Inhibition/Switching Is not Necessarily Harder than Inhibition: An Analysis of the D-KEFS Color-Word Interference Test

Sara M. Lippa, Robert N. Davis

1 University of Houston, Houston, TX, USA
2 Private Practice, Houston, TX, USA

*Corresponding author at: Private Practice, 1020 Holcombe Blvd., Suite 1304, Houston, TX 77030, USA. Tel.: +1-713-799-2818; fax: +1-206-309-0674.
E-mail address: dr.rob.davis@gmail.com (R.N. Davis).

Accepted 24 December 2009

Abstract

The Delis–Kaplan Executive Function System (D-KEFS) Color-Word Interference Test (CWIT) was designed to improve upon the Stroop task by including an inhibition/switching trial, which was designed to be more difficult than the inhibition trial in terms of time to completion and number of errors. The D-KEFS standardization data support this view. However, in clinical practice, we have noticed that many people perform better on the inhibition/switching trial than the inhibition trial. We examined the prevalence and correlates of this atypical performance pattern on the CWIT. Patients seeking outpatient neuropsychological evaluation (n = 119) completed the CWIT as part of a larger test battery. About 57.1% of patients demonstrated an atypical pattern of performance for either completion time or errors. Patients with an atypical pattern for completion time were significantly slower at color naming and word reading than patients with a typical pattern. Patients with an atypical pattern for errors performed better on measures of learning and semantic verbal fluency than patients with a typical pattern. A majority of patients in our sample exhibited atypical performance on the CWIT, and some preliminary correlates of this pattern might aid clinical interpretation.

Keywords: Executive functions; Inhibition; Switching; D-KEFS

Introduction

The Delis–Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001) is a battery of neuropsychological tests designed to measure executive functions (EF) in children and adults, ages 8–89. It is based on the Boston process approach (Milberg, Hebben, & Kaplan, 2009), which stipulates that there is a primary function that each test is designed to measure, but also component functions that contribute to performance on a particular test. Rather than a single test of EF, the D-KEFS includes a series of tests that are given to determine whether poor performance is due to specific impairment in EF, or impairment in a component function. The present study examined performance on the D-KEFS Color-Word Interference Test (CWIT), which consists of four parts: Color naming, word reading, inhibition, and inhibition/switching.

The CWIT begins with the color naming trial, in which the patient is presented with a page containing a series of red, green, and blue squares. The patient is asked to say the names of the colors as quickly as he/she can without making mistakes. The word reading trial is second, in which the patient is presented with a page containing the words “red,” “green,” and “blue” printed in black ink. The patient is asked to read the words aloud as quickly as he/she can without making mistakes. The inhibition trial is third, in which the patient is presented with a page containing the words “red,” “green,” and “blue” printed incongruently in red, green, or blue ink. The patient is asked to say the color of the ink in which each word is printed as quickly as he/she can without making mistakes. This trial is based on the Stroop (1935) procedure. Last is the inhibition/switching trial, in which the patient is presented with a page containing the words “red,” “green,” and “blue” written in red, green, or blue ink. Half of these words are enclosed within boxes. The patient is asked to say the color of the ink in which each word is printed.
Performance is measured by completion time on each of the four trials. In addition, color naming and word reading times may be summed for a composite score representing component functions. Three contrast scores (inhibition vs. color naming, inhibition/switching vs. combined color naming and word reading, and inhibition/switching vs. inhibition) may be calculated to determine whether there is a disproportionate impairment in a higher level function than the component function(s). However, these scores have been found to have poor evidence of reliability (Crawford, Sutherland, & Garthwaite, 2008). The reliability coefficients of all the CWIT contrast scores fell below 0.7, although the coefficients were higher in the youngest age group than in the older age groups. Finally, scores may be calculated for the total number of errors on each trial, as well as corrected and uncorrected errors on the inhibition and inhibition/switching trials.

The CWIT has been shown to discriminate between children with fetal alcohol syndrome and healthy children (Mattson, Goodman, Caine, Delis, & Riley, 1999). Children with fetal alcohol syndrome took longer to complete the inhibition and inhibition/switching trials, but not the color naming or word reading trials. In addition, children with fetal alcohol syndrome exhibited more errors on the inhibition and inhibition/switching trials than the healthy children.

