times 100$; Vitkovich et al., 2002; Macniven et al., 2008), “ratio” (CW/C; Lansbergen et al., 1999), and “residualized” (Capitani et al., 1999; Denney et al., 2009) scores which have been shown in recent research to better control for differences in processing speed. The residualized scores were computed by regressing CW on *C*, obtaining the unstandardized residual scores from this regression for each subject, and subtracting the unstandardized residual scores from the overall sample mean for CW. In other words, unstandardized residual scores represent the difference between a subject’s actual CW score and his or her CW score as predicted by the regression model. Thus, unstandardized residual scores represent the variance in CW not related to *C*, and thus presumably more purely related to interference.

Participants in each of the studies from which this dataset was comprised completed the Stroop as part of a larger neuropsychological protocol. With the exception of the Wechsler Adult Intelligence Scales – 3rd Edition (WAIS-III) subtests, all standardized scores for measures of processing speed and executive function were obtained using Heaton’s revised norms for the Halstead-Reitan assessment battery (Psychological Assessment Resources, 2004). Additional standardized measures of processing speed and executive function were used to examine convergent and divergent validity of the Stroop scores. These included:

Processing Speed measures

WAIS-III Digit-Symbol Coding (Wechsler, 1997): Requires rapid visual scanning and matching of symbols to numbers; dependent variable (DV) = total correct.

WAIS-III Symbol Search subtest (Wechsler, 1997): Requires rapid matching of abstract symbols; DV = total correct.

Trail Making Test, Part A (Reitan, 1992; Lezak, Howieson, & Loring, 2004) A measurement of visuomotor speed requiring the rapid sequencing of numbers shown randomly across the page; DV = time to completion, standardized score (Psychological Assessment Resources, 2004).

The scores for each of these measures were converted to a common *z*-score metric in order to be combined into a composite index of processing speed. The Processing Speed Composite was calculated as the average of the above measures' *z*-scores.

Executive Function-Related measures

Trail Making Test Part B (Reitan, 1992): Involves rapid line sequencing of alternating letters and numbers shown randomly on a page; DV = total time to completion, standardized score (Psychological Assessment Resources, 2004). Since the aim was to use this as a comparative executive function measure for convergent validity of the Stroop interference scores, we derived a "Trails B ratio" score by dividing Trails B raw scores by Trails A raw scores to arrive at a more "pure" executive function measure (see Lamberty, Putnam, Chatel, Bielauskas, & Adams, 1994).

Controlled Oral Word Association test (COWA) lexical fluency (FAS; (Tombaugh, Kozak, & Rees, 1999): Requires generating words beginning with a specific letter within 60 seconds, excluding numbers and proper nouns; DV = total words minus errors, standardized score (Psychological Assessment Resources, 2004).

Wisconsin Card Sorting Test (WCST; Berg, 1948; Grant & Berg, 1948; Heaton & PAR, 2003): Requires attention to changing rules in order to solve a problem; DV = total categories completed, standardized score (Psychological Assessment Resources, 2004).

Additional measures—Several additional measures were used to characterize the sample. All participants completed the Mini Mental State Exam (MMSE; Folstein, Folstein, & McHugh, 1975), a brief measure of cognitive status, and the Geriatric Depression Scale (Yesavage, Brink, Rose, Lum, Huang, Adey, & Leirer, 1983), a measure of depressive symptoms. Participants also completed the Unified Parkinson's Disease Rating Scale (UPDRS; Fahn & Elton, 1987) as a neurological assessment of functioning in patients with Parkinson's disease. The UPDRS total score while "on" medication at the time point closest to the neuropsychological evaluation served as the dependent variable for Model 2. The non-PD peers in this study completed the UPDRS for comparison purposes.

