*Journal of the International Neuropsychological Society* (2005), **11**, 376–385.

DOI: 10.1017/S1355617705050423

# Measuring executive dysfunction in an acute rehabilitation setting: Using the dysexecutive questionnaire (DEX)

PAULEEN C. BENNETT,[1] BEN ONG,[2] AND JENNIE PONSFORD[1]

[1]Department of Psychology, School of Psychology, Psychiatry and Psychological Medicine, Monash University, Victoria, Australia
[2]School of Psychological Science, La Trobe University, Victoria, Australia

(RECEIVED May 6, 2004; REVISED February 7, 2005; ACCEPTED February 27, 2005)

**Abstract**

It is recognized that existing neuropsychological measures of executive dysfunction lack adequate sensitivity and selectivity. While attempts have been made to develop improved measures, these have not yet been of great value to those who need to accurately identify executive deficits in a clinical setting. Several behavioral rating scales have been developed for this reason, including the 20-item Dysexecutive Questionnaire (DEX), which forms part of the Behavioral Assessment of the Dysexecutive Syndrome (BADS) test battery. To investigate the ability of the DEX to identify executive dysfunction in an acute rehabilitation setting, the BADS was administered to 64 persons who had sustained traumatic brain injury. It was found to be almost as sensitive to executive dysfunction, as measured by the total score obtained on the BADS battery, as an extended 65-item version of the scale, when completed by either the occupational therapist or clinical neuropsychologist working with each patient. Family members and the patient themselves provided, as expected, less accurate information. Our results indicate that the DEX can be used with some confidence as a screening instrument to identify executive dysfunction in an acute rehabilitation setting, provided it is completed by professional personnel, trained to be sensitive to the cognitive and behavioral concomitants of this disorder. (*JINS*, 2005, *11*, 376–385.)

**Keywords:** Dysexecutive questionnaire (DEX), Executive dysfunction (ED), Behavioral assessment of the dysexecutive syndrome (BADS), Clinical neuropsychology

## INTRODUCTION

While the term "executive dysfunction" (ED) is understood to refer to a constellation of neurobehavioral deficits, arising from disruption to the higher-order cognitive processes required to initiate, plan, execute, and monitor complex goal-directed activities (Banich, 1997; Stuss & Benson, 1986; Tranel et al., 1994), identification of the exact skills involved remains controversial (Crawford, 1998). A number of neuropsychological instruments are commonly used to assess ED. Poor performances on these multidimensional tests, however, can reflect cognitive deficits other than ED (Anderson et al., 1991; Axelrod et al., 1996; Greve et al., 1997; van den Broek et al., 1993). More importantly, existing tests sometimes fail to discriminate between patient groups classified according to whether or not ED is evident behaviorally (Baddeley et al., 1997). Some individual persons who clearly demonstrate ED in everyday life display adequate performances on relevant tests (Eslinger & Damasio, 1985; Goldstein et al., 1993; Heck & Bryer, 1986; Ponsford et al., 1995; Shallice & Burgess, 1991).

If ED represents a multifaceted group of cognitive deficits (Burgess et al., 1998; Stuss & Alexander, 2000), then identification of all cases of ED may require administration of a battery of tests, each sensitive to particular executive deficits. For this reason, Wilson et al. (1996) developed a test battery, the Behavioral Assessment of the Dysexecutive Syndrome (BADS). Each of six subtests is asserted to measure a different aspect of ED, with a composite score able to be calculated and used as an overall measure of ED. The

Reprint requests to: Pauleen C. Bennett, Department of Psychology, School of Psychology, Psychiatry and Psychological Medicine, Monash University, PO Box 197, Caulfield East Victoria 3145, Australia. E-mail: p.bennett@med.monash.edu.au.

BADS also includes a 20-item questionnaire about executive-type behavioral problems, the Dysexecutive Questionnaire (DEX). This can be administered to a brain-injured person or a person familiar with the injured person, to screen for ED.

In initial validation studies (Alderman et al., 1993; Wilson et al., 1996, 1998), the BADS was administered to neurologically intact adults and persons with a variety of neurological disorders, including closed head injury (59%), dementia (13%), and stroke (8.5%). The brain-injured participants obtained significantly lower scores on each subtest, the greatest differentiation between groups being in terms of the composite score. These results confirm that the BADS is sensitive to brain injury. To support the stronger claim that the BADS is sensitive to ED, rather than to brain injury *per se*, Wilson et al. (1996) correlated each BADS score with the DEX score obtained for the patient group. Patient (PT) and family member (FM) DEX ratings were first compared, revealing that the patients identified significantly fewer problems. Because this was thought to indicate lack of insight by the patients, FM-DEX ratings were used for subsequent analyses. The ratings were significantly negatively associated with both the overall BADS score and individual subtest scores. Provided it is accepted that the DEX accurately identifies ED, therefore, these findings indicate that all measures from the BADS are sensitive to ED.

While several studies have now evaluated the BADS in other patient groups (Evans et al., 1997; Norris & Tate, 2000), the utility of the FM-DEX as a "gold-standard" against which the BADS can be tested has not been confirmed. The DEX was not cross-validated with other rating scales reported to be sensitive to ED and it is not clear whether a brief rating scale is sufficiently comprehensive to identify all cases of ED. Also, while existing literature indicates that family members are more reliable than patients in their assessment of many cognitive areas (Prigatano, 1996; Sbordone et al., 1998), the decision to use FM-DEX rather than PT-DEX ratings was made on a *post hoc* basis and there was no opportunity for the authors to establish whether this protocol was justified. Previous reports indicate that family members sometimes fail to identify problems detected by other independent raters (Riccio et al., 1994). In addition, because the DEX fails to specify a time frame within which judgements about the target's behavior should be made, both family members and the patient themselves may be misled into assessing lifetime behaviors rather than behaviors postdevelopment of ED.

