A Brief Assessment of Executive Control Dysfunction: Discriminant Validity and Homogeneity of Planning, Set Shift, and Fluency Measures

Karl R. Hanes¹,³, David G. Andrewes¹,², Deidre J. Smith³, and Christos Pantelis³

Departments of Psychology¹ and Psychiatry², University of Melbourne; Clinical Neuropsychiatry Research Unit³, Mental Health Research Institute, Victoria, Australia

Disturbances of executive control have attracted considerable interest in recent times (Lezak, 1993; Stuss & Benson, 1986). This interest predominantly reflects the complexity of these functions, their importance in the performance of everyday life activities, and their ubiquity across a wide spectrum of psychiatric and neurological pathology (Cummings, 1990; Russell & Roxanas, 1990). Executive control comprises abilities such as strategy formation and motor programming that are involved in the organization of complex behavior. Such abilities are considered among the more complex of intellectual functions, relying on the integrity of neural structures that emerge late in the course of ontogenetic development and that lie at the apex of the neuraxial hierarchy. Most commentators have emphasized the importance of the prefrontal cortex in the mediation of executive control, with its involvement in the planning and organization of behavior (Lezak, 1993; Stuss & Benson, 1986).

The authors wish to thank Associate Professor Edmond Chiu for permission to assess patients in his care, Dr. Paul Maruff for criticism of an earlier version of this paper, and three anonymous reviewers for their useful comments. Test materials available upon request.

Address correspondence to: Karl R. Hanes, Cognitive Neuropsychiatry Research Unit, Mental Health Research Unit, Locked Bag 11, Parkville, Victoria, Australia, 3052.
Assessment of executive control is one of the most important areas of both clinical and experimental neuropsychology and its evaluation in such wide-ranging disorders as subcortical degenerative disease, schizophrenia, AIDS-dementia, and obsessive–compulsive disorder is now considered important by neuropsychologists (Cummings, 1990; Hollander, Liebowitz, & Rosen, 1991; Keefe, 1995; Maruff, Currie, Malone et al., 1994). Nonetheless, there are relatively few measures that are considered to be strictly sensitive to executive control function (Lezak, 1993), and many such tasks developed for assessing severely impaired patients (e.g., frontal lesion, head injury) may be insensitive to the less severe executive impairment often observed in extrafrontal disorders.

We have assessed the homogeneity and discriminant validity of three recognised measures of executive control function (Morris, Ahmed, Syed, & Toone, 1993; Vendrell et al., 1995), namely a fluency task, the Stroop interference test (Stroop, 1935), and a new version of the Tower of London test of planning (Shallice, 1982). The new planning task represents an extension in complexity of the original measure (Shallice, 1982) from a maximum complexity of five to eight moves and was developed to be a more sensitive index of executive function than earlier versions. We have used the original task with a number of neuropsychiatric groups and found that, although it appeared to be sensitive to severe planning disability, patients with relatively moderate loss of executive control often performed normally on all but the most complex items on this task (i.e., those involving four and five move solutions). This observation was supported by Lange and colleagues (1992) who used the computerized Tower of London to assess cognitive function in a Parkinson’s disease group and reported impaired performance of this group relative to controls, but only for the four and five move problems. These findings and the desire to develop a more sensitive index of executive function led us to consider the possibility that increasing the complexity of this task beyond five moves may afford a more sensitive measure of executive control abilities. We postulated that the more complex items on this task would more successfully discriminate among subjects than less complex items.

In order to evaluate the validity of these three indices in disorders with relatively subtle impairment of executive function, we chose to compare patients with schizophrenia, Huntington’s disease, and Parkinson’s disease, disorders that present with some breakdown of executive function (Lees & Smith, 1983; Mendez, 1994; Morris, Rushe, Woodruff, & Murray, 1995), but in which such a disturbance is not the most compelling feature. We compared performance in these groups to a control group of normal subjects. The homogeneity of these measures was determined by comparing the interrelatedness of these three measures and their association with nonexecutive control tasks. We predicted that all three patient groups would show aberrant performance on the executive control measures and that these measures would be strongly interrelated but not highly associated with test scores on the nonexecutive control tasks.

**METHOD**

**Subjects**

Four groups were tested. A group of 26 normal subjects with no head injury, substance abuse, or other neurological or psychiatric conditions served as controls. These patients were either volunteers recruited from advertisements or volunteers recruited by referrals from other volunteers who had participated in the project or spouses of patients.

Twelve patients with a diagnosis of Huntington’s disease based on choreiform movements and a positive family history of the disease were also assessed. These patients had a
mean functional disability rating of 2.8 (range 1–4) on the Shoulson–Fahn scale (Shoulson & Fahn, 1979). Patients in stage five of this scale were excluded on the grounds of severe motor disability. Mean duration of illness in these patients since formal diagnosis was 43 months, with a range of 4–120 months. Four of the patients were being treated with dopaminergic antagonistic medication, with three receiving haloperidol and one pimozide.