Analysis Plan

Demographic and other characteristics of the patient and control groups were compared using independent T-tests (Table 1). Distributions of all dependent variables were inspected and verified to be normally distributed for both patient and control groups. Models 1 and 2 employed linear regressions to examine how Stroop scores are predicted by disease status (PD vs control, Model 1) and UPDRS score (PD group only, Model 2), after controlling for age, sex, education, and depressive symptoms. Lastly, using a subset of the sample for which data were available, bivariate correlations were estimated between Stroop scores and other standardized measures of processing speed (WAIS-III Digit Symbol and Symbol Search (Arnau & Thompson, 2000; van der Heijden & Donders, 2003), Trail Making Test A (Crowe, 1998; Sanchez-Cubillo, Perianez, Adrover-Roig, Rodriguez-Sanchez, Rios-Lago,

Tirapu, & Barcelo, 2009)) and executive function (COWA letter fluency (Jurado & Rosselli, 2007; Phillips, 1997), WCST (Miyake, Friedman, Emerson, Witzki, & Howerter, 2000; Burgess, Alderman, Evans, Emslie, & Wilson, 1998), and Trails B ratio score (Arbuthnott & Frank, 2000; Lamberty, Putnam, Chatel, Bieliauskas, & Adams, 1994; Llinàs-Reglà, Vilalta-Franch, López-Pousa, Calvó-Perxas, Rodas, & Garre-Olmo, 2015; Sanchez-Cubillo et al., 2009)) to further examine whether 1) Stroop processing speed scores show convergent validity with other standardized processing speed measures and divergent validity with measures of executive function, and whether 2) Stroop interference scores show convergent validity with other executive function measures, and divergent validity with standardized measures of processing speed.

Results

Sample characteristics

Groups did not differ in age, educational attainment, or depression score (Table 1). The PD group included more males than the non-PD peer group (75.9% vs. 55.9% respectively; p = 0.02), consistent with several studies suggesting PD may be more prevalent among men (e.g. Baldereschim Di Carlo, Rocca, Vanni, Maggi, Perissinotto, & Inzitari, 2000; Van Den Eeden, Tanner, Bernstein, Fross, Leimpeter, Bloch, & Nelson, 2003; Wirdefeldt, Adami, Cole, Trichopoulos, & Mandel, 2011; Wooten, Currie, Bovbjerg, Lee, & Patrie, 2004). PD group MMSE scores were also slightly lower than non-PD peers (28.86 vs. 29.26, p = 0.03). Models 1 and 2 adjusted for effects of age, sex, education and depressive symptoms by including them as predictors in each regression. The effect of age is reported in Table 2 for all Stroop scores. Effects of sex and education were consistently non-significant, with the exception of Jensen's Difference score where there was an effect of sex in Model 1 only (β = .21, p = .03).

Processing Speed Differences

For the Stroop processing speed scores, in both Models 1 and 2 (using group membership and UPDRS score as predictors, respectively), significant differences were found for the Word Reading and Color Naming raw scores, with PD scoring lower than non-PD peers on both measures, and UPDRS scores associated with processing speed scores among patients (Tables 2 and 3). The two groups also differed on the Processing Speed Composite, with PD scoring lower than non-PD peers (Table 2).

Interference Score Differences

For Model 1, in which the five interference scores were predicted by group membership (controlling for age, sex, education, and depression) in separate linear regressions, group membership significantly predicted differences in Jensen's interference score, but not Golden's, relative, ratio, or residualized scores. Follow up tests compared the effect sizes for Jensen's score with other interference scores (Steiger, 1980; Lee & Preacher, 2013), and found that the effect sizes did not significantly differ, although for the ratio and residualized scores the differences approached significance (p = .07). For Model 2, in which each interference score was separately predicted by UPDRS score in the Parkinson's disease sample only, the same pattern held as in Model 1: that is, only Jensen's score was predicted

by UPDRS. Examining the raw scores for each interference measure (i.e., not adjusting for age, sex, education, or depressive symptoms), patients and non-PD peers significantly differed for Jensen, Golden's, relative, and residualized scores. Only for ratio scores was there no group difference before accounting for other demographic factors.

Convergent and Divergent Validity

The Stroop processing speed and interference scores were correlated with other standardized measures of speed and executive function to help determine convergent and divergent validity. The Stroop processing speed scores (Word Reading, Color Naming) correlated positively with each of the processing speed tasks (WAIS-III Digit-Symbol Coding, Symbol Search; Color Naming correlated with Trail Making Test A; Table 4).