To investigate these issues, it is necessary to examine correlations between FM-DEX scores and scores obtained when the family member is specifically directed to assess postmorbid behaviors, or perhaps to compare FM-DEX ratings with those obtained from people only familiar with the postmorbid behavior of the patient, such as rehabilitation specialists. Interestingly, two subsequent studies have reported low correlations between the FM-DEX and the total BADS score, and nonsignificant or even negative correlations between the scale and several BADS subtests (Evans et al., 1997; Norris & Tate, 2000). An additional study found that the DEX, completed by "normal" members of the population or a family member, was not associated with scores obtained by the participants on tests believed to be sensitive to ED (Chan, 2001). Yet another reported that the DEX failed to discriminate between brain-injured persons with executive-like behavioral difficulties and controls (Channon & Crawford, 1999).

Several issues, then, require investigation before the DEX can be accepted as an accurate indicator of ED, against which other measures might be compared. Some of these were addressed in the present study. Persons with traumatic brain injury (TBI) were used as participants because TBI frequently results in focal contusions and diffuse axonal injury in frontal as well as other brain regions (Gale et al., 1995; Ponsford et al., 1995). While ED is a common consequence of TBI, it occurs to varying degrees depending on the nature and location of damage sustained. The sample was expected, therefore, to provide the heterogeneity required for correlational analyses (Tabachnick & Fidell, 1996). TBI offers the additional advantage of providing a relatively young and premorbidly functionally intact population, reducing the possibility of spurious associations due to the documented effects of age on some aspects of executive functioning (Bryan & Luszcz, 2000).

The primary aim in this study was to establish the usefulness of FM-DEX ratings as an indicator of ED. To address this aim, persons who had sustained TBI completed the DEX, as did a family member (FM) and a clinical neuropsychologist (NP) and occupational therapist (OT) engaged in the person's rehabilitation program. Measures were also taken of the severity of injury sustained, the degree to which the person exhibited slowed processing speed and premorbid intelligence. While standard neuropsychological instruments are known to be limited in their ability to provide an accurate quantitative estimate of ED (Lezak, 1995), the neuropsychologist typically detects ED by looking for qualitative signs, evident as the person attempts standard cognitive tests (Ponsford et al., 1995). Each participant in this study completed a neuropsychological assessment or review within the month prior to their participation and some were reasonably well known to the treating neuropsychologist, who often had completed assessments earlier in the person's rehabilitation program. It was anticipated "*a priori*", therefore, that rating scales completed by the treating neuropsychologist would provide an accurate indicator of ED. Occupational therapists perform a functional assessment of each patient upon admission and continue to work with them on a range of activities of daily living. OT ratings were included, therefore, in order to verify whether NP ratings are indicative of functional "ED-type" difficulties and in order to ascertain whether the two professional groups concur in their perception of functional difficulties. Injury severity and impaired processing speed are amongst the more reliable predictors of outcome (Dikmen et al., 1995; Hinton-Bayre et al., 1997) and were expected to be moderately associated with the presence of ED. Premorbid intel-

https://doi.org/10.1017/S1355617705050423 Published online by Cambridge University Press

ligence, by contrast, was expected to be independent of acquired cognitive deficits.

To determine whether the 20-item DEX was sufficiently sensitive to ED to be used as a "gold standard" measure, all participants also completed a more comprehensive scale, developed using the same format and scoring procedure. This allowed for the incorporation of existing DEX items into the extended questionnaire so that participants were only required to complete one self-report measure. To generate additional items for the extended version of the DEX, the eDEX, a comprehensive literature search was conducted. This identified a number of cognitive difficulties commonly associated with ED, as summarized by Banich (1997). A number of questions were formulated to address each of these issues. These and items from the DEX were ranked according to how sensitive they were expected to be to each issue. At least eight items from each group were included in the final eDEX.

A second aim in this study was to examine the strength of association between BADS scores and rating scores. To address this aim all participants were administered the BADS. Because NP- and OT-DEX ratings were expected to be good indicators of ED, it was hypothesized that total scores on the BADS would be strongly associated with these ratings. Moderate correlations between the ratings and individual subtest scores were also expected, and it was anticipated that each subtest would make a significant unique contribution to the prediction of relevant rating scores. FM- and PT-DEX ratings were expected to be less strongly associated with BADS scores.

## METHOD

### Participants

Sixty-four persons participated in this study. All were patients at Bethesda Rehabilitation hospital, in Victoria, Australia. Sixty persons were road accident victims, the remaining four having sustained brain-injury in a workplace accident. Within the sample, the length of time spent in posttraumatic amnesia (PTA) ranged from zero to greater than 3 months ($M = 23$ days, $SD = 23$ days). Glasgow Coma Scale scores obtained at acute hospital admission ranged from 3 to 15 ($M = 7$, $SD = 4$).

Participants were excluded from the study if they did not speak fluent English or were cognitively unable to undertake a formal assessment. Descriptive information was obtained with respect to each person's age (range = 17–73 years, $M = 33$, $SD = 14$), gender (47 male and 17 female) and educational achievements (range = 9–19 years, $M = 12$, $SD = 2.3$). The type of accident experienced was noted, as was the time postinjury at which assessment occurred (range = 15 days–30 years, *Median* = 63 days). All except five participants were seen within one year of injury, 48 remained hospitalized at the time of assessment. Also obtained was information regarding previous difficulties with learning (4), substance use (6), psychopathology (4), neurological illness (2), and head injury (1).

Adequate documentation concerning the location of brain injury was often unavailable, mainly because TBI is typically associated with diffuse microscopic injury not revealed by widely used scanning techniques (Levin & High, 1989; Wilson et al., 1992). Of the 64 participants, 24 had documented damage to anterior sections of the brain (8 bilateral, 8 left hemisphere, and 8 right hemisphere). The remaining 40 had documented lesions elsewhere or no evidence of localized injury.