Twenty-five medicated Parkinson’s disease patients were recruited, all of whom exhibited a bilateral akinetic-rigid syndrome with rest tremor. Mean duration of disease since formal diagnosis in this group was 84 months, with a range of 3–180 months and mean ratings on the Hoehn and Yahr (1967) scale of the latter group were 3.0 (range 1–4). Patients in stage five of this scale were excluded due to severe motor impairment. All but five of these patients were being treated with a dopa-decarboxylase inhibitor (sinemet), while the five patients not receiving sinemet treatment were on either benzhexol alone or in combination with bromocriptine.

The schizophrenic group comprised 20 patients with chronic schizophrenia, all of which were receiving neuroleptic medication, diagnosed according to DSM-III-R (American Psychiatric Association, 1987) criteria. Mean duration of illness since initial diagnosis was 88 months, with a range of 16–192 months. Exclusion criteria for the schizophrenic group were age above 60, and duration of illness below 1 year and above 20 years. This group showed an equal proportion of positive [38.1(12.1)] and negative [38.5(16.2)] symptoms, as assessed by the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984a) and Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984b). Table 1 presents sociodemographic information of the groups. A chi-square test indicated no differences across the groups in terms of sex distribution.

**MATERIALS AND PROCEDURE**

The measures employed included the Stroop task (Stroop, 1935), a category fluency task with a 60-second response interval and a new planning measure. These tests are, for the most part, described in Lezak (1995). The categories chosen for the fluency task were: musical instruments, vehicles, fruit, trees, clothing, furniture, metal objects, food, games, and tools.

The New Tower of London is a variant of the original measure (Shallice, 1982) and consists of 10 items, ranging from two to eight moves in level of difficulty. The full version of this task includes two problems per level of difficulty, although it is possible to employ a shortened version consisting of one item per level of difficulty, perhaps in working with groups with motivational difficulties. Because a difficulty of eight moves represented the upper range of available complexity of the Tower of London puzzle, items were included...
that assess planning requiring a sequence of two, four, six, seven, and eight steps. Cards illustrating the starting state, goal position, and the number of permissible moves have been devised based on the original version (Shallice, 1982). Subjects are permitted a maximum of 90 seconds to complete each trial, and initial response latency is recorded as a measure of planning time, defined as the time between card presentation and displacement of the first ball from one of the pegs. The performance measure employed was accuracy of performance, which has been found to be most sensitive to executive control function.

Other measures not related to executive control abilities included the average number of peg placements on the Purdue Grooved Pegboard (Purdue Research Association, 1948) and recall at 30 min delay on the Rey Complex Figure (Rey, 1941). Motor performance was controlled by the Grooved Pegboard, premorbid IQ was estimated with the National Adult Reading Test (NART; Nelson, 1982), and depression was quantified using the Beck Depression Inventory (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961).

**RESULTS**

Performance indices for the executive control measures were the mean number of appropriate words for the fluency task across the 10 categories, a Stroop score, representing the time taken to name colors subtracted from the time taken to read words, and the number of trials to completion on a modified Tower of London task. The number of trials to completion was chosen as the performance measure for the modified planning task, because this was highly interrelated with all other performance measures. Due to the significant demographic variability among the groups, alpha was set at 0.01 in order to reduce the probability of a Type I error.

**Group Differences**

Repeated measures analysis of covariance (ANCOVA), controlling for age, premorbid IQ, motor performance, and depression, revealed a significant statistical group effect for all the measures: planning \( F(3, 77) = 4.6, p < .005 \); Stroop, \( F(3, 77) = 5.4, p < .003 \); fluency, \( F(3, 77) = 7.0, p < .002 \). Multiple post hoc comparisons, using the Scheffe procedure, revealed that all three patient groups performed more poorly than controls on all executive control measures, but there were no significant differences between the performance of individual patient groups.

With regard to the modified planning task, post hoc comparisons for each difficulty level showed that poor performance of the patient groups was specific to the more complex items. There were no differences between patient groups and controls for the two and four move problems, with the exception of the schizophrenic group which differed from controls on the four move problem \( p < .01 \). An effect for item complexity was also observed for this measure \( F(3, 77) = 66.4, p < .0001 \), as well as an interaction for group \( \times \) item complexity \( F(12, 316) = 8.4, p < .001 \). Pair-wise comparisons for the interaction effect revealed significant effect for group \( \times \) item complexity when each of the patient groups were entered separately with the control group \( p < .01 \).