Among the interference scores, Jensen's correlated positively with all three standardized processing speed measures and did not correlate with any measures of executive function. Golden's interference score did not correlate with any measures of processing speed or executive function. The relative and residualized interference scores were both positively and significantly correlated with the Trails B ratio score, to a similar degree (r=.27, p<.05). The ratio score revealed a somewhat stronger (not statistically different) correlation with the Trails B ratio score (r=.39, p<.01), and was also significantly correlated with WCST category score (r=−.27, p<.05) such that greater ratio interference scores were associated with fewer categories on the WCST.

Discussion

The purpose of this study was to examine the utility of three "alternative" calculations of Stroop interference, compared to two traditionally-used scores, in controlling for the confounding effects of processing speed among individuals with known processing speed deficits. Overall, this study added to the body of evidence suggesting these three methods – relative, ratio, and residualized scores – adequately control for confounds associated with processing speed. That is, they were uncorrelated with the Word Reading and Color Naming scores, uncorrelated with other standardized measures of processing speed, correlated with classical measures of executive function evaluating cognitive flexibility and response inhibition in particular, did not yield differences in scores for patients with processing speed impairments versus healthy non-PD peers, and did not yield differences in scores for patients with varying degrees of disease severity. While all three alternative scoring approaches performed similarly, the ratio interference score was slightly better-correlated with the Trails B ratio score, was additionally correlated with the second measure of executive function (WCST), and yielded no group differences for patients versus non-PD peers even without controlling for effects of age, education, sex, and depression symptoms. There has also been more extensive validation of a ratio score to control for processing speed in another well-established measure of executive function (Trails B ratio score; Lamberty et al., 1994; Arbuthnott & Frank, 2000), lending this particular scoring approach extra support. To our knowledge, testing norms have not been developed for the relative, ratio or residualized interference scores.

A practical consideration is the ease of clinical application for each of these scores. The relative ([(*C*− CW)/*C*] × 100) and ratio (*CW*/C) could readily be calculated by hand from a single patient's Stroop scores. The residualized Stroop score, on the other hand, would more likely require statistical software and database of scores in order to be calculated for a single patient, and therefore seems more appropriate for application in research.

This study also added to evidence suggesting that Jensen's Difference score is not a valid measure of interference, and likely is confounded by processing speed: it correlated with other standard processing speed measures, did not correlate with other measures of executive function, and yielded group differences in scores for patients versus non-PD peers, as well as for patients with varying degrees of disease severity.

Regarding Golden's interference score, this study found different results than a prior, similar study: Denney and Lynch (2009) reported differences for Golden's interference score in a sample of people with MS compared to non-PD peers, suggesting it may have been confounded by impairments in processing speed. Our findings did not find differences in Golden's score between Parkinson's patients and non-PD peers. This could be attributable to a difference of method: this study used the original paper version of the Stroop, and Denney and Lynch used a computerized version. A recent study (Penner et al., 2012) comparing the two versions found that while both had high test-retest reliability and could detect interference, the original and computerized task scores did not correlate, and interference effects appeared to be diminished in the computerized version.

This study has limitations. Not all participants completed exactly the same neuropsychological measures beside the Stroop, so correlating Stroop scores with other speed and executive function measures had to be done using parts of the full sample. Also, the tasks used to examine convergent validity measured aspects of executive function that are not perfectly comparable to interference. For example the COWA task measures verbal generation, which may be considered quite different from inhibitory abilities. This may explain why interference scores were uncorrelated with COWA scores. Additionally, the participants were tested 'on' dopaminergic medications, which can improve cognitive function in some patients but may make them worse in others (see Vaillancourt, Schonfeld, Kwak, Bohnen, & Seidler, 2013). Use of the data provided for normative PD reference is therefore cautioned. Finally, there is evidence that in addition to speed of processing deficits, degraded color vision in PD (Archibald, Clarke, Mosimann, & Burn, 2009) may also account for patient-control differences on the Stroop (e.g., Ben-David & Schneider, 2010; Ben-David, Tewari, Shakuf, & Van Lieshout, 2014). Although participants reporting color-blindness were excluded from this sample, this study did not formally assess disease-related changes in color vision, and thus cannot speak to whether such changes also played a role in group differences on the Stroop.