Among healthy participants, the CWIT completion time scores had low positive correlations with a measure of verbal memory, the California Verbal Learning Test-II (CVLT-II). The highest correlations (range: $r = .30$ to $r = .35$) were between the CWIT inhibition and inhibition/switching trials and CVLT-II total immediate recall, short delayed free recall, and recognition discriminability (Delis et al., 2001). The CWIT inhibition and inhibition/switching trials also had low positive correlations with perseverative responses ($r = .23$ and $r = .20$, respectively) on the Wisconsin Card Sorting Test (WCST). In addition, the CWIT inhibition trial completion time was strongly negatively correlated with the number of WCST categories completed ($r = -.53$), whereas the inhibition/switching trial had a lower negative correlation ($r = -.31$) with this measure (Delis et al., 2001). According to the D-KEFS authors, the CWIT improves upon the Stroop procedure because it includes normative data for the number of errors, as well as completion time. In addition, the CWIT includes an inhibition/switching trial. This trial is similar to an earlier modification of the Stroop task that was found to discriminate between patients with mild head injury from healthy participants more effectively than the original Stroop procedure (Bohnen, Jolles, & Twijnstra, 1992).

The inhibition/switching trial was designed to be the most difficult in terms of completion time and the number of errors. We attempted to explore this issue with the D-KEFS standardization data, but were denied permission to perform the analyses (W. Schryver, personal communication, January 22, 2009). Nonetheless, inspection of the published normative data appears to support this intention, in that raw completion time on the inhibition/switching trial is slower than on the inhibition trial. Moreover, in most cases (ages 14–69), the mean number of errors on the inhibition/switching trial is greater than on the inhibition trial. Interestingly, however, from ages 8–13 and 70–89, the mean number of errors on the inhibition/switching trial is less than or equal to the inhibition trial.

Pilot data presented in the D-KEFS manual have suggested that the inhibition/switching trial captures aspects of executive functioning that the inhibition trial does not. In a study comparing performance of Alzheimer’s disease (AD) and Huntington’s disease patients on the D-KEFS (Delis et al., 2001), there were significant differences in the pattern of performance between these groups on the CWIT. Alzheimer’s disease patients performed within the normal range on color naming, word reading, and inhibition, but showed a decrease in performance on the inhibition/switching trial for both completion time and errors. Huntington’s disease patients performed slowly on all four components of the CWIT, but had significantly fewer errors than the AD patients on the inhibition/switching trial.

These findings appear to support the idea that the inhibition/switching trial captures a relatively distinct aspect of executive functioning, which may facilitate differential diagnosis of EF deficits among various patient groups. Despite these initially promising findings, however, in clinical practice, we have observed that some patients perform better on the inhibition/switching trial than the inhibition trial. Although uncommon, some patients even perform at a deficient level on the inhibition trial, but at an average level on the inhibition/switching trial. On many occasions, patients have spontaneously commented that the inhibition/switching trial was easier than the inhibition trial.

Practice effects may be one possible reason for this occurrence. The effort it takes to inhibit the word reading response may be reduced after the inhibition trial is performed. Patients may be able to practice the inhibition process on the inhibition trial and put this skill to use on the inhibition/switching trial, which may decrease their time and errors on the inhibition/switching trial. Hypothetically, if the inhibition/switching trial preceded the inhibition trial, then it might have more novelty and difficulty than it has under the standard administration, in which inhibition precedes inhibition/switching. Novelty may be decreased in the standard administration since patients have already completed the inhibition trial (with its respective cognitive demands) before attempting the inhibition/switching trial. Alternatively, it is possible that the inhibition trial takes more time to complete than the inhibition/switching trial for some patients because the inhibition trial requires more color naming. Color naming requires more completion time than word reading; this fact is readily apparent in the normative data on the component
Irrespective of any executive functioning involved, the inhibition trial requires color naming for 100% of the items, whereas the inhibition/switching trial requires color naming for only 50% of the items. The other 50% of the items on the inhibition/switching trial require word reading. The simple fact that the inhibition trial requires more color naming than the inhibition/switching trial might result in lengthier completion time on the inhibition trial for some patients.

Anecdotal observations from clinical practice suggest that a fair number of patients perform better on the inhibition/switching than the inhibition trial, but the prevalence and correlates of such a performance pattern remain unclear. We sought to determine the prevalence of such atypical performance on the CWIT. Atypical performance was defined as: (a) completion time on the inhibition/switching trial \( \leq \) completion time on the inhibition trial, or (b) errors on the inhibition/switching trial \( \leq \) errors on the inhibition trial. We also hoped to reveal correlates of this atypical performance in order to clarify what differentiates these patients from those who demonstrate the typical pattern of performance on the CWIT.