### Materials

#### *Clinical measures of injury severity*

The duration of PTA (days) reported by each participant was, as suggested by Brooks et al. (1980), used as a measure of injury severity.

#### *Subjective reports of executive dysfunction*

The 20-item DEX was embedded within the 65-item eDEX, which was administered to four persons associated with each TBI participant; the patient (PT-eDEX), a family member (FM-eDEX), a clinical neuropsychologist (NP-eDEX), and an occupational therapist (OT-eDEX). The version administered to the patient was formatted in first person, while that administered to all other participants was formatted in third person.

#### *Neuropsychological measures*

The Symbol Digit Modalities Test (SDMT), as suggested by Hinton-Bayre et al. (1997), was used as a measure of processing speed. This test is sensitive to the effects of generalized cognitive slowing and has strong prognostic significance (Ponsford & Kinsella, 1992). Premorbid intelligence (EST-IQ) was estimated using the National Adult Reading Test (NART: Nelson, 1982). All participants were administered each subtest from the BADS; the Rule Shift Card Test (RULSFT), the Action Program Test (ACTPRO), the Key Search Test (KEYSRT), the Temporal Judgement Test (TEMPJ), the Zoo Map Test (ZOOMAP), and the Modified Six Elements Test (MSET). The scores for each subtest were reduced, as advised by Wilson et al. (1996), to a single score. A previous study has reported that the norms provided for the Temporal Judgement Test may be inappropriate for an Australian population (Norris & Tate, 2000). Because Australian norms are not available, two composite scores for the BADS were calculated. The first (BADS) included the score for the Temporal Judgement Test, calculated using the formula provided by Wilson et al. (1996). The second (MBADS) was comprised of summed profile scores for each of the remaining five BADS subtests.

### Procedure

Over the course of 10 months, all admissions to Bethesda who met the criteria for inclusion in the study were invited to act as participants. Each participant was administered the BADS and then asked to complete the eDEX. The eDEX was also given to a family member, subject to availability, and to the treating clinical neuropsychologist and occupational therapist. These professionals had worked with each patient during their time as an inpatient at the hospital, which ranged from 2 weeks to several months. Some of the longer term patients had been part of the rehabilitation community for several years and were well known to staff members. Others were less well known, but all had been seen by each professional at least three times during their period of admission. All testing was conducted by a single experimenter. The treating neuropsychologist, who was not the experimenter, was asked to complete the rating scale on the basis of their own cognitive assessment of the participant, prior to accessing any test data from this study. Unfortunately, it was not possible for ethical reasons to access previous test data to determine whether any of the participants had been administered any of the BADS subtests in prior assessments, but the number likely to have done so was considered small by the clinicians involved and none had been administered the entire BADS. Also, no participants had been administered any BADS subtests during their most recent assessment.

Midway through the study, administrative changes prevented the continued participation of occupational therapists. OT rating scales were obtained, therefore, for only 45 participants. Available rating data from participants (55) and family members (42) was also reduced due to noncompliance or unavailability. All data analysis was conducted using SPSS for Windows. Missing data were excluded on a pairwise basis. Alpha was set at .05 for all analyses but, due to the large number of comparisons reported, results which remained significant at $\alpha < .01$ are indicated in the relevant tables.

## RESULTS

### The FM-DEX as an Indicator of Executive Dysfunction

The internal consistency of the DEX, as completed by each of the four respondent groups, was examined (Table 1) using Cronbach's $\alpha$ coefficient. All scales had $\alpha$ levels in excess of .91. The eDEX scales were marginally more reliable than corresponding DEX scales, as were the scales completed by professional raters rather than by family members or participants. Table 1 also shows the group data obtained for PTA ($M = 23$ days, $SD = 23$ days), SDMT ($M = 42.41$ s, $SD = 12.79$ s) and EST-IQ ($M = 105$, $SD = 8.3$).

The mean scores from all respondents on both the DEX and eDEX were calculated (Table 1) and analyzed using separate one-way within-group analyses of variance. The mean scores on both scales varied only marginally across groups, although participant ratings, as found in previous studies (Burgess et al., 1998; Knight et al., 2002; Wilson et al., 1996) tended to be lower than those from independent sources. No significant differences were found for DEX ratings [$F(3,202) = .59$, $p = .62$] or eDEX ratings [$F(3,202) = 1.34$, $p = .26$].

Correlations were computed to investigate whether DEX ratings obtained from each respondent group were statistically associated with those obtained from other groups, and with those obtained from the same group using the eDEX (see Table 2). Each scale was also correlated with the two indicators of injury severity and estimated premorbid intelligence.

There was a moderate negative association between length of PTA and scores from the SDMT. Neither variable was associated with EST-IQ. All DEX scores were strongly associated (.96–.98) with corresponding scores on the eDEX. NP and OT ratings on the DEX were strongly associated, both with each other and with corresponding ratings on the eDEX. In addition, DEX and eDEX ratings from both professional groups were moderately associated with the duration of PTA. NP-DEX ratings were also moderately associated with SDMT. While OT-DEX ratings were not statistically associated with this measure, the degree of association approached statistical significance ($p = .07$) and both NP- and OT-eDEX scales were moderately associated with SDMT. NP and OT ratings from the DEX or eDEX were not associated with EST-IQ. PT-DEX and PT-eDEX ratings failed to be significantly associated with ratings provided by professional groups or with SDMT. They were also only weakly correlated with PTA, but were moderately associated with EST-IQ. FM-DEX ratings were moderately associated with NP- and OT-DEX ratings, and with corresponding eDEX ratings. They were not significantly associated with PTA and were only weakly associated with SDMT, both expected *a priori* to be associated with ED. FM-DEX ratings were associated with EST-IQ, and with PT-DEX and PT-eDEX ratings, variables not thought to be indicative of acquired ED.