These findings may be useful to clinicians using the Stroop in patient groups with known processing speed impairments. Jensen's Difference score is not recommended as a measure of interference in such populations. Golden's interference score may not have produced group differences in this study, but performed in an unexpected direction in a previous study (Denney & Lynch, 2009). This study difference suggests caution when applying Golden's

algorithm. By contrast, the relative, ratio, and residualized scores appear to adequately control for the effects of processing speed so that interference can be ascertained. The sizes of their effects in predicting group differences were not significantly different from that of Jensen's; their relative advantage should thus not be overstated. Overall, these metrics performed similarly in terms of eliminating group differences in interference, demonstrating convergent validity with other executive function measures and demonstrating divergent validity from standardized processing speed measures. Among the three interference scores, the ratio interference score performed slightly better than the others on several dimensions, is practically applicable for clinical use as it can be easily calculated with a single patient's Stroop scores, and there is prior empirical validation for using this calculation in a similar well-established measure of executive function. Thus, while all three alternative interference scores show acceptable control for processing speed on the Stroop, the ratio interference score may be the most appropriate and clinically useful choice.

Acknowledgments

This study was partially supported by NINDS K23NS60660 (CP), The University of Florida Neurorestoration and Movement Disorder Center, and the NIH General Clinical Research Center Grant # UL1 TR000064.

References Cited

Arbuthnott K, Frank J. Trail making test, part B as a measure of executive control: validation using a set-switching paradigm. Journal of clinical and experimental neuropsychology. 2000; 22(4):518–528. [PubMed: 10923061]

Archibald NK, Clarke MP, Mosimann UP, Burn DJ. The retina in Parkinson's disease. Brain. 2009:awp068.

Arnau RC, Thompson B. Second order confirmatory factor analysis of the WAIS-III. Assessment. 2000; 7(3):237–246. [PubMed: 11037391]

Baldereschi M, Di Carlo A, Rocca WA, Vanni P, Maggi S, Perissinotto E, … Inzitari D. Parkinson's disease and parkinsonism in a longitudinal study: Two-fold higher incidence in men. Neurology. 2000; 55(9):1358–1363. [PubMed: 11087781]

Ben-David BM, Schneider BA. A sensory origin for color-word Stroop effects in aging: Simulating age-related changes in color-vision mimics age-related changes in Stroop. Aging, Neuropsychology, and Cognition. 2010; 17(6):730–746.

Ben-David BM, Tewari A, Shakuf V, Van Lieshout PH. Stroop effects in Alzheimer's disease: Selective attention speed of processing, or color-naming? A meta-analysis. Journal of Alzheimer's Disease. 2014; 38(4):923–938.

Berg EA. A simple objective test for measuring flexibility in thinking. Journal of General Psychology. 1948; 39:15–22. [PubMed: 18889466]

Brown RG, Marsden CD. Dual task performance and processing resources in normal subjects and patients with Parkinson's disease. Brain. 1991; 114(1):215–231. [PubMed: 1998883]

Brück A, Portin R, Lindell A, Laihinen A, Bergman J, Haaparanta M, … Rinne JO. Positron emission tomography shows that impaired frontal lobe functioning in Parkinson's disease is related to dopaminergic hypofunction in the caudate nucleus. Neuroscience letters. 2001; 311(2):81–84. [PubMed: 11567783]

Capitani EM, Laiacona MM, Barbarotto RM, Cossa FM. How can we evaluate interference in attentional tests? A study based on bi-variate non-parametric tolerance limits. Journal of Clinical and Experimental Neuropsychology. 1999; 21:216–228. [PubMed: 10425518]

Carter CS, Van Veen V. Anterior cingulate cortex and conflict detection: an update of theory and data. Cognitive, Affective, & Behavioral Neuroscience. 2007; 7(4):367–379.

