

Published in final edited form as:

Neuropsychology. 2011 September ; 25(5): 655–665. doi:10.1037/a0023863.

Let's Inhibit Our Excitement: The Relationships Between Stroop, Behavioral Disinhibition, and the Frontal Lobes

Lara H. Heflin, Victor Laluz, Jung Jang, Robin Ketelle, Bruce L. Miller, and Joel H. Kramer
Department of Neurology, Memory and Aging Center, University of California, San Francisco Medical Center, San Francisco, CA.

Abstract

Objective—The Stroop is a frequently used neuropsychological test, with poor performance typically interpreted as indicative of disinhibition and frontal lobe damage. This study tested those interpretations by examining relationships between Stroop performance, behavioral disinhibition, and frontal lobe atrophy.

Method—Participants were 112 well-characterized patients with mild cognitive impairment or dementia, recruited through UCSF's Memory and Aging Center. Participants received comprehensive dementia evaluations including structural MRI, neuropsychological testing, and informant interviews. Freesurfer, a semi-automated parcellation program, was used to analyze 1.5T MRI scans. Behavioral disinhibition was measured using the Disinhibition scale of the Neuropsychiatric Inventory. The sample ($n=112$) mean age was 65.40 ($SD=8.60$) years, education was 16.64 ($SD=2.54$) years, and MMSE was 26.63 ($SD=3.32$). Hierarchical linear regressions were used for data analysis.

Results—Controlling for age, MMSE, and Color Naming performance, Stroop performance was not significantly associated with behavioral disinhibition ($\beta=0.01$, $\Delta R^2=0.01$, $p=0.29$). Hierarchical regressions controlling for age, MMSE, Color Naming, intracranial volume, and temporal and parietal lobes, examined whether left hemisphere or right hemisphere regions predict Interference speed. Bilaterally, parietal lobes were the brain region in which atrophy best predicted poorer Stroop (left: $\beta=0.0004$, $\Delta R^2=0.02$, $p=0.002$; right: $\beta=0.0004$, $\Delta R^2=0.02$, $p=0.002$). Of frontal regions, only dorsolateral prefrontal cortex atrophy predicted poorer Stroop ($\beta=0.001$, $\Delta R^2=0.01$, $p=0.03$); left and right anterior cingulate cortex (ACC) atrophy predicted better Stroop (left: $\beta=-0.003$, $\Delta R^2=0.01$, $p=0.02$; right: $\beta=-0.004$, $\Delta R^2=0.01$, $p=0.02$).

Conclusions—These findings suggest Stroop performance is a poor measure of behavioral disinhibition and frontal lobe atrophy even among a relatively high-risk population.

Keywords

Stroop; disinhibition; MRI; frontal; neuropsychological

Among the primary uses of neuropsychological assessment is measuring impairments that either reflect real-world skills and behaviors, or provide information about neurological

Please address correspondence concerning this article to Lara Heflin, Psychology 116B, VA Palo Alto Healthcare System, 3801 Miranda Ave., Palo Alto, CA 94304. Lara.Heflin@va.gov..

Lara H. Heflin's current institutional affiliation is with the Veterans Affairs Palo Alto Healthcare System, Palo Alto, CA.

Publisher's Disclaimer: The following manuscript is the final accepted manuscript. It has not been subjected to the final copyediting, fact-checking, and proofreading required for formal publication. It is not the definitive, publisher-authenticated version. The American Psychological Association and its Council of Editors disclaim any responsibility or liabilities for errors or omissions of this manuscript version, any version derived from this manuscript by NIH, or other third parties. The published version is available at www.apa.org/pubs/journals/neu

structure. Although the latter use has been construed as the more inferential of the two, because it presupposes a relationship between task performance and neurological function, both uses are highly inferential. The supposition that a neuropsychological test measures real-world skills and behaviors is an assumption, and one whose merits are called into question by weak or inconsistent relationships between neuropsychological test scores and other measures of the skill or behavior (Shallice & Burgess, 1991; Ready et al., 2001; Chaytor et al., 2006) and poor construct differentiation (Dodrill, 1997; Duncan et al., 1997), particularly for executive functioning (Burgess, 1998). Put simply, the tests upon which neuropsychologists rely may not, in fact, measure the skills and behaviors they are assumed to measure. This assumption should be tested empirically.

One of the most widely used neuropsychological tests is the Stroop Color-Word Interference Test (Stroop, 1935; Rabin, Barr, & Burton, 2005), which typically is used to measure inhibition of a prepotent response (Miyake et al., 2000; Vendrell et al., 1995). Although variations exist, it typically consists of three conditions: a color naming condition for which individuals are asked to name patches of colors on a page as quickly as possible, a color reading condition for which individuals are asked to read color names on a page as quickly as possible, and a color-word interference condition that contains color names written in incongruent colors (e.g., the word “blue” written in red ink) for which individuals are asked to name ink colors as quickly as possible. The increased time needed to name the colors on the color-word interference condition and the increased number of errors committed, relative to the simple color naming condition, are referred to as “the Stroop effect.” This effect is believed to be due to the difficulty of suppressing the natural tendency to read the words, which is thought to be a more practiced and thus automatic action (Stroop, 1935; Delis et al., 2001).

Stroop (1935) originally conceptualized the task as one of interference or inhibition, noting that the terms had been used indiscriminately. Since his publication, researchers have been active in investigating the Stroop effect (see MacLeod, 1991 for a comprehensive review of the experimental literature), and in trying to understand the cognitive processes, largely from an experimental cognitive framework, responsible for producing the effect. The research from experimental psychology, however, appears almost entirely separate from the research and use in clinical neuropsychology, as Stuss and Levine (2002) have noted.

Neuropsychologists may have adopted the Stroop, as Alvarez & Emory (2006) suggest, after one study found that patients with left frontal brain lesions showed significantly more slowing from word reading and color naming to Stroop inhibition than did patients with brain lesions in temporal, posterior, or right frontal regions (Perret, 1974). However, subsequent lesion studies with the Stroop have been inconsistent (Vendrell et al., 1995; Stuss et al., 2001), leading some to describe the Stroop as a “fundamentally ineffective test” for “evaluating frontal lobe functioning,” (Dodrill, 1999, p.563). Increased time taken to complete the Stroop interference typically is predictive of frontal lesions only when color naming speed is not taken into consideration (Foong et al., 1997; Stuss et al., 2001), suggesting that frontal lobe lesions or atrophy cause slowing in general, rather than a true exaggerated interference effect. When controlling for color naming, Stuss et al. (2001) found exaggeration of the Stroop interference only for patients with superior medial frontal lesions. A meta-analytic review that compared the sensitivity to frontal damage of the Wisconsin Card Sorting Test, verbal fluency, and the Stroop, showed the Stroop to be the least sensitive to frontal damage, with $d=-0.30$ (Alvarez & Emory, 2006).

In adopting the Stroop for neuropsychological assessment, clinicians and researchers use an inverted interpretation of the interference effect: while Stroop's original study and subsequent experimental research found slowing in *normal* people on the interference trial,

neuropsychologists tend to conclude that when individuals are abnormally slowed on the interference trial, it is due to impaired inhibition or concentration (Lezak et al., 2004). That is, when a normal process—interference—occurs at abnormally high levels, it is taken as evidence of a pathological process of disinhibition. Studies addressing the construct validity of the Stroop provide poor support for this interpretation, however, and suggest it may measure basic attention, working memory, and processing speed better than inhibition. Correlations with other neuropsychological tests have found that the Stroop shares a large amount of variance (e.g., correlations of .95) with measures of processing speed, leading the authors of one study to conclude that “interference scores from Stroop tasks are not simply measures of inhibition because they share most of their age-related effects with other measures of processing speed,” (Salthouse & Meinz, 1995). Indeed, processing speed appears to mediate the relationship between age and poorer Stroop performance (Earles et al., 1997). In contrast, substantial variation has been found between the Stroop interference task and other tests believed to measure inhibition (Shilling et al., 2002; Earles et al., 1997), suggesting the Stroop is a poor measure of inhibition. Factor analysis has shown the Stroop interference task to load on a measure of “sustained selective processing” along with Serial 7 and Serial 13 tasks, tasks that have little to do with inhibition (Shum et al., 1990). There thus appears to be little evidence from other neuropsychological tasks that Stroop interference is a measure of inhibition rather than simply attention and processing speed.