**Materials and Methods**

**Participants**

Participants underwent a comprehensive neuropsychological assessment, including tests of intellectual, language, visuospatial/constructional, attention, memory, executive, and motor functions, as well as mood and/or personality. Assessments were completed in one session lasting roughly 3 hr. Only a representative subset of the tests administered was included in this study. All testing was conducted individually in a private testing room. Before the testing session, participants signed an informed consent form stating that deidentified (anonymous) data may be used for research purposes, and a clinical interview was conducted by a board certified clinical neuropsychologist (RND).

Participants included 119 adult patients (57 men and 62 women) seen consecutively at a private neuropsychology practice in Houston, TX. Of these, 100 (84%) identified as Caucasian, 8 (6.7%) identified as African American, 7 (5.9%) identified as Hispanic, and 4 (3.4%) identified as another ethnicity. Mean age was 63.2 years (range: 22–88 years; \( SD = 15.7 \)), with 84.3% of the sample falling between the ages of 43 and 81. Mean education was 14.8 years (range: 0–20 years; \( SD = 3.0 \)), with 85.7% of the sample having completed between 12 and 18 years of education. The mean estimated Full-Scale IQ of the sample was within the average range, at 104.9 (range: 80–125; \( SD = 10.5 \)). None of the participants were involved in a legal case with which their neuropsychological testing results might be relevant. Neurologic diagnoses included mild cognitive impairment (31.1%), dementia due to AD (4.2%), dementia of vascular etiology (4.2%), dementia due to other causes (5.0%), Parkinson’s disease (2.5%), traumatic brain injury (2.5%), and multiple sclerosis (1.7%). Psychiatric diagnoses included depression (13.5%), anxiety (5.9%), mixed depressed and anxious mood (15.9%), adjustment disorder (3.3%), attention deficit/hyperactivity disorder (2.5%), somatoform disorder (2.5%), and bipolar disorder (1.7%).

**Measures**

Table 1 summarizes the tests and measures used in the study. As we were looking for general correlates of the atypical pattern of performance, we chose primary variables, rather than breaking down large constructs into multiple components (i.e., total digit span rather than digits forward and digits backward; list learning over 3 trials rather than delayed recall and recognition) to reduce the number of analyses conducted. Normative data for the CWIT scores were obtained by using age-scaled scores from the test manual. Normative data for the other neuropsychological test scores were obtained from commonly used sources (Heaton, Taylor, Miller, & Grant, 2004; Ivnik, Malec, Smith, Tangelos, & Petersen, 1996; Lucas et al., 2005). We estimated FSIQ using two WAIS-III subtests, similarities and matrix reasoning, according to Sattler’s (2008) method.

**Statistical Analyses**

Statistical Package for the Social Sciences (SPSS for Windows, version 14.0, Chicago, IL, USA) was used for all statistical analyses. Patients were classified as “atypical” if their raw score performance (as measured by completion time and number of errors) on the CWIT inhibition/switching trial was less than or equal to their performance on the inhibition trial. (Faster completion times and fewer errors indicate better performance.) The sample was first divided into two dichotomous groups: Patients with an atypical pattern for time and patients with a typical pattern for time. Patients were classified as having an atypical pattern for time if their raw completion time on the inhibition/switching trial was equal to or faster than their raw completion time on the inhibition trial. Patients were classified as having a typical pattern for time if their raw completion time on the inhibition/switching trial was slower than their raw completion time on the inhibition trial. Bivariate correlations
Table 1. Tests and measures

<table>
<thead>
<tr>
<th>Name of test</th>
<th>Citation</th>
<th>Function(s) assessed</th>
<th>Score(s) used</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-KEFS CWIT</td>
<td>Delis and colleagues (2001)</td>
<td>Color naming speed, word reading speed, response inhibition, and cognitive flexibility</td>
<td>CWIT inhibition and inhibition/switching trials (completion time and errors) raw and scaled scores</td>
</tr>
<tr>
<td>WAIS-III</td>
<td>Wechsler (1997)</td>
<td>Intelligence, attention, and working memory</td>
<td>Similalirities, matrix reasoning, and digit span scaled scores</td>
</tr>
<tr>
<td>HVLT-R</td>
<td>Brandt and Benedict (2001)</td>
<td>Verbal memory</td>
<td>Total immediate recall (trials 1–3) T score</td>
</tr>
<tr>
<td>BNT</td>
<td>Kaplan, Goodglass, and Weintraub (1983)</td>
<td>Visual confrontation naming</td>
<td>Total correct without cues T score</td>
</tr>
<tr>
<td>FAS Verbal Fluency</td>
<td>Gladso and colleagues (1999)</td>
<td>Phonemic verbal fluency</td>
<td>Total correct T score</td>
</tr>
<tr>
<td>Animal Naming</td>
<td>Gladso and colleagues (1999)</td>
<td>Semantic verbal fluency</td>
<td>Total correct T score</td>
</tr>
<tr>
<td>Orientation</td>
<td>Benton, Sivan, Hamsher, Varney, and Spreen (1994)</td>
<td>Visuospatial judgment</td>
<td>Total correct T score</td>
</tr>
<tr>
<td>Trail Making Test</td>
<td>Reitan (1958)</td>
<td>Simple and complex sequencing</td>
<td>Completion time for parts A and B T scores</td>
</tr>
<tr>
<td>PHQ-9</td>
<td>Spitzer and colleagues (1999)</td>
<td>Depressed mood</td>
<td>Total score</td>
</tr>
<tr>
<td>GAD-7</td>
<td>Spitzer, Korenke, Williams, and Lowe (2006)</td>
<td>Anxious mood</td>
<td>Total score</td>
</tr>
</tbody>
</table>