Cohen JD, Dunbar K, McClelland JL. On the control of automatic processes: a parallel distributed processing account of the Stroop effect. Psychological review. 1990; 97(3):332. [PubMed: 2200075]

Cooper JA, Sagar HJ, Tidswell P, Jordan N. Slowed central processing in simple and go/no-go reaction time tasks in Parkinson's disease. Brain. 1994; 117(3):517–529. [PubMed: 8032862]

Crowe SF. The differential contribution of mental tracking, cognitive flexibility, visual search, and motor speed to performance on parts A and B of the Trail Making Test. Journal of clinical psychology. 1998; 54(5):585–591. [PubMed: 9696108]

Denney DR, Lynch SG. The impact of multiple sclerosis on patients' performance on the Stroop Test: processing speed versus interference. J Int Neuropsychol Soc. 2009; 15(3):451–458. DOI: 10.1017/S1355617709090730 [PubMed: 19402931]

Djamshidian A, O'Sullivan SS, Lees A, Averbeck BB. Stroop test performance in impulsive and non impulsive patients with Parkinson's disease. Parkinsonism & related disorders. 2011; 17(3):212–214. [PubMed: 21247790]

Fahn, S.; Elton, RL. Recent developments in Parkinson's disease. Florham Park. New York: Macmillan; 1987. Unified rating scale for Parkinson's disease; p. 153-163.

Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975; 12(3):189–198. 0022-3956(75)90026-6. [PubMed: 1202204]

Golden CJ. A Manual for the Clinical and Experimental Use of the Stroop Color and Word Test. 1978

Golden, CJ.; Freshwater, SM. The Stroop Color and Word Test: A Manual for Clinical and Experimental Uses. Chicago, IL: Stoelting; 2002.

Grant DA, Berg E. A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. Journal of experimental psychology. 1948; 38(4): 404. [PubMed: 18874598]

Grossman M, Zurif E, Lee C, Prather P, Kalmanson J, Stern MB, Hurtig HI. Information processing speed and sentence comprehension in Parkinson's disease. Neuropsychology. 2002; 16(2):174. [PubMed: 11949709]

Heyder K, Suchan B, Daum I. Cortico-subcortical contributions to executive control. Acta psychologica. 2004; 115(2):271–289. [PubMed: 14962404]

Heaton, RK. Psychological Assessment Resources. Wisconsin Card Sorting Test: Computer Version 4. Research Edition. PAR Psychological Assessment Resources; Ind., Lutz, Florida: 2003. www.parinc.com

Hsieh YH, Chen KJ, Wang CC, Lai CL. Cognitive and motor components of response speed in the stroop test in Parkinson's disease patients. The Kaohsiung journal of medical sciences. 2008; 24(4):197–203. [PubMed: 18424356]

Jacobs DM, Marder K, Cote LJ, Sano M, Stern Y, Mayeux R. Neuropsychological characteristics of preclinical dementia in Parkinson's disease. Neurology. 1995; 45(9):1691–1696. [PubMed: 7675228]

Jensen AR. Scoring the Stroop test. Acta Psychologica. 1965; 24:398–408. [PubMed: 5841721]

Jokinen P, Karrasch M, Brück A, Johansson J, Bergman J, Rinne JO. Cognitive slowing in Parkinson's disease is related to frontostriatal dopaminergic dysfunction. Journal of the neurological sciences. 2013; 329(1):23–28. [PubMed: 23561982]

Jurado MB, Rosselli M. The elusive nature of executive functions: a review of our current understanding. Neuropsychology review. 2007; 17(3):213–233. [PubMed: 17786559]

Karayanidis F. Parkinson's disease: A conceptualization of neuropsychological deficits within an information-processing framework. Biological psychology. 1989; 29(2):149–179. [PubMed: 2697379]

Kehagia AA, Barker RA, Robbins TW. Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. The Lancet Neurology. 2010; 9(12):1200–1213. [PubMed: 20880750]

Kornblum S, Stevens GT, Whipple A, Requin J. The effects of irrelevant stimuli: 1. The time course of stimulus–stimulus and stimulus–response consistency effects with Stroop-like stimuli, Simon-like

tasks, and their factorial combinations. Journal of Experimental Psychology: Human Perception and Performance. 1999; 25(3):688.