More troublesome is that neuropsychologists' interpretation of the Stroop may go beyond cognitive interference to inferences about behavioral disinhibition. Such inferences occur despite the dearth of evidence that difficulty with Stroop inhibition has any relationship with behavioral disinhibition. Two studies using the Neuropsychiatric Inventory (NPI) have failed to find relationships between the Stroop and behavioral disinhibition in samples of people with probable Alzheimer's disease (AD) and Parkinson's disease (PD) (Aarsland et al., 1999; Back-Madruga et al., 2002), although it was significantly correlated with NPI total score (Back-Madruga et al., 2002), aggression (Back-Madruga et al., 2002), and apathy (Aarsland et al., 1999). In contrast, Chan (2001) found a borderline significant association ($r=.18$, $p=.05$) between an inhibition factor of the Disexecutive Questionnaire (DEX) and time taken to complete the Stroop among a sample of cognitively normal Hong Kong Chinese people aged 18 – 50.

Despite the research that has cast doubt on the validity of Stroop as a measure of disinhibition, and the inconsistencies in its relationship with the frontal lobes, neuropsychologists continue to use it in these ways. The first aim of the present study is to directly examine how well the Stroop Interference test predicts behavioral disinhibition in daily life. The second aim is to examine the extent to which poor performance on the Stroop Interference test is associated with atrophy in frontal regions believed to be necessary for inhibition.

Method

Participants

Participants were 112 individuals with mild cognitive impairment (MCI) or dementia, who were recruited through the University of California, San Francisco (UCSF) Memory and Aging Center. They were drawn from the Clinical Core of a Program Project Grant investigating frontotemporal dementia and UCSF's Alzheimer's Disease Research Center. Participants received comprehensive dementia evaluations including a neurological exam, neuropsychological assessment, and an informant interview. Within the neuropsychological assessment, participants were administered the Stroop Interference test (Interference). As part of the informant interview, an individual trained on administering the Neuropsychiatric Interview (NPI) administered the NPI to a family member, spouse, or close friend of the

patients. A team of neurologists, neuropsychologists, nurses, and social workers made consensus diagnoses based upon information from medical, social, psychiatric, and family history; neurological exam; neuropsychological evaluations; and clinical inspection of brain imaging. These diagnoses were made according to published criteria (e.g., McKhann et al., 1984; McKeith et al., 1992; Peterson et al., 1997; Neary, Snowden, & Mann, 2000). The UCSF IRB approved the research in which these individuals were involved, and written informed consent was obtained from participants.

Participants were included in the study if they had: 1) a diagnosis of MCI or dementia, 2) completed the Interference test, and 3) received MRI. Individuals were excluded if they had: 1) a MMSE score less than 15, or 2) abnormally low scores (classified as outliers) of 20 or fewer correct on the Color Naming task, which indicated a speech or language impairment that would impair Interference performance irrespective of inhibition skills. Of 137 individuals who had completed the Interference test, were not cognitively normal, and who had either NPI or structural imaging data, 22 were excluded for having indeterminate diagnoses. One person was excluded for having a MMSE of less than 15, and two were excluded for having abnormally low scores on Color Naming. Eleven people were missing NPI data, but had available data for the other variables. Table 1 summarizes demographic, neuropsychological, and NPI Disinhibition data for participants, while Table 2 shows the diagnostic distribution of the sample.

Stroop Interference Test

The Stroop Interference test contained two conditions: Color Naming and Interference. In the first condition, Color Naming, participants were shown a stimulus with 126 (18 rows of 7) word-length strings of 'X's (e.g., XXX XXXXX XXXX) printed in blue, red, or green ink, and asked to name the ink color of each, row by row. In the second condition, Interference, participants were shown a stimulus with 77 (11 rows of 7) color words (blue, red, and green) written in incongruent colors and asked to name the ink color in which the words are written. For each task, participants were instructed to work as quickly as possible without skipping any or making mistakes. Two scores were recorded for each condition: the number of correct responses completed in one minute, and the number of errors committed. The variable of interest was the Stroop effect, or the number of correct responses on the Interference condition after controlling statistically for the number of correct responses on the Color Naming condition. A higher number of correct responses is indicative of better scores and faster performance.

NPI

The Neuropsychiatric Inventory (NPI) is a structured interview designed to assess psychiatric symptoms in individuals with dementia (Cummings, 1997; Cummings et al., 1994). It was administered to a caregiver, relative, or other person familiar with the individual. Twelve psychiatric symptoms were assessed: delusions, hallucinations, depressed mood, anxiety, agitation, euphoria, apathy, irritability, disinhibition, aberrant motor behavior, night-time behavior disturbances, and eating behavior changes. For each symptom, the interviewer asked a screening question; if the informant endorsed it, seven or eight follow-up questions were asked. Then the informant rated the frequency of the symptom on a 4-point scale (1-4), and the severity of the symptom on a 3 point scale (1-3), with higher numbers indicating greater frequency and severity. A total score for each symptom, including disinhibition, was obtained by multiplying the frequency by the severity. This total score for the disinhibition symptom was the score for the Disinhibition scale.

Neuroimaging

MRI scans were acquired on a 1.5T Magnetom VISION system (Siemens Inc., Iselin, NJ, USA) using a standard quadrature head coil. Volumetric magnetization-prepared rapid gradient echo (MP-RAGE) MRI (TR/TE/inversion time [TI] = 10/4/300 msec) obtained T1-weighted structural images of the entire brain. The T1 images were in a coronal orientation, with a 15° flip angle, with 1.0 × 1.0 mm² in-plane resolution and 1.5 mm slab thickness.

The T1 MPRAGE structural MR images were analyzed using Freesurfer, which is documented and freely available for download online at: <http://surfer.nmr.mgh.harvard.edu/>. Previous publications have detailed and validated the software (Segonne et al., 2004; Dale et al., 1999; Fischl et al., 1999a; Fischl et al., 2001). Freesurfer is a surface-based structural MRI analysis tool that segments white matter and tessellates both grey and white matter surfaces. The procedure, in brief, involves the removal of non-brain tissue using a hybrid watershed/surface deformation procedure (Segonne et al., 2004) and intensity normalization (Sled et al., 1998), followed by automated Talairach transformation and volumetric segmentation of cortical and subcortical gray and white matter, subcortical limbic structures, basal ganglia and ventricles, used to calculate total intracranial volume (ICV) (Fischl et al., 2002; Fischl et al., 2004a). The surfacing algorithm uses intensity and continuity data, and corrects topological defects to generate a continuous cortical ribbon used to calculate gray matter volume and thickness (Segonne et al., 2004; Fischl et al., 2001; Segonne et al., 2007; Fischl & Dale, 2000), a procedure validated against histological analysis (Rosas et al., 2002) and manual measurements (Kuperberg et al., 2003; Salat et al., 2004). This cortical surface is then inflated and registered to a spherical atlas and parcellated into regions of interest (ROI) based on gyral and sulcal structure (Fischl et al., 1999a; Fischl et al., 2004b; Fischl et al., 1999b; Desikan et al., 2006).

Cortical regions of interest were defined as described in Desikan et al. (2006). For data reduction, some frontal sub-regions were combined to create broader regions of interest for each hemisphere, as were temporal and parietal regions. The middle frontal cortex (MFC) was the sum of the caudal middle frontal and rostral middle frontal cortices for each hemisphere; the inferior frontal cortex (IFC) was the sum of the pars opercularis, pars orbitalis, and pars triangularis; the orbitofrontal cortex (OFC) was the sum of the lateral and medial orbitofrontal cortex; and the anterior cingulate cortex (ACC) was the sum of the caudal and rostral anterior cingulate volumes. The temporal lobe cortices for each hemisphere were created by summing the entorhinal, parahippocampal, temporal pole, fusiform, superior temporal sulcus, inferior temporal, middle temporal, superior temporal, and transverse temporal volumes. The parietal lobe cortices for each hemisphere were created by summing the post-central, supramarginal, superior parietal, inferior parietal, and precuneus volumes.