**Discriminant Function Analysis**

In order to evaluate the relative utility of the executive control measures in discriminating patients with schizophrenia, HD, and PD from normal controls, a series of discriminant analyses were conducted, corrected for the effects of age, premorbid IQ, depression, and
motor performance. In these analyses, patient groups were pooled together, because these groups performed very similarly on all these measures (see Table 2). Findings indicated statistically significant separation between the two groups, Wilks $L' = .54$, $p < .0003$.

Overall, the three tasks discriminated between the patient groups and controls with 81% accuracy. The loading matrix of correlations between test scores and discriminant functions suggested that the order of tasks in terms of discriminant power were the planning measure, the fluency task, with the Stroop task being the least discriminating.

Pearson product-moment partial correlations, controlling for age, premorbid IQ, motor performance, and depression between performance of each of the groups on all of the measures were also conducted. Findings indicate consistent strong relationships between the three executive control measures, but only moderate associations between these tasks and measures not considered to be associated with executive control or problem solving abilities (see Table 3).

**DISCUSSION**

We have assessed the discriminant validity and homogeneity of a brief battery of measures of executive control function in a number of neuropsychiatric groups who frequently manifest difficulties on measures of executive control function. Findings indicated aberrant performance of the three patient groups on all the executive control measures used, relative to controls and discriminant validity of these tasks, which were strongly interrelated. Significantly, a strong interaction was observed for group and item complexity on the planning measure, with the more complex items more successfully discriminating between performance of patient groups and controls. Indeed, there were no consistent differences between patient groups and controls on the two and four move problems, suggesting that the current version of the Tower of London task may be more sensitive to subtle problem-solving perturbations than previous versions.

The finding of impaired performance of the three patient groups on the measures of executive function adds to the literature on problem solving and executive function disturbance in HD, PD, and schizophrenia (Lees & Smith, 1983; Mendez, 1994; Morris et al., 1995). The strong relationship between test scores on the three executive control tasks, but only moderate correlations of these tasks with other measures, is consistent with the homogeneity of the executive control tasks in the recruitment of executive control and problem solving abilities (Morris et al., 1993; Vendrell et al., 1995).

**TABLE 2**

<table>
<thead>
<tr>
<th>Measures and Standard Deviations of Huntington's Disease (HD), Parkinson's Disease (PD), Schizophrenic (SC) Patients, and Normal Controls on Tasks of Executive Control Function and Other Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Controls</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Executive control tasks</td>
</tr>
<tr>
<td>NTOL*</td>
</tr>
<tr>
<td>Fluency*</td>
</tr>
<tr>
<td>Stroop*</td>
</tr>
<tr>
<td>Other measures</td>
</tr>
<tr>
<td>Rey Figure*</td>
</tr>
<tr>
<td>Purdue Pegboard</td>
</tr>
</tbody>
</table>

*ANCOVA significance, $p < .01$.

Note. NTOL = a New Tower of London measure.
TABLE 3
Correlations Between Poor Performance on Measures of Executive Control and Tests not Linked to Executive Control Function in PD, HD, and Schizophrenia Patients and Normal Controls

<table>
<thead>
<tr>
<th>Executive Control Measures</th>
<th>Other Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>NTOL</td>
<td>FL</td>
</tr>
<tr>
<td>NTOL</td>
<td>-</td>
</tr>
<tr>
<td>FL</td>
<td>0.71*</td>
</tr>
<tr>
<td>STR</td>
<td>0.65*</td>
</tr>
</tbody>
</table>

*< 0.01.

Note. Pearson product-moment partial correlations, corrected for age, premorbid IQ, depression, and motor performance. NTOL = New Tower of London task. FL = fluency task. STR = Stroop test. REY = Rey Complex Figure. PP = Purdue Pegboard.

The current battery of executive control function may provide a useful and sensitive assessment of cognitive dysfunction in a wide range of disorders. It may, therefore, be of some use in the increasingly diverse application of neuropsychological assessment to groups with moderate impairment of executive abilities (i.e., as compared to patients with focal frontal lobe injury). Additionally, the current measure of planning appears to be the most effective of these three measures in discriminating among the groups and represents a useful refinement of this paradigm of problem-solving behavior, because previous versions of this task have often failed to discriminate among patients in whom executive control deficits may be subtle, such as in basal ganglia groups. In the study of Robbins, James, Lange, and colleagues (1994), for example, who used the computerised version of the Tower of London to assess cognitive function in Parkinson’s disease, progressive supranuclear palsy, and multiple system atrophy, it was reported that patients with Parkinson’s disease showed normal accuracy on this task. This is not surprising because the measure employed in that study extended to a difficulty of only five moves, a level at which the current Parkinsonian group also appeared to demonstrate largely intact performance. It is only on the more complex items, however, that a problem-solving difficulty in this group became evident. The potential utility of the current measure of planning, therefore, may extend to cases where executive deficits are subtle or where the diagnosis of such impairment with earlier versions of the current measure or other instruments is somewhat unclear.

REFERENCES