Lamberty GJ, Putnam SH, Chatel DM, Bieliauskas LA, Adams KM. A Preliminary Report. Cognitive and Behavioral Neurology. 1994; 7(3):230–234.

Lansbergen MM, Kenemans JL, van Engeland H. Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. Neuropsychology. 2007; 21(2):251. [PubMed: 17402825]

Lee C, Grossman M, Morris J, Stern MB, Hurtig HI. Attentional resource and processing speed limitations during sentence processing in Parkinson's disease. Brain and Language. 2003; 85(3): 347–356. [PubMed: 12744946]

Lezak, MD.; Howieson, DB.; Loring, DW. Neuropsychological Assessment. 4. New York: Oxford University Press; 2004. p. 371-374.

Llinàs-Reglà J, Vilalta-Franch J, López-Pousa S, Calvó-Perxas L, Rodas DT, Garre-Olmo J. The Trail Making Test Association With Other Neuropsychological Measures and Normative Values for Adults Aged 55 Years and Older From a Spanish-Speaking Population-Based Sample. Assessment. 2015 1073191115602552.

Logan GD. Attention and automaticity in Stroop and priming tasks: Theory and data. Cognitive psychology. 1980; 12(4):523–553. [PubMed: 7418368]

MacDonald AW, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 2000; 288(5472):1835–1838. [PubMed: 10846167]

Macniven JAB, Davis C, Ho MY, Bradshaw CM, Szabadi E, Constantinescu CS. Stroop performance in multiple sclerosis: information processing, selective attention, or executive functioning? Journal of the International Neuropsychological Society. 2008; 14(05):805–814. [PubMed: 18764975]

Mahurin RK. Frontal-subcortical determinants of processing speed in Parkinson's disease. Information processing speed in clinical populations. 2013:195.

Melara RD, Algom D. Driven by information: a tectonic theory of Stroop effects. Psychological review. 2003; 110(3):422. [PubMed: 12885110]

Muslimović D, Post B, Speelman JD, De Haan RJ, Schmand B. Cognitive decline in Parkinson's disease: a prospective longitudinal study. Journal of the International Neuropsychological Society. 2009; 15(03):426–437. DOI: 10.1017/S1355617709090614 [PubMed: 19402929]

Lee, IA.; Preacher, KJ. Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. 2013 Sep. Available from http://quantpsy.org

Pate DS, Margolin DI. Cognitive slowing in Parkinson's and Alzheimer's patients Distinguishing bradyphrenia from dementia. Neurology. 1994; 44(4):669–669. [PubMed: 8164823]

Penner IK, Kobel M, Stöcklin M, Weber P, Opwis K, Calabrese P. The Stroop task: comparison between the original paradigm and computerized versions in children and adults. Clinical Neuropsychology. 2012; 26(7):1142–53. DOI: 10.1080/13854046.2012.713513

Phaf RH, Van der Heijden AHC, Hudson PT. SLAM: A connectionist model for attention in visual selection tasks. Cognitive psychology. 1990; 22(3):273–341. [PubMed: 2376113]

Phillips LH. Do "frontal tests" measure executive function? Issues of assessment and evidence from fluency tests. Methodology of frontal and executive function. 1997:191–213.

Price CC, Tanner JJ, Nguyen PT, Schwab NA, Mitchell S, Slonena E, Brumback B, Okun MS, Mareci TH, Bowers D. Gray and white matter contributions to cognitive frontostriatal deficits in non-demented Parkinson's disease. PLOS One. 2016 in press.

Psychological Assessment Resources. Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults, professional manual. 2004.

Reitan, RM. Trail Making Test: Manual for administration and scoring. Reitan Neuropsychology Laboratory; 1992.

Sanchez-Cubillo I, Perianez JA, Adrover-Roig D, Rodriguez-Sanchez JM, Rios-Lago M, Tirapu JEEA, Barcelo F. Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. Journal of the International Neuropsychological Society. 2009; 15(03):438–450. [PubMed: 19402930]

Sarason, IG.; Pierce, GR.; Sarason, BR. Cognitive interference: Theories, methods, and findings. Routledge; 2014.