Data Analysis

First, we examined data for univariate outliers and normality. As described in the 'Participants' section, outliers on the Color Naming condition were excluded from the sample due to probable language problems and were not included in the sample. To deal with non-normality and outliers on other variables, two strategies were used: first, we conducted logarithmic data transformations and evaluated their improvements on the variable distributions and outliers; second, if data for variables remained significantly non-normal, we conducted non-parametric analyses in order to confirm the findings from more commonly used parametric statistics. Because parametric statistics are more powerful than non-parametric statistics, we report the findings of both approaches, when used. Before starting hypothesis-testing analyses, we plotted the relationship between the color-word interference tasks and NPI Disinhibition, and the relationship between color-word

interference tasks and lobar volumes. Correlations were then calculated between the variables. Because the NPI Disinhibition remained non-normally distributed after the log-transformation, we used Spearman's correlations for the NPI; Pearson's correlations were used for all other variables. After each hypothesis-testing hierarchical linear regression, plots of residuals were examined for normality, linearity, and homoscedasticity. Collinearity diagnostics were also examined to assess for collinearity.

To examine whether the Stroop Interference task is a significant predictor of behavioral disinhibition, we first regressed NPI Disinhibition on Interference using linear regression. We then used hierarchical linear regression to determine whether Interference was a significant predictor of NPI Disinhibition after controlling for the contributions of Color Naming, age, and general cognitive impairment (MMSE). Because the results of the regressions were counterintuitive, we then included Interference errors in the last step of the hierarchical linear regression to ascertain whether inaccuracy on the task was a significant predictor of NPI Disinhibition. These analyses were then replicated using two strategies: 1) logistic regressions, since NPI Disinhibition is more accurately an ordinal than a continuous variable, and 2) excluding individuals with no evidence of behavioral disinhibition.

To examine whether specific frontal lobe and anterior cingulate region volumes are significant predictors of Interference, we first examined partial correlations between Interference and each region of interest (OFC, MFG, IFC, superior frontal cortex, or SFC, and ACC volumes), controlling for Color Naming speed, intracranial volume (ICV), age, and MMSE. These regions were chosen based upon findings from previous research of an association between frontal lobe or anterior cingulate functioning and performance on Stroop Interference (Perret, 1974; Pardo et al., 1990; Stuss et al., 2001; Brass et al., 2005; Alvarez & Emory, 2006). Each region of interest that was correlated with Interference time at a significance level of $p < 0.20$ was included in the hypothesis-testing analyses. To test our hypotheses, we used hierarchical linear regression to determine whether these brain regions were significant predictors of Interference time after controlling for Color Naming speed, MMSE, age, ICV, and non-frontal brain regions (temporal and parietal lobe volumes). Because left and right hemisphere volumes for each region of interest were very highly correlated—most above $r = 0.80$ —separate regressions were conducted for left and right hemispheres to avoid multicollinearity. Color Naming was entered into the first step of the hierarchical regression; MMSE, age, and ICV were entered into the second step; temporal and parietal lobes were entered into the third step; and the brain regions of interest that met the criteria described above were entered into the last step. The significance level for each variable was set at $p < 0.05$.

Results

Examination of the variables distributions showed Color Naming and Interference were mildly non-normally distributed, and NPI Disinhibition was substantially non-normally distributed. Applying log transformations to the data improved normality for NPI Disinhibition (original skew=1.65, kurtosis=1.54; log-transformed skew=1.06, kurtosis=-0.55), but the variable remained non-normal. Log transformations of Color Naming and Interference exacerbated the non-normality of the distributions, so the original variables were retained. All the regional brain volumes were normally distributed and without outliers. Plots of bivariate correlations between Color Naming, Interference, log-transformed NPI Disinhibition, and regional brain volumes suggested linear relationships between the variables of interest.

Interference and NPI Disinhibition

Bivariate correlations between Interference and Disinhibition are summarized in Table 3. Both Interference correct and Color Naming correct were significantly negatively correlated with Disinhibition, indicating that better performance on Color Naming and Interference was associated with less disinhibition. Partial correlations between Interference and Disinhibition are also summarized in Table 3, and show a non-significant negative correlation between Interference and Disinhibition, after controlling for Color Naming, MMSE, and age.

To test whether Interference performance predicts behavioral disinhibition, we ran a series of regressions, summarized in Table 4. First, we regressed NPI Disinhibition on Interference, to see whether the Stroop interference task alone predicted behavioral disinhibition. Interference performance alone was a significant predictor of NPI Disinhibition. When controlling for Color Naming performance, Interference performance was no longer a significant predictor of NPI Disinhibition. After controlling for age, Color Naming, and MMSE, the association between Interference performance and NPI Disinhibition remained non-significant. Collinearity diagnostics showed no problematic collinearity between Color Naming and Interference despite the high correlation between these two variables (Tolerance=0.24, VIF=4.13, Condition Index=3.80).

It was possible that the poor association between Interference performance and Disinhibition was due to a speed-accuracy tradeoff, wherein individuals with greater levels of disinhibition worked quickly and achieved an equivalent number of items correct in the time limit as individuals with less or no disinhibition, but at the expense of accuracy, committing more errors. If this were true, the number of Stroop Interference items correct within a time limit would be a poor predictor of disinhibition, but errors would be a better predictor. To explore the possibility of a speed-accuracy tradeoff accounting for the poor association between Interference performance and Disinhibition, we regressed Disinhibition on Interference correct and Interference errors, controlling for age, Color Naming, and MMSE (Model 4, Table 4). Errors on the Interference task were not significant predictors of Disinhibition ($\beta=0.10$, $p>0.05$), and the association between Interference correct and Disinhibition remained non-significant ($\beta=0.02$, $p>0.05$).

Because the findings ran counter to the conventional wisdom in neuropsychology, additional exploratory analyses were conducted to ensure these findings were not a statistical artifact. In the first of these analyses, we treated Disinhibition as an ordinal, rather than a continuous variable, and used proportional odds logistic regressions to explore the relationship between Interference and un-transformed Disinhibition. Because the original 13-point (0-12) distribution of the NPI Disinhibition scores led to violation of the proportional odds assumption, the distribution was re-examined and categorized into a 4-point ordinal variable (0-3) based upon natural cut-points in the distribution. Using this variable, the proportionality and linearity assumptions of the statistic were met. We ran logistic regressions using Interference to predict this 4-point ordinal NPI Disinhibition variable, both alone and in conjunction with the covariates described above. Interference was not a statistically significant predictor of NPI Disinhibition using logistic regression either alone (OR=0.98, $p>0.05$), after controlling for Color Naming (OR=1.01, $p>0.05$), or after controlling for Color Naming, age, education, and MMSE (OR=1.02, $p>0.05$).

In subsequent exploratory analyses, individuals with NPI Disinhibition scores of '0'—individuals whose caregivers reported no symptoms of disinhibition—were excluded from the sample. This was done to explore whether the relatively high number of people without reported behavioral disinhibition were skewing the results despite the data transformation and use of logistic regression, and to ascertain whether the relationship between behavioral

disinhibition and Stroop interference may be seen for the most relevant population: those who are showing behavioral disinhibition. Spearman's correlations were calculated between the Stroop variables and NPI Disinhibition for the 35 individuals with NPI Disinhibition scores greater than 0. The associations between NPI Disinhibition and Color Naming ($\rho = -0.13$), and NPI Disinhibition and Interference ($\rho = -0.05$) were each non-significant. Interference was not a statistically significant predictor of NPI Disinhibition using linear regression either alone ($\beta = -0.00$, $p > 0.05$), after controlling for Color Naming in a hierarchical regression ($\beta = 0.01$, $p > 0.05$), or after controlling for Color Naming, age, education, and MMSE ($\beta = 0.01$, $p > 0.05$).