**Table 1.** Means and standard deviations of DEX scales (including indices of internal consistency) and measures of injury severity, processing speed, and estimated premorbid IQ

| Measure | *N* | *M* | *SD* | Standardized $\alpha$ |
|---|---|---|---|---|
| FM-DEX | 42 | 20.64 | 14.72 | .93 |
| NP-DEX | 64 | 19.35 | 16.04 | .95 |
| OT-DEX | 45 | 21.71 | 16.85 | .94 |
| PT-DEX | 55 | 18.02 | 13.16 | .92 |
| FM-eDEX | 42 | 66.33 | 43.91 | .98 |
| NP-eDEX | 64 | 71.01 | 54.09 | .99 |
| OT-eDEX | 45 | 76.40 | 58.01 | .99 |
| PT-eDEX | 55 | 59.24 | 39.95 | .97 |
| PTA (days) | 62 | 23.28 | 22.79 | |
| SDMT | 64 | 42.41 | 12.64 | |
| EST-IQ | 64 | 105.30 | 8.29 | |

**Table 2.** Correlations between DEX scores obtained from each respondent group and clinical measures of injury severity, processing speed and estimated premorbid IQ

| Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 FM-DEX | 1.00 (42) | .42** (42) | .45* (30) | .68** (42) | .08 (40) | −.33* (42) | −.49** (42) |
| 2 NP-DEX | | 1.00 (64) | .79** (45) | .24 (55) | .53** (62) | −.37** (64) | −.14 (64) |
| 3 OT-DEX | | | 1.00 (45) | .23 (40) | .59** (43) | −.27 (45) | −.23 (45) |
| 4 PT-DEX | | | | 1.00 (55) | .20 (53) | −.20 (55) | −.39** (55) |
| 5 PTA (days) | | | | | 1.00 (62) | −.47** (62) | −.07 (62) |
| 6 SDMT | | | | | | 1.00 (64) | .23 (64) |
| 7 EST-IQ | | | | | | | 1.00 (64) |
| 8 FM-eDEX | .96** (42) | .42** (42) | .50** (30) | .68** (42) | .14 (40) | −.33* (42) | −.45** (42) |
| 9 NP-eDEX | | .98** (64) | .79** (45) | .21 (55) | .56** (62) | −.43** (64) | −.13 (64) |
| 10 OT-eDEX | | | .98** (45) | .22 (40) | .65** (43) | −.34* (45) | −.25 (45) |
| 11 PT-eDEX | | | | .97** (55) | .28* (53) | −.21 (55) | −.39** (55) |

*$p < .05$, **$p < .01$. Numbers in parentheses indicate $n$ following pairwise deletion of missing data.

## Predicting DEX Ratings with the BADS

The reliability of the BADS was assessed statistically. For the 61 participants with valid data for all six subtests, the results achieved a standardized Cronbach $\alpha$ coefficient of .60. This indicates that the total profile score from the BADS is not a particularly reliable measure of ED, although a moderate Cronbach $\alpha$ coefficient is desirable in a test battery designed to assess a multidimensional construct. Because the validity of the Temporal Judgement Test for an Australian population has been challenged (Norris & Tate, 2000), the reliability of combining the other five subtests was also calculated. This marginally improved the standardized $\alpha$, but it remained low (.63).

To determine whether measures from the BADS were associated with scores from the NP and OT rating scales, the appropriate DEX and eDEX scales were correlated with all variables from the BADS (Table 3). FM- and PT-DEX ratings were included to facilitate direct comparisons with previous studies. Because very similar results were obtained with corresponding DEX and eDEX instruments, to simplify the data presented and facilitate direct comparisons with previous reports, only DEX scores are reported. A moderate association was found between NP- or OT-DEX ratings and the total BADS score. The degree of association between the total BADS score and FM-DEX ratings or PT-DEX ratings was poor and did not attain statistical significance. Removing the Temporal Judgement Test from the total BADS score failed to strengthen the association between NP- or OT-DEX ratings and the total score for the remaining BADS subtests.

Only three individual subtests from the BADS were significantly associated with the NP-DEX and OT-DEX; RULSFT, ACTPRO, and MSET. A moderate degree of association was found for all of these variables except ACTPRO, which was moderately associated with NP-DEX ratings but more strongly associated with OT-DEX ratings. The magnitude of association between rating scores and the MSET was the same for both groups of professional raters, although this was more significant statistically for the NP-DEX ratings because of higher participant numbers. The other three BADS subtests, KEYSRT, TEMPJ, and ZOOMAP, achieved only poor correlations with both NP- and OT-DEX ratings.

No single test was significantly associated with FM-DEX ratings, although RULSFT approached significance with a weak correlation of $-.29$ ($p = .06$). PT-DEX ratings were not statistically associated with any of the test variables. Consistent with the results of the reliability analysis, between subtest correlations were moderate at best. RULSFT was significantly associated with ACTPRO, KEYSRT and ZOOMAP, ACTPRO also being associated with the KEYSRT and MSET. Consistent with claims that the Temporal Judgement Test may be inappropriate for an Australian population (Norris & Tate, 2000), it was not associated with any other subtest and only weakly associated with the BADS score.

https://doi.org/10.1017/S1355617705050423 Published online by Cambridge University Press