Steiger JH. Tests for comparing elements of a correlation matrix. Psychological Bulletin. 1980; 87(2): 245–251.

Stroop JR. Studies for interference in serial verbal reactions. Journal of Experimental Psych. 1935; 18(6):643–662.

Taylor AE, Saint-Cyr JA. The neuropsychology of Parkinsons-disease. Brain and cognition. 1995; 28(3):281–296. [PubMed: 8546855]

Tombaugh TN, Kozak J, Rees L. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. Arch Clin Neuropsychol. 1999; 14(2):167–177. [PubMed: 14590600]

Vaillancourt DE, Schonfeld D, Kwak Y, Bohnen NI, Seidler R. Movement Disorders. 2013; 28(14): 1920–1929. [PubMed: 24123087]

Van Den Eeden SK, Tanner CM, Bernstein AL, Fross RD, Leimpeter A, Bloch DA, Nelson LM. Incidence of Parkinson's disease: variation by age, gender, and race/ethnicity. American journal of epidemiology. 2003; 157(11):1015–1022. [PubMed: 12777365]

van der Heijden P, Donders J. A Confirmatory Factor Analysis of the WAIS—III in Patients with Traumatic Brain Injury. Journal of clinical and experimental neuropsychology. 2003; 25(1):59–65. [PubMed: 12607172]

Vitkovitch M, Bishop S, Dancey C, Richards A. Stroop interference and negative priming in patients with multiple sclerosis. Neuropsychologia. 2002; 40(9):1570–1576. [PubMed: 11985838]

Wechsler, D. Wechsler Adult Intelligence Scale. 3. San Antonio, Texas: The Psychological Corporation; 1997.

Wirdefeldt K, Adami HO, Cole P, Trichopoulos D, Mandel J. Epidemiology and etiology of Parkinson's disease: a review of the evidence. European journal of epidemiology. 2011; 26(1):1–58. [PubMed: 20845063]

Woodward TS, Bub DN, Hunter MA. Task switching deficits associated with Parkinson's disease reflect depleted attentional resources. Neuropsychologia. 2002; 40(12):1948–1955. [PubMed: 12207992]

Wooten GF, Currie LJ, Bovbjerg VE, Lee JK, Patrie J. Are men at greater risk for Parkinson's disease than women? Journal of Neurology, Neurosurgery & Psychiatry. 2004; 75(4):637–639.

Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: a preliminary report. Journal of psychiatric research. 1983; 17(1):37–49. [PubMed: 7183759]

Zgaljardic DJ, Borod JC, Foldi NS, Mattis P. A review of the cognitive and behavioral sequelae of Parkinson's disease: relationship to frontostriatal circuitry. Cognitive and behavioral neurology. 2003; 16(4):193–210. [PubMed: 14665819]

Zhang HH, Zhang J, Kornblum S. A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. Cognitive psychology. 1999; 38(3):386–432. [PubMed: 10328858]

Table 1

Demographic and self-report characteristics of the patient and control groups

	PD patients (N = 58)		Non-PD peers (N = 68)		t(*df*)	*p*
	Range	**M (SD)**	**Range**	**M (SD)**		
Age (years)	60–79	67.91 (5.10)	60–83	69.60 (5.53)	1.77 (124)	.079
Education (years)	10–22	15.79 (2.95)	9–24	15.61 (3.02)	−.43 (124)	.672
MMSE score	26–30	28.86 (1.13)	26–30	29.26 (0.91)	2.22 (124)	.029*
GDS score	0–26	5.13 (5.54)	0–14	4.4 (3.41)	−.897 (99.67)	.372
UPDRS [1]	8–70	29.55 (14.41)	0–6	3.19 (2.04)	−9.26 (75)	<.001**

*$p < 0.05$

PD = Parkinson's Disease

MMSE = Mini Mental State Exam

GDS = Geriatric Depression Scale

UPDRS – Unified Parkinson's Disease Rating Scale

[1]UPDRS scores were collected for a subset of participants ($N = 77$; non-PD = 26, PD = 51)

Table 2

Patient and Control Scores for Tests of Processing Speed and Interference[1]