Interference and frontal lobe volumes

Bivariate correlations between Interference and frontal lobe volumes are summarized in Table 3. Interference performance was significantly positively correlated with left and right orbitofrontal cortex (OFC), left and right middle frontal gyrus (MFG), left and right inferior frontal cortex (IFC), left and right superior frontal cortex (SFC), left and right temporal cortex, and left and right parietal lobe volumes. Partial correlations between Interference and frontal lobe volumes, which controlled for Color Naming, MMSE, age, and ICV are also summarized in Table 3. After controlling for these covariates, Interference was significantly positively correlated with left and right MFG, and left and right parietal volumes.

To test whether the volume of specific frontal lobe regions predicts Interference performance, we used the results of the partial correlations between Interference and frontal regions to guide model-building, including in the model the frontal regions that showed associations with Interference at the level of $p < 0.20$. Within the left hemisphere, the MFG, SFC, and ACC were the only frontal regions associated with Interference ($p < 0.20$), and so warranted inclusion into the model. Within the right hemisphere, the MFG and ACC were the only frontal regions associated with Interference ($p < 0.20$), and so warranted inclusion into the model. To examine whether any frontal regions of interest within the left hemisphere predicted Interference performance, we used a hierarchical linear regression, with Color Naming time entered into the first step, MMSE, age, and ICV into the second step, left temporal and left parietal lobes into the third step, and left MFG, left SFC, and left ACC into the final step. The same approach was used to examine whether any frontal regions of interest within the right hemisphere predicted Interference: using hierarchical linear regression, Color Naming was entered into the first step, MMSE, age, and ICV into the second step, right temporal and right parietal lobes into the third step, and right MFG and right ACC into the final step.

As the change in R squared shows in Table 5 (Models 1 & 2), Color Naming performance accounted for nearly 73% of the variance in Interference performance. The remaining variables accounted for relatively little—less than 3% each—of the variance in Interference performance for each model. When examining the extent to which left hemisphere brain regions predicted Interference performance (Table 5, Model 1), the left parietal lobe was the best brain region predictor ($\Delta R^2 = 0.02$, $p < 0.01$). After controlling for Color Naming correct, MMSE, age, ICV, left parietal lobe, and left temporal lobe, left MFG was a significant positive predictor of Interference ($\Delta R^2 = 0.01$, $p < 0.05$) and left ACC was a significant negative predictor ($\Delta R^2 = 0.01$, $p < 0.05$). These findings show smaller left parietal lobe, smaller left MFG and larger left ACC volumes predicted poorer Interference performance. When examining the extent to which right hemisphere brain regions predicted Interference performance (Table 5, Model 2), the right parietal lobe was the best predictor ($\Delta R^2 = 0.02$, $p < 0.01$). After controlling for Color Naming time, MMSE, age, ICV, right parietal lobe, and right temporal lobe, right ACC was a significant negative predictor of Interference ($\Delta R^2 = 0.01$, $p < 0.05$). These findings show smaller right parietal lobe and larger right ACC predicted poorer Interference performance. 1

It was possible that the atrophy pattern of participants with AD was driving the association between poor Stroop performance and atrophy in the parietal cortex and middle frontal gyri. To explore this possibility, we re-ran the analyses excluding individuals diagnosed with AD (N=87 without AD), using the same data analytic strategy described for the original analyses. Within the left hemisphere, the MFG, IFC, and SFC were the only frontal regions associated with Interference ($p < 0.20$), and so warranted inclusion into the model. Within the right hemisphere, the MFG and SFC were the only frontal regions associated with Interference ($p < 0.20$), and so warranted inclusion into the model. We then used hierarchical linear regressions identical to those described in the original analyses to examine whether any frontal regions predicted Interference scores. For the left hemisphere, the only brain region that was significantly associated with Interference was the parietal lobe ($\Delta R^2 = 0.02$, $p < 0.05$). None of the frontal regions included in the left hemisphere regression were significant predictors of Interference ($p > 0.10$), after excluding individuals with AD from the sample. For the right hemisphere, the parietal lobe remained the best predictor of Interference ($\Delta R^2 = 0.02$, $p < 0.01$), but the MFG was also a significant predictor of Interference ($\Delta R^2 = 0.01$, $p < 0.05$).

Discussion

Contrary to clinical lore, this study did not find the expected relationships between the Stroop Interference test and behavioral disinhibition or frontal lobe atrophy. In the first of the study's two main findings, performance on the Stroop Interference test showed no association with behavioral disinhibition after controlling for color naming speed, age, and global cognitive functioning. Indeed, Color Naming performance showed a stronger relationship with behavioral disinhibition than did Stroop Interference performance, and this association remained significant after controlling for age and MMSE. In the second main finding, poorer Stroop Interference performance was associated most significantly with greater parietal lobe atrophy bilaterally, but also with greater left MFG atrophy, and less ACC atrophy bilaterally, after controlling for color naming speed, age, global cognitive functioning, and intracranial volume. The relationship between Interference and parietal lobe atrophy remained even after individuals with Alzheimer's disease, who are liable to show the highest rates of parietal lobe atrophy, were removed from the sample.

The relationship between the Stroop Interference test and behavioral disinhibition highlighted the importance of accounting for Color Naming speed when using the Stroop for clinical or research purposes. Without controlling color naming speed when investigating the Interference effect, any slowing on the Interference task could be due simple to slowing in processing speed. That is, if color naming speed is not controlled for when investigating the relationship between Stroop Interference and inhibition difficulty, it would be impossible distinguish between general slowing and true inhibition difficulty. Indeed, our results suggest that a very large proportion of variance in Stroop Interference scores can be attributed to simple color naming speed. Subsequent analyses showed that before accounting for color naming speed, Stroop Interference performance showed a weak, but statistically significant, negative correlation with behavioral disinhibition. After controlling for Color Naming speed using hierarchical regression, Interference speed was no longer a significant predictor of behavioral disinhibition.

¹Repeating the analyses described in this paper using a version of the Stroop task that included a word reading component, which may elicit a stronger interference effect by priming individuals for reading words, resulted in a very similar set of findings and did not substantially change the conclusions of the present study. In brief, using this alternate Stroop task, poor performance on Stroop Interference failed to significantly predict behavioral Disinhibition, and failed to predict atrophy in frontal brain regions. Neither approach found support for the idea that poor performance on Stroop Interference predicted greater behavioral Disinhibition, and both approaches found that atrophy in the parietal cortex was most significant predictor of poor Stroop Interference performance.

This pattern of results was consistent across several statistical approaches. In each set of analyses, Interference performance failed to show the expected significant negative association with behavioral disinhibition, after simple color naming speed had been accounted for. When using logistic regression or the sample of people who had some evidence of behavioral disinhibition (i.e., ratings of '1' or higher on NPI Disinhibition), the results were identical to those of the hierarchical linear regression. The lack of a statistically or clinically-significant relationship between slowed Interference speed and behavioral disinhibition in the analyses focusing on only individuals who were showing some level of behavioral disinhibition suggests that, even before controlling for Color Naming speed, Stroop Interference tasks are insensitive to behavioral disinhibition. Moreover, it does not appear that an association between Interference speed and behavioral disinhibition in the present study was attenuated to non-significance due to a speed-accuracy tradeoff; Interference errors were not significant predictors of behavioral disinhibition. Put simply, it does not appear that people with good inhibition slowed down in order to minimize errors, and that individuals with high levels of disinhibition failed to slow down and consequently committed more errors.

Previous studies have similarly failed to find an association between Stroop Interference and measures of behavioral disinhibition after accounting for Color Naming speed (Cheung, Mitsis, & Halperin, 2004; Back-Madruga et al., 2002). Indeed, Marra and colleagues (2007) found that patients with frontotemporal dementia (FTD) completed the Stroop test more quickly than patients with Alzheimer's disease (AD) or Progressive Nonfluent Aphasia (PNFA), and these same FTD patients were more likely to be rated as disinhibited on the NPI than the AD or PNFA patients, although the latter finding did not meet the Bonferroni-adjusted significance level. Taken together, FTD patients were both faster on the Stroop, and also more disinhibited, than AD or PNFA patients (Marra et al., 2007), suggesting that individuals with greater levels of disinhibition might complete the Stroop more quickly than individuals with lower levels of disinhibition.