Notes: D-KEFS CWIT = Delis–Kaplan Executive Function System Color-Word Interference Test; HVLT-R = Hopkins Verbal Learning Test-Revised; BNT = Boston Naming Test; PHQ = Patient History Questionnaire; GAD = Generalized Anxiety Disorder; SS = scaled score; WAIS-III = Wechsler Adult Intelligence Scale - Third Edition.

were calculated between (a) patient group (atypical vs. typical for time) and (b) demographic variables and scores on other neuropsychological tests.

Next, in a separate set of analyses, the sample was divided into two different groups: Patients with an atypical pattern for errors and patients with a typical pattern for errors. Patients were classified as having an atypical pattern for errors if their raw number of errors on the inhibition/switching trial was equal to or less than their raw number of errors on the inhibition trial. Patients were classified as having a typical pattern for errors if their raw number of errors on the inhibition/switching trial was greater than their raw number of errors on the inhibition trial. Bivariate correlations were calculated between (a) patient group (atypical vs. typical for errors) and (b) demographic variables and scores on other neuropsychological tests.

Results

Nearly one third (31.1%) of patients demonstrated an atypical pattern of performance for completion time. Slightly less than one half (42.9%) of patients demonstrated an atypical pattern of performance for errors. Among all patients, 16.8% demonstrated an atypical pattern of performance for both completion time and errors, 40.3% of patients demonstrated an atypical pattern for either completion time or errors (but not both), and 42.9% of patients demonstrated a typical pattern of performance for both completion time and errors. Hence, 57.1% of patients exhibited some form of atypical performance with respect to the inhibition and the inhibition/switching trials.

The atypical pattern of performance for completion time was not significantly correlated with age, education, gender, estimated IQ, other neuropsychological test scores, or depressive/anxiety symptoms. However, patients with an atypical pattern of performance for completion time were significantly slower at color naming \(r = .31, p < .01\), and word reading \(r = .20, p < .05\), when compared with patients with a typical pattern of performance for completion time. See Table 2 for the complete results.

The atypical pattern of performance for errors was not significantly correlated with education, gender, estimated IQ, or mood. It was, however, related to younger age \(r = - .19, p < .05\), decreased completion time on the inhibition/switching trial \(r = - .33, p < .01\), and better demographically adjusted scores on the Boston Naming Test (BNT) \(r = .23, p < .05\), Trails B \(r = .24, p < .05\), Hopkins Verbal Learning Test-Revised (HVLT-R; \(r = .26, p < .01\), and animal naming \(r = .19, p < .05\). See Table 3 for the complete results.

Note in the preceding paragraph that the atypical pattern of performance for errors was related to younger age, among other factors. This finding raises the possibility that age differences might explain the putative correlations between the atypical performance pattern and demographically adjusted scores on the BNT, Trails B, HVLT-R, and animal naming. One would think that these scores, which are demographically adjusted, would in fact be independent of demographic influences, including age. Nevertheless, we wanted to rule out possible age confounds. In other words, the unwanted possibility would be that younger patients are more likely than older patients to have an atypical performance pattern for errors and better scores on certain neuropsychological tests. We calculated a series of residuals from the raw scores of four tests (BNT, Trails B, HVLT-R, and animal naming) that eliminated age influences to ensure that age could not explain the correlations between the atypical pattern of performance for errors and these...
scores. Using this approach, the atypical pattern of performance for errors remained significantly associated with HVLT-R (r = .25, p < .05) and animal naming (r = .19, p < .05) performance. Figure 1 displays these results. In contrast, the atypical pattern of performance for errors was no longer significantly related to BNT and Trails B performance.