**Table 3.** Correlations between DEX scores and BADS tests

| Measure | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|
| 1 FM-DEX | −.14 | −.09 | −.29 | −.17 | −.03 | −.17 | −.06 | .06 |
| | (42) | (42) | (42) | (42) | (42) | (42) | (42) | (41) |
| 2 NP-DEX | −.37** | −.32** | −.32** | −.36** | −.17 | −.19 | −.09 | −.36** |
| | (64) | (64) | (63) | (63) | (64) | (64) | (64) | (61) |
| 3 OT-DEX | −.39** | −.39** | −.39** | −.51** | −.14 | −.09 | −.08 | −.36* |
| | (45) | (45) | (45) | (45) | (45) | (45) | (45) | (45) |
| 4 PT-DEX | −.13 | −.12 | −.19 | −.09 | −.17 | −.05 | .09 | −.04 |
| | (55) | (55) | (55) | (55) | (55) | (55) | (55) | (55) |
| 5 BADS | 1.00 | .96** | .76** | .63** | .57** | .28* | .53** | .49** |
| | (64) | (64) | (63) | (63) | (64) | (64) | (64) | (61) |
| 6 MBADS | | 1.00 | .77** | .66** | .59** | −.00 | .55** | .54** |
| | | (64) | (63) | (63) | (64) | (64) | (64) | (61) |
| 7 RULSFT | | | 1.00 | .50** | .51** | .09 | .38** | .21 |
| | | | (63) | (63) | (63) | (63) | (63) | (61) |
| 8 ACTPRO | | | | 1.00 | .32* | −.00 | .24 | .39** |
| | | | | (63) | (63) | (63) | (63) | (61) |
| 9 KEYSRT | | | | | 1.00 | .01 | .15 | .15 |
| | | | | | (64) | (64) | (64) | (61) |
| 10 TEMPJ | | | | | | 1.00 | .02 | −.07 |
| | | | | | | (64) | (64) | (61) |
| 11 ZOOMAP | | | | | | | 1.00 | .07 |
| | | | | | | | (64) | (61) |
| 12 MSET | | | | | | | | 1.00 |
| | | | | | | | | (61) |

*$p < .05$, **$p < .01$. Numbers in parentheses indicate *n* following pairwise deletion of missing data.

The three subtests found to be significantly associated with NP-DEX ratings were almost as strongly associated with this measure as was the total BADS score. A similar situation occurred with respect to OT-DEX ratings, except that the association obtained for ACTPRO did exceed that obtained for both the standard and modified composite BADS scores. To begin to address the question of unique predictive validity, regression analyses were conducted, in which the BADS subtest scores were entered as potential predictors of scores on the NP- or OT-DEX (Table 4).

The regression analyses (Table 4) confirm that the combined BADS subtests act as a significant predictor of NP-DEX ratings ($R^2 = .25$, $F(6,54) = 3.04$, $p = .01$) and OT-DEX ratings ($R^2 = .34$, $F(6,38) = 3.24$, $p = .01$). Analy-

**Table 4.** Standard multiple regression of BADS subtests predicting NP- and OT-DEX ratings

| Neuropsychological measures | $\beta$ | *r* | *sr* | *t* | *p* |
|---|---|---|---|---|---|
| Predictors of NP-DEX ratings | | | | | |
| Rule Shift Card Sort Test | −.19 | −.32 | −.14 | −1.20 | .24 |
| Action Program Test | −.18 | −.36 | −.14 | −1.20 | .23 |
| Key Search Test | .02 | −.17 | .02 | .15 | .88 |
| Temporal Judgement Test | .20 | −.19 | −.19 | −1.68 | .09 |
| Zoo Map Test | .04 | −.09 | .04 | .34 | .74 |
| Modified Six Element Test | −.27 | −.36 | −.25 | −2.08 | .04* |
| Predictors of OT-DEX ratings | | | | | |
| Rule Shift Card Sort Test | −.26 | −.39 | −.19 | −1.47 | .15 |
| Action Program Test | −.37 | −.51 | −.29 | −2.26 | .03* |
| Key Search Test | .12 | −.14 | .16 | .80 | .43 |
| Temporal Judgement Test | −.08 | −.09 | −.08 | −.59 | .56 |
| Zoo Map Test | .11 | −.08 | .09 | .74 | .46 |
| Modified Six Element Test | −.19 | −.36 | −.17 | −1.29 | .20 |

*$p < .05$.

https://doi.org/10.1017/S1355617705050423 Published online by Cambridge University Press

sis of the contribution made by each test, however, demonstrated that only one BADS measure, MSET, made a significant unique contribution to NP-DEX ratings. Similarly, only one BADS measure made a significant unique contribution to OT-DEX ratings, although, interestingly, the measure found to be significant was ACTPRO.

## DISCUSSION

The primary aim in this study was to establish whether DEX ratings, particularly FM-DEX ratings, provide an accurate source of information about the presence of ED. The DEX was found to be a statistically reliable scale. In addition, DEX ratings by clinical neuropsychologists and occupational therapists, each familiar with the participant and each qualified to assess the cognitive and functional consequences of TBI, were found to be strongly associated. NP- and OT-DEX ratings were moderately associated with impaired processing speed and length of PTA, both good predictors of outcome following TBI. They were not, however, associated with estimated premorbid intelligence. Scores from the DEX were strongly associated with corresponding scores from the eDEX, a comprehensive 65-item scale developed on the basis of a functional model of ED formulated by Banich (1997).

While mean family member rating scores for both the DEX and eDEX were statistically indistinguishable from those provided by neuropsychologists and occupational therapists, FM ratings on both scales were only moderately associated with ratings provided by professionals. In addition, FM ratings were moderately negatively associated with estimated premorbid intelligence, and were associated only weakly with impaired processing speed and not at all with PTA. This indicates that the family members in this study did not provide an accurate assessment of the participant's postinjury level of functioning.