A similar pattern of counterintuitive results emerged when examining the relationship between Stroop Interference and frontal lobe atrophy, with Stroop Interference showing weak relationships with frontal lobe regions after accounting for Color Naming. In simple bivariate correlations, all the frontal regions, but not left or right ACC, showed significant positive associations with Stroop Interference performance, suggesting initially that individuals with frontal lobe atrophy tended to complete fewer items on Stroop Interference than individuals with less frontal lobe atrophy. These correlations are consistent with previous research that has found an association between slowed Interference speed and frontal lobe atrophy or frontal lobe lesions when Color Naming speed was not accounted for (Foong et al., 1997; Stuss et al., 2001; Soderlund et al., 2004). Once Color Naming performance was accounted for using partial correlations, right and left MFG were the only frontal regions that were significantly positively associated with Interference speed. After controlling for Color Naming speed, age, intracranial volume, MMSE, and non-frontal brain regions using hierarchical regressions, poorer Stroop Interference performance was significantly associated with bilateral parietal atrophy and left MFG atrophy, but larger left and right ACC volumes. This finding was also generally consistent with previous research: when Color Naming speed is accounted for, Stroop Interference speed tends to show poor associations with frontal lobe damage (Hanninen et al., 1997; Stuss et al., 2001).

The strength of the association between poor Stroop Interference performance and parietal lobe atrophy, rather than frontal lobe atrophy, was among the most unexpected findings in the study. Parietal atrophy accounted for a larger amount of variance in Stroop Interference performance than did left MFG atrophy, the only frontal lobe region in which atrophy was significantly associated with poorer performance. Due to the association between bilateral

parietal atrophy and poorer Stroop performance, the present study suggests that the Stroop is not a specific indicator of atrophy in frontal lobe regions among individuals with mild cognitive impairment or dementia. Replication of this finding would substantially alter the interpretation of Stroop performance when it is used in neuropsychological assessment for localization, since poor performance may be more indicative of parietal atrophy than left dorsolateral prefrontal cortex atrophy.

One potential explanation for the poor relationship between the Stroop, behavioral disinhibition, and most of the regions of frontal lobe atrophy measured is suggested by a distinction between different types of inhibition and the neurological circuitry that may underlie them. Inhibition may not be a unitary construct (Friedman & Miyake, 2004; Nigg, 2000; Harnishfeger, 1995), and while there is little evidence from this or other studies that the Stroop measures behavioral disinhibition (Aarsland et al., 1999; Back-Madruga et al., 2002; Cheung, Mitsis, & Helperin, 2004), it may better measure a cognitive type of inhibition. Indeed, recent research examining frontal brain regions associated with behavioral disinhibition and executive dysfunction found a double dissociation, with atrophy in the OFC predicting NPI behavioral disinhibition, and atrophy in the MFG predicting executive functioning, including Stroop performance (Krueger et al., under review). The possibility that neuroanatomically distinct pathways are necessary for cognitive and behavioral inhibition could help explain the poor association between Interference and behavioral disinhibition, the counterintuitive finding of an association between ACC atrophy and superior Interference performance, and why Interference performance was positively associated only with left MFG (which includes the left dorsolateral prefrontal cortex) rather than diffuse frontal regions.

In his theoretically-guided taxonomy of inhibitory functions, Nigg (2000) classified the Stroop as measuring 'interference control,' or the ability to suppress competing internal or external stimuli when making a response. He distinguished interference control from intentional motor inhibition, which he described as "inhibiting a dominant or prepotent response" (p.223), and from inhibition in the context of attentional orienting, oculomotor inhibition, motivational inhibition, and automatic inhibition of attention (Nigg, 2000). Using an empirical approach, Friedman and Miyake (2004) used confirmatory factor analysis and structural equation modeling to investigate the relationship between different types of inhibition and their measurement. In their most parsimonious model, they found the Stroop task to load on a factor of Response-Distractor Inhibition along with antisaccade, stop-signal, Eriksen flanker, word naming, and shape matching tasks. Demonstrating an ability to discriminate between different types of inhibition, the Response-Distractor Inhibition factor correlated very poorly (0.01) with a Resistance to Proactive Interference factor of inhibition, which included three tasks that required individuals to inhibit irrelevant intrusions from working memory. They did not, however, include measures of behavioral or motivational disinhibition, so it is impossible to know from this research the extent to which behavioral/motivational and cognitive disinhibition can be empirically dissociated. Nevertheless, both of these taxonomies of inhibitory functions would likely predict the Stroop to have a poor relationship with behavioral disinhibition such as that measured by the NPI.

A differentiation between inhibition subtypes is supported by the distinct roles that are theorized for different frontal-subcortical circuits, with dorsolateral frontal-subcortical circuitry believed to be involved in working memory and attentional or cognitive control (Petrides, 2000; MacDonald et al., 2000), orbitofrontal-subcortical circuitry involved in social and behavioral inhibition (Fuster, 1997), and ACC circuitry involved during tasks with conflicting information (Carter, 2000; MacDonald et al., 2000). Within this context, real-world behavioral inhibition might be expected to be most reliant upon healthy orbitofrontal cortex and circuitry, whereas the more cognitive Stroop task would be expected

to be most reliant upon healthy dorsolateral and ACC cortex and circuitry, if fast Stroop performance requires healthy frontal cortex. Indeed, out of all the frontal regions, behavioral disinhibition on the NPI showed the strongest negative correlation with right and left OFC volumes ($r=-0.43$ for left OFC and $r=-0.54$ for right OFC), suggesting that greater OFC atrophy is associated with greater behavioral disinhibition. This correlation is also consistent with the distinction that Dillon and Pizzagalli (2007) made between response inhibition and cognitive set inhibition in their review of the neuroanatomical substrates of three types of inhibition, insofar as it suggests that OFC is involved in inhibition of affectively meaningful, or reinforced, responses such as those that likely guide much of our real-world behavior. Of the frontal lobe regions, the Stroop Interference task, in contrast, showed (in the hierarchical regressions) a significant negative association with only left middle frontal gyrus volume, the area that corresponds to left dorsolateral prefrontal cortex. This relationship between left dorsolateral prefrontal cortex atrophy and poorer Stroop Interference performance is consistent with MacDonald et al.'s (2000) finding that dorsolateral prefrontal cortex is involved in attentional control necessary for performing the Stroop Interference task.

Although its contribution is small, explaining around 1% of the variance, we did find that ACC volumes bilaterally were negatively associated with Interference performance. Conceptualizing the ACC as a structure that is involved in conflict monitoring (e.g., Carter & van Veen, 2007), rather than one involved in inhibition, renders less surprising the present finding. Within this model, when an individual performs the Stroop task, the normally-functioning ACC should become activated upon recognizing the conflict in the task (i.e., the conflict between the overlearned word reading response and the incongruent ink color identification response), and the individual's performance should slow modestly. The ACC should then recruit frontal regions including the dorsolateral prefrontal cortex (DLPFC) and possibly the medial frontal cortex to improve cognitive control (e.g., improve focus on the response-appropriate stimuli) and perform more the task more efficiently (Carter & van Veen, 2007; Nessler et al., 2007). However, if the dorsolateral DLPFC is working sub-optimally, the individual may have difficulty improving cognitive control, thus leading to poorer performance on the Stroop task despite good ACC functioning. Increased conflict monitoring in the ACC within the context of sub-optimal DLPFC functioning thus may impede Stroop performance by causing individuals to slow down despite being unable to benefit from improved cognitive control. Indeed, even cognitively normal older adults may show intact conflict monitoring but reduced engagement of cognitive control processes when faced with high-conflict tasks such as the Stroop (Nessler et al., 2007).