		Patients (*N*=58)		Non-PD peers (*N*=68)			
Measure	**Formula**	**Range Min Max**	**M (SD)**	**Range Min Max**	**M (SD)**	**t (*df*)**	***p***
Processing Speed Composite		−1.32 – .66	−.22 (.58)	−,02 – 1.86	.73 (.55)	6.54 (61)	<.001
Stroop Processing Speed							
Word reading	W	61 124	87.21 (14.35)	59 129	92.25 (12.40)	2.12 (124)	.04*
Color naming	C	40 95	60.66 (12.00)	42 90	64.84 (10.70)	2.07 (124)	.04*
Stroop Interference							
Difference score	C CW	7 60	28.26 (10.51)	15 59	33.24 (9.08)	2.85 (124)	.005**
Golden's score	[(W × C)/(W + C)] − CW	−22.90 14.24	−3.22 (7.61)	−22.73 12.60	−6.22 (6.89)	−2.32 (124)	.02*
Relative score	[(C CW)/C] × 100	12.28 77.27	46.13 (12.84)	27.78 77.78	51.11 (10.57)	2.39 (124)	.02*
Ratio score	(C/CW) × 10	11.40 44.00	19.71 (5.31)	13.85 45.00	21.69 (6.14)	1.92 (124)	.06
Residualized score		10.27 49.31	29.52 (7.62)	15.00 46.93	32.19 (6.79)	2.08 (124)	.04*

*$p < 0.05$

**$p < 0.01$

M = mean; SD = standard deviation; Min = minimum, Max = maximum

Raw scores are the number of items correctly read aloud.

[1]Reported values and significance tests are associated with raw data, not adjusted for age, sex, education, or depression scores.

Clin Neuropsychol. Author manuscript; available in PMC 2016 October 01.

Table 3

Predictor Effects for Stroop Scores Regressed with Model 1 and Model 2

	Model 1[1]		Model 2[2]	
Measure	**Group β (p)**	**Age β (p)**	**UPDRS β (p)**	**Age β (p)**
Processing speed				
Word reading	−.21 (.02*)	−.02 (.83)	−.36 (.02*)	.11 (.47)
Color naming	−.19 (.04*)	−.27 (.004*)	−.36 (.02*)	−.12 (.42)
Interference				
Difference score	−.20 (.03*)	−.09 (.32)	−.31 (.04*)	−.01 (.94)
Golden's score	.14 (.14)	−.16 (.08)	.18 (.26)	−.15 (.34)
Relative score	−.14 (.14)	.08 (.36)	−.20 (.20)	.09 (.57)
Ratio score	−.11 (.25)	.05 (.57)	−.14 (.36)	.06 (.72)
Residualized score	−.11 (.23)	.11 (.22)	−.12 (.46)	.09 (.57)

[1] Model 1 predictors included disease status group (0 = control and 1 = PD patient), age, education, sex, and depression score, and examined the full data sample (N = 126). Each Stroop score was predicted in a separate linear regression using these five predictors.

[2] Model 2 predictors included UPDRS score, age, education, sex, and depression score (N = 51). Model 2 examined effects within the PD patient sample only. Each Stroop score was predicted in a separate linear regression using these five predictors.

β = standardized beta coefficient

* $p < 0.05$

Table 4

Correlations between Stroop interference scores and other standardized measures of processing speed and executive functions (*N*=136; patient and control groups).

	Jensen	Golden	Relative	Ratio	Residualized
Color Naming *N*=126	-	−.05	.07	−.06	−.01
Digit Symbol *z*-score (*N*=125)	.34**	−.06	.05	.01	.01
Symbol Search *z*-score (*N*=63)	.25*	.002	−.04	−.04	−.06
Trail Making A *z*-score (*N*=125)	.20*	−.02	.01	−.03	.01
Trail Making B ratio score (*N*=55)	.11	−.21	.27*	.39**	.27*
COWA *z*-score (*N*=63)	−.06	.05	−.19	−.13	−.18
WCST Categ. *z*-score (*N*=53)	−.10	.21	−.18	−.27*	−.18

* $p < 0.05$

** $p < 0.01$

WCST = Wisconsin Card Sorting Test