This finding is potentially problematic, as family members and/or carers have been used in previous studies to complete the DEX (Burgess et al., 1998; Chan, 2001; Channon & Crawford, 1999; Evans et al., 1997; Norris & Tate, 2000; Wilson et al., 1996). It is not surprising, however, that family members in this study were less sensitive to ED than were professional persons trained to identify this disorder. The population was relatively acute, with many participants remaining hospitalized at the time of assessment. Even a family member who visits the TBI patient on a regular basis is unlikely to observe subtle signs of ED while the patient remains in a structured setting, since this environment obviates the need for "executive" skills. In addition, the instructions for completing the DEX do not specify a time frame, so it is possible that at least some family members adopted a long-term perspective, assessing the patient's premorbid behavior rather than their behavior following TBI. This is consistent with the finding that family member ratings were moderately negatively associated with estimated premorbid intelligence and could easily be avoided by providing more specific instructions with the DEX. The results, however, serve to caution against using family member ratings without careful consideration of the population being studied, and may partially account for conflicting results obtained previously.

In most previous studies using the DEX, the brain-injured participants came from a chronic population. Family members were therefore probably better able to assess their daily functional capacity, and it would be more appropriate to adopt a long-term perspective when doing so. This might explain why FM-DEX ratings have previously been reported to be strongly associated with other putative measures of ED (Burgess et al., 1998), including the BADS total score (Wilson et al., 1996). In one study, however, a poor correlation between FM-DEX ratings and the total BADS score was obtained for one of two patient groups (Evans et al., 1997). This may indicate that FM-DEX ratings failed to provide an appropriate reference point in this population, or that the BADS was inaccurate as a measure of ED. Another relevant study by Norris and Tate (2000) tested patients with either TBI or multiple sclerosis. This study found that neither the overall BADS score nor five of the subtest scores were significantly associated with FM-DEX ratings. Even more problematic, the significant association between the final subtest and the FM-DEX was opposite to that expected. Again, therefore, this may indicate that the FM-DEX is not always the best measure against which to evaluate test performances or, alternatively, that the tests used were not sensitive to ED, as measured by the DEX in these samples.

If it is not appropriate to assume that FM-DEX ratings are always an accurate indicator of ED, then an alternative source of information is required in order to evaluate the usefulness of formal testing instruments. Family members are generally acknowledged to be more accurate raters than are participants (Evans et al., 1997; Prigatano, 1996; Sbordone et al., 1998) and this was confirmed in this study, with participant ratings not being associated with any other indicator of ED. In the present study, however, a far more accurate assessment was found to be provided by the DEX (and eDEX) ratings provided by the treating clinical neuropsychologist and occupational therapist, an approach previously suggested by Channon and Crawford (1999) and employed by Knight et al. (2002).

Unlike family members, the neuropsychologists in our study had the advantage of having conducted at least one formal assessment of each patient, using standard cognitive tests. This has the potential to confound the use of neuropsychologists ratings as a "gold standard" against which other ED measures can be compared, since their ratings may at least partially reflect the same test results being evaluated as putative indicators of ED. To guard against this possibility, the neuropsychologists in this study were required to complete the rating scales prior to accessing the test data collected as part of the study. While they did potentially have access to test data obtained during previous assessments for some participants, the entire BADS had never been administered previously to any participant. At most, a

https://doi.org/10.1017/S1355617705050423 Published online by Cambridge University Press

small number of participants may have previously been administered one or two individual subtests from the BADS (Zoo Map Test and Key Search Test). This could not be ascertained but no BADS subtest had been administered within the most recent assessment.

In addition, while the occupational therapists had access to the neuropsychologist's report for some participants, they had no direct access to neuropsychological test data. The fact that rating scores from the two professional groups were so strongly associated, and that many of the same measures were found to be associated with both the NP- and OT-DEX, therefore strengthens the argument that the professional rating scores provided unbiased and accurate indicators against which test performances could be assessed. ED-related behavior observed by occupational therapists is clearly reasonably consistent with that identified by neuropsychologists on the basis of a formal assessment, although those few differences observed may indicate subtle differences in the way that particular behaviors are perceived and this should be followed up in subsequent research.

A secondary aim in this study was to determine the sensitivity of the BADS to ED, as measured using the NP- and OT-DEX scores. Moderate negative associations between the BADS total score and the NP- and OT-DEX were found, supporting the hypothesis that the BADS may be reasonably sensitive to ED. The degree of association between several subtests and NP- or OT-DEX ratings was equally if not more robust, however, indicating that some of the subtests may be effective in isolation. Of more concern, three BADS subtests were not significantly associated with either NP- or OT-DEX ratings. Moreover, stepwise regression analyses revealed that only one subtest, the Modified Six Elements Test, made a significant unique contribution to the prediction of NP-DEX scores. Similarly, only the Action Program Test made a significant unique contribution to the prediction of OT-DEX ratings. These results indicate that the BADS battery may be of only limited utility to those seeking to identify ED in a clinical population; one or two subtests from the BADS may be just as sensitive to ED as is the entire battery.

If ED represents a multidimensional construct, then it is likely and indeed desirable, that individuals with different types of ED should fail only selected tests within a comprehensive test battery. However, the population in this study was particularly selected because of the known heterogeneity of ED associated with TBI. Thus, it is difficult to explain why this study should fail to find significant correlations between each subtest and the DEX, when these were reported by Wilson et al. (1996) in a similar population. Differences on the Temporal Judgement Test can probably be attributed to the use of norms inappropriate to an Australian population, as reported by Norris and Tate (2000). This cannot account, however, for the failure of the Zoo Map Test and Key Search Test to be associated with NP- or OT-DEX ratings.