If we consider regional volume to be an indicator of possible neurological activity such that smaller ACC volumes are likely to be less active than larger ACC volumes, our present finding is also consistent with previous findings of an association between greater activation and slower Stroop performance (Raz et al., 2005; MacDonald et al., 2000). Similarly, Milham et al. (2002) found greater ACC activation using functional MRI imaging during the Stroop amongst older adults compared to younger adults, and interpreted this ACC activation as a response to poorer attentional control and greater propensity for errors amongst older adults, rather than a necessary part of inhibition. Within the context of the ACC's apparent role in conflict-monitoring (Carter & van Veen, 2007; Carter et al., 2000; Gehring & Knight, 2000; Carter et al., 1998), the relationship between larger ACC volume and slower Stroop Interference speed is consistent with the idea that ACC involvement during tasks of conflict might serve a cautionary function of slowing responses, or that slowing may be a by-product of monitoring during conflict-laden tasks.

Despite the null findings in the first portion of the study, it remains possible that the Stroop could better measure inhibition deficits in this population if other—possibly more cognitive or attentional—types of inhibition were investigated. Because we did not include measures

of different types of inhibition, one limitation of the present study is that it cannot shed any light on whether the Stroop Interference task may be a better measure of another type of disinhibition than behavioral disinhibition. This may be an important avenue of future research within samples of individuals who are at high risk for inhibitory deficits, given the evidence and theoretical basis for distinct inhibitory processes (Friedman & Miyake, 2004; Nigg, 2000; Harnishfeger, 1995). In addition, because our study focused entirely on older adults with neurodegenerative disorders, it is possible that the Stroop may have stronger associations with disinhibition in other populations of individuals with inhibitory deficits.

The sole use of an informant report as a measure of behavioral disinhibition is another limitation of the present study. Although informant ratings of patient behavior have several advantages over clinician's ratings of a patient's behavioral disinhibition—such as the knowledge of what a patient was like prior to the onset of a neurodegenerative disease, and the ability to observe the patient over a longer period of time and a range of situations—rating scales based upon informant reports may have their own problems. Because different caregivers may have different internal criteria for endorsing symptoms of behavioral disinhibition in the patient, inter-informant error is introduced that would not be present if one clinician rated each patient on the level of disinhibition exhibited during an evaluation. Such error variance could come from inter-informant differences in observational skills, willingness to endorse undesirable behavior, and overall level of stress. In the present study, error variance is a particular problem since we are reporting primarily null findings, and increased error variance reduces one's ability to find significant results.

However, if added inter-informant error variance played a large role in preventing us from finding significant results, we would expect that Interference would show poor associations with other indices from the NPI, since all indices from the NPI would logically be expected to have similar inter-informant error variance. When we examined bivariate associations between Interference and other indices from the NPI, Interference was significantly negatively correlated with Agitation ($p < 0.01$), Depression ($p < 0.01$), Anxiety ($p < 0.01$), Euphoria ($p < 0.05$), Apathy ($p < 0.0001$), Irritability ($p < 0.01$), Aberrant Motor Behavior ($p < 0.05$), Night-time Behavior Disturbances ($p < 0.01$), and Eating Behavior Changes ($p < 0.001$). After controlling for Color Naming, only NPI Anxiety remained significantly associated with Stroop Interference ($\rho = -0.21$, $p < 0.05$). As mentioned earlier, previous research has also found an association between Stroop Interference and other indices of the NPI, but not Disinhibition (Aarsland et al., 1999; Back-Madruga et al., 2002), suggesting that informant-based measures of behavioral disinhibition can show significant correlations with neuropsychological test results.

Because statistical factors can also influence one's ability to find significant results, multiple statistical approaches were used within the first part of the study, with each approach compensating for weaknesses in the other approaches. Hierarchical linear regression allowed us to model the data in a manner that is among the most commonly used in psychology research, and which preserves good statistical power. However, because the NPI Disinhibition variable was ordinal rather than continuous, and could not be made to approximate a good normal distribution even with data transformations, we also used the more statistically appropriate—albeit less powerful and less commonly used—approach of proportional odds ordinal logistic regression. Finally, we examined the association between Color-Word Interference and behavioral disinhibition using only individuals who had received ratings of '1' or higher on the Disinhibition rating scale, replicating the poor relationship between Color-Word Interference and behavioral disinhibition even among a sample of individuals who showed some evidence of disinhibition. This last approach reduced the sample size, therefore reducing statistical power, but it allowed a focus upon the individuals at highest risk for behavioral disinhibition. As mentioned earlier, each of these

approaches resulted in null findings, showing no association between Stroop Interference and behavioral disinhibition after controlling for color naming.

Although the failure to reject a null hypothesis does not mean that the null hypothesis should be accepted, we thought it noteworthy that the Stroop interference task failed to be a significant predictor of behavioral disinhibition and atrophy in most regions of the frontal lobes. This failure occurred despite a large sample size, multiple methods of examining the data, and focusing upon a population at high risk for behavioral disinhibition, a population for whom the Stroop is thought to be a useful and meaningful measure. The findings suggest not simply that there was insufficient evidence for the null hypothesis to be rejected, but also show a clinically insignificant relationship between the Stroop and behavioral disinhibition, once basic color naming speed is accounted for. While the present study did not examine other types of inhibition, previous studies have not found strong evidence that Stroop Interference is a good measure of inhibition after color naming time is accounted for (Shilling et al., 2002; Earles et al., 1997; Salthouse & Meinz, 1995; Shum et al., 1990), as noted earlier. Moreover, the present findings showed unimpressive or counterintuitive relationships between the Stroop and atrophy in frontal lobe regions, with parietal lobe atrophy predicting poor Stroop performance better than dorsolateral prefrontal cortex atrophy, and with ACC atrophy predicting improved Stroop performance. Hence, the data compelled us to come to the unconventional conclusion of accepting the null hypothesis in this instance: the Stroop does not appear to be a meaningful measure of behavioral disinhibition—although it may be a measure of cognitive inhibition—or a specific indicator of frontal lobe damage. We suggest that neuropsychologists be cautious in interpreting results from the Stroop, and should avoid interpreting it in isolation from color naming speed.

Acknowledgments

This study was supported in part by the National Institute on Aging (NIA) grant 1 R01 AG022983-01A1, NIA grant 5 P01 AG019724, NIA grant 3 P50 AG023501, and Alzheimer's Disease Research Center of California grant 01-154-20. This work was presented at the 37th Annual Meeting of the International Neuropsychological Society, Atlanta, GA.

References

- Aarsland D, Larsen JP, Lim NG, Janvin C, Karlsen K, Tandberg E, et al. Range of neuropsychiatric disturbances in patients with Parkinson's disease. *Journal of Neurology Neurosurgery and Psychiatry*. 1999; 67:492–496.
- Alvarez JA, Emory E. Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review*. 2006; 16:17–42. [PubMed: 16794878]
- Back-Madruga C, Boone KB, Briere J, Cummings J, McPherson S, Fairbanks L, et al. Functional ability in executive variant Alzheimer's disease and typical Alzheimer's disease. *The Clinical Neuropsychologist*. 2002; 16:331–340. [PubMed: 12607146]
- Brass M, Derrfuss J, von Cramon DY. The inhibition of imitative and overlearned responses: a functional double dissociation. *Neuropsychologia*. 2005; 43:89–98. [PubMed: 15488909]
- Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*. 1998; 4:547–558. [PubMed: 10050359]
- Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD. Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*. 1998; 280:747–749. [PubMed: 9563953]
- Carter CS, Macdonald AM, Botvinick M, Ross LL, Stenger VA, Noll D, et al. Parsing executive processes: Strategic vs. evaluative functions of the anterior cingulate cortex. *Proceedings of the National Academy of Sciences*. 2000; 97:1944–1948.