Discussion

As expected, a substantial proportion of 119 consecutive patients from our private neuropsychology practice demonstrated an atypical pattern of performance for either completion time or errors on the inhibition and inhibition/switching trials of the
D-KEFS CWIT. In total, 68 (57.1%) patients exhibited an atypical pattern of performance on the inhibition and the inhibition/switching trials of the D-KEFS CWIT, considering either completion time or errors. Thirty seven (31.1%) of patients demonstrated an atypical pattern of performance for completion time. Fifty one (42.9%) patients demonstrated an atypical pattern of performance for errors.

Patients with an atypical pattern of performance for completion time were significantly slower at color naming and word reading compared with patients with a typical pattern of performance for completion time. Though both of these findings were significant, the effect size for color naming speed ($R^2 = .10$) was more than twice as large as the effect size for word reading speed ($R^2 = .04$). It is possible that patients who completed the inhibition/switching trial in less time than the inhibition trial were faster because they had to name fewer colors in the inhibition/switching trial than the inhibition trial. Color naming typically takes longer to execute than word reading, as evidenced by the D-KEFS and Stroop normative data (Delis et al., 2001; Golden & Freshwater, 2002).

Participants with an atypical pattern of performance for errors were significantly younger, took longer to complete the inhibition/switching trial, and had better performance on BNT, Trails B, HVLT-R, and animal naming. After controlling for age, the atypical pattern of performance for errors was still correlated significantly with HVLT-R and animal naming performance. These results suggest that individuals who demonstrate an atypical pattern for errors are better able to learn and spontaneously generate verbal information. Being better able to learn verbal information may be what makes them perform better on the inhibition/switching trial than the inhibition trial. These patients may be better able to apply the practice from the inhibition trial to their inhibition/switching trial performance than patients who demonstrate the typical pattern. These results are consistent with previous research that found positive correlations between CVLT-II total immediate recall and scores on the inhibition and the inhibition/switching trials (Delis et al., 2001).

Although we investigated the differences in demographics and test scores between the typical and atypical groups, we did not have the sample size to investigate how the pattern related to specific diagnostic groups. Moreover, we were primarily interested in the prevalence and correlates of typical and atypical CWIT performance in a heterogeneous clinical sample, which affords greater generalizability to similar outpatient settings. Future studies could investigate whether the atypical pattern of performance is associated with specific diagnostic groups.

The inhibition/switching trial is presented as a more demanding task on which patients are more likely to make errors, and which takes longer to complete than the inhibition trial, but our results suggest that this is commonly not the case. It is possible that reversing the order of the inhibition and the inhibition/switching trials would affect our results; having to complete the inhibition/switching trial before the inhibition trial would not afford a practice trial for the process of inhibiting an automatic response. However, subsequent studies are needed to determine if this hypothesis is accurate. In our sample, some patients performed at a deficient level on the inhibition trial, but at an average level on the inhibition/switching trial. As response inhibition is conceptualized as being a necessary skill for performing the inhibition/switching trial, this patient profile is interesting and raises questions as to how such a profile should be interpreted. It is clear that further studies are necessary to investigate how this particular profile should be interpreted, as well as to further investigate why this atypical pattern of performance is exhibited by some patients, but not others.

In sum, we have demonstrated that a majority (57.1%) of consecutive outpatients in a private practice setting had faster completion times and/or fewer errors on the inhibition/switching trial than the inhibition trial of the D-KEFS CWIT. Such a finding appears to constitute atypical performance inasmuch as the standardization data generally follow the opposite pattern, and the inhibition/switching trial has been found to capture deficits in EF that the inhibition trial has not. The findings
call into question the idea that the inhibition/switching trial is consistently more difficult than the inhibition trial, and subsequently questions the incremental validity of the inhibition/switching trial. Furthermore, our data suggest two possible mechanisms for this atypical performance pattern. First, patients who perform the inhibition/switching trial faster than the inhibition trial may do so because the inhibition/switching trial involves only half as much color naming as the inhibition trial, and color naming is typically a slower process than word reading. Patients who were slower at color naming were more likely to show the atypical pattern of performance than those who were faster at color naming. Second, patients who demonstrate fewer errors on the inhibition/switching trial than the inhibition trial may have learning characteristics that interact with the order of administration of the CWIT (i.e., inhibition always precedes inhibition/switching). These patients tend to show better verbal learning and semantic verbal fluency than patients who show the typical pattern of CWIT performance.

Conflict of Interest
None declared.

References