Several previous studies that have directly compared BADS scores obtained by brain-injured patients and intact controls have reported similarly disturbing results. For example, while Knight et al. (2002) found that most BADS subtests correlated with the DEX, with the Zoo Map test being most sensitive, Evans et al. (1997) found that only the Action Program Test, Temporal Judgement Test, and Modified Six Elements Test discriminated between their sample of persons with schizophrenia and normal controls. Norris and Tate (2000) also found that only three BADS subtests (Action Program test, Modified Six Elements Test, and Zoo Map Test) discriminated between a sample comprised of people with either TBI or multiple sclerosis and uninjured controls, even though the groups were distinguished by several other putative tests of ED. A study by Chan (2001) found no association between the Modified Six Elements Test and DEX ratings in a normal population. Differences between available studies may be partially attributable to the exact type of ED associated with specific clinical populations, although these are not sufficiently characterized as yet to permit this hypothesis to be tested. At present, it would clearly be remiss to conclude that measurement difficulties traditionally associated with ED have been satisfactorily resolved.

The results of this study have significant implications. Importantly, while neuropsychologists often reported only limited contact with the participants in this study, sometimes as little as three one-hour assessment sessions, neuropsychologists' ratings were strongly associated with objective measures of injury severity and with ratings provided by occupational therapists, who often worked with the participants on a daily basis. This suggests that even standard neuropsychological instruments, while not providing quantitative evidence of ED, may provide a context within which subtle cognitive deficits can be noted by a trained observer. In their original description of the Six Elements Test and Multiple Errands Test, Shallice and Burgess (1991) emphasized the importance of observing qualitative difficulties on these tasks, a view echoed by Knight et al. (2002). Perhaps, as previously suggested by Rosenthal and Millis (1992), therefore, future research should concentrate not so much on criticizing the utility of existing measures of ED, but on examining how these measures might best be scored and interpreted so as to convey the information they elicit. Assessment of subtle functional deficits should take into account not only whether a person is able to perform a test or activity, for example, but also the manner by which they reach the solution, and also perhaps how well they may have been expected to perform given premorbid skills and personal characteristics.

The present study was limited by the use of a relatively acute sample, this being necessary to obtain participants familiar to both a neuropsychologist and occupational therapist. It was also limited by the use of a pure TBI sample since, if ED is as multifactorial as claimed (Burgess et al., 1998; Stuss & Alexander, 2000), it might be expected that other aetiologies may lead to systematic differences in the way that ED is expressed behaviorally. It would be useful to collect sufficient data so that participant groups could be analyzed according to the location of brain injury sustained. One might well find that ED is manifest differently according to the exact location of damage (see Stuss & Alexander, 2000, and Burgess et al., 2000, for discussions relevant to this issue) but this could not be addressed in this study. Recent attempts to factor analyze multiple measures from individual neuropsychological tests sensitive to ED suggest that a more comprehensive analysis of available indicators, requiring far more participants that were available in the present study, may prove useful in identifying the range of skills which contribute to ED (Greve et al., 1997).

In summary, family members may not provide an accurate source of information concerning the presence of ED in a population of TBI patients presenting for rehabilitation, many still in an acute stage of recovery. This is particularly problematic when an instrument is used that fails to specify the period during which the patient's behavior should be assessed. With the exception of this oversight in the instructions for the scale, the DEX seems to provide a useful instrument. When administered to an appropriate respondent it can be used as an indicator of ED, against which other putative indicators can be evaluated. Using this approach, it seems that the BADS contains several useful subtests. Other components of the BADS may be less useful and, in a TBI sample at least, may not add to the sensitivity of the overall battery. When used judiciously a combination of neuropsychological instruments may facilitate detection of ED. It seems, however, that qualitative information may be equally useful, perhaps because the concept of ED is simply not yet sufficiently specified to permit a rigorous, quantitative approach. Executive difficulties in the initiation and regulation of complex behaviors are likely to reflect a number of syndromes, not yet clearly defined or able to be distinguished empirically. If so, then it is not surprising that different studies, using different participant populations and different combinations of tests, continue to report conflicting results. Future research, using larger sample numbers and either qualitative data analyses or multivariate statistical techniques, is needed to address a number of outstanding issues.

## ACKNOWLEDGMENTS

We gratefully acknowledge the participants in this study, who gave willingly of their time and energy during a period of personal crisis. We also thank the professional raters who contributed to the study and to three anonymous reviewers whose insightful comments significantly improved our interpretation of the data collected.

## REFERENCES

Alderman, N., Evans, J.J., Burgess, P., & Wilson, B.A. (1993). Behavioural assessment of the dysexecutive syndrome. *Journal of Clinical and Experimental Neuropsychology*, *15*, 69–70.

Anderson, S.W., Damasio, H., Jones, R.D., & Tranel, D. (1991). Wisconsin Card Sorting Test performance as a measure of frontal lobe damage. *Journal of Clinical and Experimental Neuropsychology*, *13*, 909–922.

Axelrod, B.N., Goldman, R.S., Heaton, R.K., Curtiss, G., Thompson, L.L., Chelune, G.J., & Kay, G.G. (1996). Discriminability of the Wisconsin Card Sorting Test using the standardization sample. *Journal of Clinical and Experimental Neuropsychology*, *18*, 338–342.

Baddeley, A., Della Sala, S., Papagno, C., & Spinnler, H. (1997). Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion. *Neuropsychology*, *11*, 187–194.

Banich, M.T. (1997). *Neuropsychology: The neural basis of mental function*. Boston, Massachusetts: Houghton Mifflin.

Brooks, D.N., Aughton, M.E., Bond, M.R., Jones, P., & Rizvi, S. (1980). Cognitive sequelae in relationship to early indices of severity of brain damage after severe blunt head injury. *Journal of Neurology, Neurosurgery and Psychiatry*, *43*, 529–534.

Bryan, J. & Luszcz, M.A. (2000). Measurement of executive function: Considerations for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology*, *22*, 40–55.

Burgess, P.W., Alderman, N., Evans, J., Emslie, H., & Wilson, B.A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*, 547–558.