- Carter CS, van Veen V. Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, and Behavioral Neuroscience*. 2007; 7:367–379.
- Chan RCK. Dysexecutive symptoms among a non-clinical sample: A study with the use of the Dysexecutive Questionnaire. *British Journal of Psychology*. 2001; 92:551–565.
- Chaytor N, Schmitter-Edgecombe M, Burr R. Improving the ecological validity of executive functioning assessment. *Archives of Clinical Neuropsychology*. 2006; 21:217–227. [PubMed: 16554143]
- Cheung AM, Mitsis EM, Halperin JM. The relationship of behavioral inhibition to executive functions in young adults. *Journal of Clinical and Experimental Neuropsychology*. 2004; 26:393–404. [PubMed: 15512928]
- Cummings JL. The Neuropsychiatric Inventory: Assessing psychopathology in dementia patients. *Neurology*. 1997; 48:S10–S16. [PubMed: 9153155]
- Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: Comprehensive assessment of psychopathology in dementia. *Neurology*. 1994; 44:2308–2314. [PubMed: 7991117]
- Dale AM, Fischl Bruce, Sereno MI. Cortical Surface-Based Analysis I: Segmentation and Surface Reconstruction. *NeuroImage*. 1999; 9:179–194. [PubMed: 9931268]
- Delis, DC.; Kaplan, E.; Kramer, JH. *Delis-Kaplan Executive Function System Examiner's Manual*. The Psychological Corporation; San Antonio, TX: 2001.
- Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006; 31:968–980. [PubMed: 16530430]
- Dillon DG, Pizzagalli DA. Inhibition of action, thought, and emotion: A selective neurobiological review. *Applied and Preventive Psychology*. 2007; 12:99–114. [PubMed: 19050749]
- Dodrill CB. Myths of neuropsychology. *The Clinical Neuropsychologist*. 1997; 11:1–17.
- Dodrill CB. Myths of neuropsychology: Further considerations. *The Clinical Neuropsychologist*. 1999; 13:562–572. [PubMed: 10806469]
- Duncan J, Johnson R, Swales M, Freer C. Frontal lobe deficits after head injury: unity and diversity of function. *Cognitive Neuropsychology*. 1997; 14:713–741.
- Earles JL, Connor LT, Frieske D, Park DC, Smith AD, Zwahr M. Age differences in inhibition: Possible causes and consequences. *Aging, Neuropsychology, and Cognition*. 1997; 4:45–57.
- Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Science*. 2000; 97:11050–11055.
- Fischl B, Liu A, Dale AM. Automated Manifold Surgery: Constructing Geometrically Accurate and Topologically Correct Models of the Human Cerebral Cortex. *IEEE Transactions on Medical Imaging*. 2001; 20:70–80. [PubMed: 11293693]
- Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002; 33:341–355. [PubMed: 11832223]
- Fischl B, Salat DH, van der Kouwe AJ, Makris N, Segonne F, Quinn BT, et al. Sequence-independent segmentation of magnetic resonance images. *Neuroimage*. 2004a; 23:S69–S84. [PubMed: 15501102]
- Fischl B, Sereno MI, Dale AM. Cortical Surface-Based Analysis II: Inflation, Flattening, and a Surface-Based Coordinate System. *NeuroImage*. 1999a; 9:195–207. [PubMed: 9931269]
- Fischl B, Sereno MI, Tootell RB, Dale AM. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*. 1999b; 8:272–284. [PubMed: 10619420]
- Fischl B, van der Kouwe A, Destrieux C, Halgren E, Ségonne F, Salat DH, et al. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*. 2004b; 14:11–22. [PubMed: 14654453]
- Foong J, Rozewicz L, Quaghebeur G, Davie CA, Kartsounis LD, Thompson AJ, et al. Executive function in multiple sclerosis: The role of frontal lobe pathology. *Brain*. 1997; 120:15–26. [PubMed: 9055794]

- Friedman NP, Miyake A. The relations among inhibition and interference control functions: A latent-variable analysis. *Journal of Experimental Psychology: General*. 2004; 133:101–135. [PubMed: 14979754]
- Fuster, JM. *The prefrontal cortex: Anatomy, physiology, and the neuropsychology of the frontal lobe*. 3rd Ed.. Lippencott-Raven. J; Philadelphia: 1997.
- Hanninen T, Hallikainen M, Koivisto K, Partanen K, Laakso MP, Riekkinen PJ, Soininen H. Decline of frontal lobe functions in subjects with age-associated memory impairment. *Neurology*. 1997; 48:148–153. [PubMed: 9008510]
- Harnishfeger, KK. The development of cognitive inhibition: Theories, definitions, and research evidence. In: Dempster, FN.; Brainerd, CJ., editors. *Interference and inhibition in cognition*. Academic Press; New York: 1995. p. 175-204.
- Krueger CK, Laluz V, Rosen HJ, Neuhaus JM, Miller BL, Kramer JH. Double dissociation in the anatomy of socioemotional disinhibition and executive functioning in dementia. under review.
- Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, et al. Regionally localized thinning of the cerebral cortex in schizophrenia. *Archives of General Psychiatry*. 2003; 60:878–888. [PubMed: 12963669]
- Lezak, MD.; Howieson, DB.; Loring, DW. *Neuropsychological Assessment*. 4th ed.. Oxford University Press; New York, NY: 2004.
- MacLeod CM. Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*. 1991; 109:163–203. [PubMed: 2034749]
- MacDonald AW III, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*. 2000; 288:1835–1838. [PubMed: 10846167]
- Marra C, Quaranta D, Zinno M, Misciagna S, Bizarro A, Masullo C, et al. Clusters of cognitive and behavioral disorders clearly distinguish progressive primary aphasia from frontal lobe dementia and Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*. 2007; 24:317–326. [PubMed: 17851236]
- McKeith IG, Perry RH, Fairbairn AF, Jabeen S, Perry EK. Operational criteria for senile dementia of Lewy body type. *Psychological Medicine*. 1992; 22:911–922. [PubMed: 1362617]
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984; 34:939–944. [PubMed: 6610841]
- Milham MP, Erickson KI, Banich MT, Kramer AF, Webb A, Wszalek T, et al. Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition*. 2002; 49:277–296. [PubMed: 12139955]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wager TD. The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology*. 2000; 41:49–100. [PubMed: 10945922]
- Neary D, Snowden JS, Mann DMA. Classification and description of frontotemporal dementias. *Annals of the New York Academy of Sciences*. 2000; 920:46–51. [PubMed: 11193176]
- Nessler D, Friedman D, Johnson R Jr, Bersick M. ERPs suggest that age affects cognitive control but not response conflict detection. *Neurobiology of Aging*. 2007; 28:1769–1782. [PubMed: 16930775]
- Nigg JT. On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin*. 2000; 126:220–246. [PubMed: 10748641]
- Pardo JV, Pardo PJ, Janer KW, Raichle ME. The anterior cingulate cortex mediates processing selection in the Stroop attentional conflict paradigm. *Proceedings of the National Academy of Science*. 1990; 87:256–259.
- Perret E. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia*. 1974; 12:323–330. [PubMed: 4421777]
- Peterson RC, Smith GE, Waring SC, Ivnik RJ, Kokmen E, Tangelos EG. Aging, memory, and mild cognitive impairment. *International Psychogeriatrics*. 1997; 9:65–69.

- Petrides M. The role of the mid-dorsolateral prefrontal cortex in working memory. *Experimental Brain Research*. 2000; 133:44–54.
- Rabin LA, Barr WB, Burton LA. Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*. 2005; 20:33–65. [PubMed: 15620813]
- Raz A, Kirsch I, Pollard J, Nitkin-Kaner Y. Suggestion reduces the Stroop effect. *Psychological Science*. 2005; 17:91–95. [PubMed: 16466414]
- Ready RE, Stierman L, Paulsen JS. Ecological validity of neuropsychological and personality measures of executive functions. *The Clinical Neuropsychologist*. 2001; 15:314–323. [PubMed: 11778769]
- Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, et al. Regional and progressive thinning of the cortical ribbon in Huntington's disease. *Neurology*. 2002; 58:695–701. [PubMed: 11889230]
- Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, et al. Thinning of the cerebral cortex in aging. *Cerebral Cortex*. 2004; 14:721–730. [PubMed: 15054051]
- Salthouse TA, Meinz EJ. Aging, inhibition, working memory, and speed. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 1995; 50:P297–P306.
- Segonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, et al. A hybrid approach to the skull stripping problem in MRI. *Neuroimage*. 2004; 22:1060–1075. [PubMed: 15219578]
- Segonne F, Pacheco J, Fischl B. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *IEEE Transactions on Medical Imaging*. 2007; 26:518–529. [PubMed: 17427739]
- Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. *Brain*. 1991; 114:727–741. [PubMed: 2043945]
- Shilling VM, Chetwynd A, Rabbitt PMA. Individual inconsistency across measures of inhibition: An investigation of the construct validity of inhibition in older adults. *Neuropsychologia*. 2002; 40:605–619. [PubMed: 11792402]
- Shum DHK, McFarland KA, Bain JD. Construct validity of eight tests of attention: Comparison of normal and closed head injured samples. *The Clinical Neuropsychologist*. 1990; 4:151–162.
- Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging*. 1998; 17:87–97. [PubMed: 9617910]
- Soderlund H, Nyberg L, Nilsson L-G. Cerebral atrophy as a predictor of cognitive function in old, community-dwelling individuals. *Acta Neurologica Scandinavica*. 2004; 109:398–406.
- Stroop JR. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*. 1935; 18:643–662.
- Stuss DT, Floden D, Alexander MP, Levine B, Katz D. Stroop performance in focal lesion patients: Dissociation of processes and frontal lobe lesion location. *Neuropsychologia*. 2001; 39:771–786. [PubMed: 11369401]
- Stuss DT, Levine B. Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology*. 2002; 53:401–433.
- Vendrell P, Junque C, Pujol J, Jurado MA, Molet J, Grafman J. The role of prefrontal regions in the Stroop task. *Neuropsychologia*. 1995; 33:341–352. [PubMed: 7792000]

Table 1

Demographics, Neuropsychological Test Scores, and NPI Disinhibition

	Mean	SD	Max-Min
Age (years)	65.40	8.60	40 - 93
Education (years)	16.64	2.54	9 - 20
MMSE	26.63	3.32	16 - 30
Color Naming correct	65.93	22.65	22 - 110
Color-Word Interference correct	34.41	18.40	0 - 70
Design Fluency correct	8.03	4.02	0 - 18
Digit Span Backward Span	4.51	1.42	2 - 7
D-words (phonemic fluency)	11.96	5.68	2 - 28
Animals (semantic fluency)	14.25	6.68	1 - 36
GDS	7.01	5.60	0 - 25
CDR	0.70	0.52	0 - 2
CDR Sum of Box Scores	3.45	3.30	0 - 14
NPI Disinhibition (total score)	1.97	3.41	0 - 12

Note: NPI = Neuropsychiatric Inventory; MMSE = Mini-Mental State Examination; GDS = Geriatric Depression Scale; CDR = Clinical Dementia Rating. N=112 for age, education, MMSE, Color Naming correct, Interference correct, D-words; N=111 for Animals; N=110 for GDS; N=109 for Design Fluency correct ; N=108 for Digit Span Backward Span; N=106 for CDR N=101 for NPI Disinhibition.

Table 2

MCI and dementia diagnoses

Diagnosis	Frequency	% total sample
MCI	44	39.3%
AD (probable or possible)	25	22.3%
bvFTD (probable or possible)	20	17.9%
SD	9	8.0%
CBD (probable or possible)	6	5.4%
ALS	3	2.7%
PPA	3	2.7%
PSP (probable or possible)	2	1.8%

Note: MCI=mild cognitive impairment; AD=Alzheimer's disease; bvFTD= behavioral variant fronto-temporal degeneration; SD=semantic dementia; CBD=corticobasal degeneration; ALS=amyotrophic lateral sclerosis; PSP=progressive supranuclear palsy; PPA=primary progressive aphasia. N=112.

Table 3

Bivariate and partial correlations: Interference, Disinhibition, and brain volumes.

	Bivariate Correlations		Partial Correlations
	Color Naming	Interference	Interference
Color Naming		0.85 §	
Interference Errors	-0.35 §	-0.55 §	-0.48 §
NPI Disinhibition	-0.26 **	-0.22 *	0.06
Left OFC	0.23 *	0.25 **	0.03
Right OFC	0.23 *	0.22 *	-0.05
Left MFG	0.41 §	0.50 §	0.29 **
Right MFG	0.37 §	0.44 §	0.20 *
Left IFC	0.35 §	0.36 §	0.02
Right IFC	0.28 **	0.32 §	0.09
Left SFC	0.39 §	0.43 §	0.15 †
Right SFC	0.31 **	0.33 §	0.06
Left ACC	0.17 †	0.12 †	-0.14 †
Right ACC	0.12	0.07	-0.17 †
Left temporal	0.32 §	0.35 §	0.07
Right temporal	0.28 **	0.26 **	-0.04
Left parietal	0.43 §	0.53 §	0.30 **
Right parietal	0.39 §	0.47 §	0.24 *

Note. NPI=Neuropsychiatric Inventory; OFC=orbitofrontal cortex; MFG=middle frontal gyrus; IFC=inferior frontal cortex; SFC=superior frontal cortex; ACC=anterior cingulate cortex; ICV = intracranial volume.

† = <0.20

* = p < 0.05

** = p < 0.01

§ = p < 0.001. Partial correlations control for Color Naming, ICV, age, and MMSE.

Table 4

Hierarchical linear regression predicting NPI Disinhibition with Interference performance

Model	Dependent Variable	Step	Predictor	β	ΔR^2	F	df	ΔR^2 p-value
1	NPI Disinhibition	1	Interference correct	-0.01	0.04	4.12	100	0.045
2	NPI Disinhibition	1	Color Naming correct	-0.02	0.07	7.43	1	0.008
		2	Interference correct	0.01	0.00	0.40	1	0.53
3	NPI Disinhibition	1	Age	-0.02	0.01	1.08	1	0.30
		1	MMSE	-0.05	0.01	0.87	1	0.35
		1	Color Naming correct	-0.02	0.07	7.43	1	0.008
		2	Interference correct	0.01	0.01	1.12	1	0.29
4	NPI Disinhibition	1	Age	-0.02	0.01	1.08	1	0.30
		1	MMSE	-0.05	0.01	0.87	1	0.35
		1	Color Naming correct	-0.02	0.07	7.43	1	0.008
		2	Interference correct	0.02	0.01	1.12	1	0.29
		2	Interference Errors	0.10	0.01	0.69	1	0.41

Note. N=101.

Table 5
Hierarchical linear regressions predicting Interference performance with left and right hemisphere brain volumes

Model	Step	Independent Variable	β	ΔR^2	F	df	ΔR^2 p-value
1	1	Color naming correct	0.53	0.7259	291.34	1	<0.0001
	2	MMSE	0.74	0.0294	13.08	1	0.0005
	2	Age	0.12	0.0008	0.35	1	0.56
	2	ICV	-0.00001	0.0000	0.00	1	>0.99
2	3	Left temporal	-0.00004	0.0010	0.45	1	0.50
	3	Left parietal	0.0004	0.0222	10.55	1	0.002
	4	Left MFG	0.001	0.0100	4.94	1	0.03
	4	Left SFC	-0.00003	0.0004	0.18	1	0.68
2	4	Left ACC	-0.003	0.0106	5.62	1	0.02
	1	Color naming correct	0.53	0.7259	291.34	1	<0.0001
	2	MMSE	1.03	0.0294	13.08	1	0.0005
	2	Age	0.1275	0.0008	0.35	1	0.56
3	3	ICV	-0.000006	0.0000	0.00	1	>0.99
	3	Right temporal	-0.0003	0.0003	0.13	1	0.71
	3	Right parietal	0.0004	0.0219	10.35	1	0.002
	4	Right MFG	0.0009	0.0060	2.91	1	0.09
4	4	Right ACC	-0.004	0.0120	6.04	1	0.02

Note. ICV=intracranial volume; MFG= middle frontal gyrus; SFC=superior frontal gyrus; ACC= anterior cingulate cortex. N=112