Burgess, P.W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*, 848–863.

Chan, R.C.K. (2001). Dysexecutive symptoms among a nonclinical sample: A study with the use of the dysexecutive questionnaire. *British Journal of Psychology*, *92*, 551–565.

Channon, S. & Crawford, S. (1999). Problem-solving in real-life-type situations: The effect of anterior and posterior lesions on performance. *Neuropsychologia*, *37*, 757–770.

Crawford, J.R. (1998). Introduction to the assessment of attention and executive functioning. *Neuropsychological Rehabilitation*, *8*, 209–211.

Dikmen, S.S., Machamer, J.E., Winn, H.R., & Temkin, N.R. (1995). Neuropsychological outcome at 1-year post head injury. *Neuropsychology*, *9*, 80–90.

Eslinger, P.J. & Damasio, A.R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, *35*, 1731–1741.

Evans, J.J., Chua, S.E., McKenna, P.J., & Wilson, B.A. (1997). Assessment of the dysexecutive syndrome in schizophrenia. *Psychological Medicine*, *27*, 635–646.

Gale, S.D., Johnson, S.C., Bigler, E.D., & Blatter, D.D. (1995). Nonspecific white matter degeneration following traumatic brain injury. *Journal of the International Neuropsychological Society*, *1*, 17–28.

Goldstein, L.H., Bernard, S., Fenwick, P.B., Burgess, P.W., & McNeil, J. (1993). Unilateral frontal lobectomy can produce strategy application disorder. *Journal of Neurology, Neurosurgery and Psychiatry*, *56*, 274–276.

Greve, K.W., Brooks, J., Crouch, J.A., Williams, M.C., & Rice, W.J. (1997). Factorial structure of the Wisconsin Card Sorting Test. *British Journal of Clinical Psychology*, *36*, 283–285.

Heck, E.T. & Bryer, J.B. (1986). Superior sorting and categorizing ability in a case of bilateral frontal atrophy: An exception to the rule. *Journal of Clinical and Experimental Neuropsychology*, *8*, 313–316.

Hinton-Bayre, A.D., Geffen, G., & McFarland, K. (1997). Mild head injury and speed of information processing: A prospective study of professional rugby league players. *Journal of Clinical and Experimental Neuropsychology*, *19*, 275–289.

Knight, C., Alderman, N., & Burgess, P.W. (2002). Development of a simplified version of the multiple errands test for use in hospital settings. *Neuropsychological Rehabilitation*, *12*, 231–255.

Levin, H.S. & High, W.M. (1989). Contributions of neuroimaging to neuropsychological research on closed head injury. *Neuropsychology*, *3*, 243–254.

Lezak, M.D. (1995). *Neuropsychological assessment*. New York: Oxford University Press.

Nelson, H.E. (1982). *National Adult Reading Test (NART): Test Manual*. Windsor, Berks (UK): NFER-Nelson.

Norris, G. & Tate, R.L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, concurrent and construct validity. *Neuropsychological Rehabilitation*, *10*, 33–45.

Ponsford, J.L. & Kinsella, G. (1992). Attentional deficits following closed-head injury. *Journal of Clinical and Experimental Neuropsychology*, *14*, 822–838.

Ponsford, J.L., Sloan, S., & Snow, P. (1995). *Traumatic brain injury: Rehabilitation for everyday adaptive living*. Hove (UK): Lawrence Erlbaum Associates.

Prigatano, G.P. (1996). TBI patients tend to underestimate: A replication and extension to patients with laterlized cerebral dysfunction. *Clinical Neuropsychologist*, *10*, 191–201.

Riccio, C.A., Hall, J., Morgan, A., & Hynd, G.W. (1994). Executive function and the Wisconsin Card Sorting Test: Relationship with behavioural ratings and cognitive ability. *Developmental Neuropsychology*, *10*, 215–229.

Rosenthal, M. & Millis, S.R. (1992). Relating neuropsychological indicators to psychosocial outcome after traumatic brain injury. *NeuroRehabilitation*, *2*, 1–8.

Sbordone, R.J., Seyranian, G.D., & Ruff, R.M. (1998). Are the subjective complaints of traumatically brain injured patients reliable? *Brain Injury*, *12*, 505–515.

Shallice, T. & Burgess, P.W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*, 727–741.

Stuss, D.T. & Alexander, M.P. (2000). Executive functions and the frontal lobes: A conceptual review. *Psychological Research*, *63*, 289–298.

Stuss, D.T. & Benson, D.F. (1986). *The frontal lobes*. New York: Raven Press.

Tabachnick, B.G. & Fidell, L.S. (1996). *Using multivariate statistics*. California: Harper Collins.

Tranel, D., Anderson, S.W., & Benton, A.L. (1994). Development of the concept of 'executive function' and its relationship to the frontal lobes. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 125–148). Amsterdam: Elselvier.

van den Broek, MD, Bradshaw, C.M., & Szabadi, E. (1993). Utility of the Modified Wisconsin Card Sorting Test in neuropsychological assessment. *British Journal of Clinical Psychology*, *32*, 333–343.

Wilson, B.A., Alderman, N., Burgess, P.W., Emslie, H., & Evans, J.J. (1996). *Behavioural assessment of the dysexecutive syndrome: Test manual*. England: Thames Valley Test Company.

Wilson, B.A., Evans, J.J., Emslie, H., Alderman, N., & Burgess, P. (1998). The development of an ecologically valid test for assessing patients with dysexecutive syndrome. *Neuropsychological Rehabilitation*, *8*, 213–228.

Wilson, J.T., Hadley, D.M., Wiedmann, K.D., & Teasdale, G.M. (1992). Intercorrelation of lesions detected by magnetic resonance imaging after closed head injury. *Brain Injury*, *6*, 391–399.