Behavioural Assessment of the Dysexecutive Syndrome (BADS)

by Barbara A Wilson, Nick Alderman, Paul W Burgess, Hazel Emslie, Jonathan J Evans
(published by Thames Valley Test Company, 1996 version)

Reviewed by Elaine Chamberlain
Rehab UK, London Brain Injury Vocational Centre, 21 St Thomas's Street, London

PURPOSE

This is a battery of tests designed to assess the effects of dysexecutive syndrome, a cluster of impairments generally associated with damage to the frontal lobes of the brain. These impairments include difficulties with high-level tasks such as planning, organising, initiating, monitoring and adapting behaviour. The tests have also been used in research with people with schizophrenia to assess the same characteristics.

The tests were developed in response to concerns over the low ecological validity of neuropsychological measures. Ecological validity refers to the extent to which the results of controlled tests generalise onto performance in naturalistic settings. One of the arguments against these measures is that, while neuropsychological tests assess single, isolated abilities, real-life tasks involve a complex, messy array of skills. The authors of the BADS assert that their tests assess executive functioning in more complex, real life situations than do established measures and thereby improve their ability to predict day-to-day difficulties.

FORMAT

The BADS is a battery of six tests and two questionnaires. The tests require participants to plan, initiate, monitor and adjust behaviour in response to the explicit and implicit demands of a series of tasks. A profile score, ranging from 0–4, is calculated for each test and an overall profile score is produced as a sum of individual test scores. Profile scores can be converted to standard scores with a mean of 100 and a standard deviation of 15, enabling comparisons to be made with WAIS and WMS scores. This makes it possible to identify discrepancies between expected and achieved BADS scores and to classify performance in a similar qualitative manner, from impaired to very superior. The

authors recommend that all six tests are administered but state that the final score can be prorated based on five. The questionnaire scores are not standardised and do not contribute to the final score.

1. Rule Shift Cards (RS) – This test purports to identify perseverative tendencies and its obverse, mental flexibility (perseveration refers to a difficulty in adjusting behaviour to meet the demands of a changing situation). It requires participants to respond to stimuli (red or black playing cards) according to one of two rules that are presented consecutively. Performance is scored according to how successfully the respondent shifts from applying the first to the second rule. A penalty is imposed for lack of speed.

2. Action Programme (AP) – This test was designed to assess ability to devise and implement a solution to a practical problem (getting a cork out of a narrow plastic tube) while not contravening a set of rules. The score is based on the number of steps completed without assistance. Penalties are imposed for rule breaks.

3. Key Search (KS) – It is claimed that this test assesses ability to plan a strategy to solve a problem (finding a key lost in a field). The score is based on a number of criteria, including whether the rater believes the strategy to be systematic, efficient and likely to be effective. A penalty is imposed for lack of speed.

4. Temporal Judgement (TJ) – It is not made clear what this test has been designed to assess. It appears to involve judgement and abstract thinking based on common knowledge, as the respondent is required to estimate times for everyday events, such as the life span of a dog. The score is based on the accuracy of the estimate.

5. Zoo Map (ZM) – This is a test to assess ability independently to formulate and implement a plan (high demand condition) and to follow a pre-

formulated plan (low demand condition). It involves plotting or following a route through a map that does not contravene a set of rules. The score is based on the successful implementation of the plan. Penalties are imposed for rule breaks and lack of speed.

6. Modified Six Elements (6E) – This test was designed to assess the ability to time-manage. It involves dividing the available time between a number of simple tasks (picture naming, arithmetic and dictation) while not contravening a set of rules. The score is based on the number of tasks attempted. Penalties are imposed for rule breaks and for sharing time unequally between tasks.

7. Dysexecutive Questionnaire (DEX) – This is a 20-item questionnaire describing behaviour associated with dysexecutive syndrome. Ratings of the frequency with which the particular behaviour occurs are made on a Likert-type scale (from 'never' to 'often'). There is one version for the respondent to use and a second version for a family member, colleague or carer to rate the respondent.

Generally, the format of the tests has been thoughtfully designed. For example, certain tests (such as 6E) have particularly complex instructions that can be difficult to grasp by participants with memory or receptive language difficulties. Some appropriate safeguards have been put into place (such as providing summarised test instructions on printed cards or requiring the respondent to paraphrase the instructions to the tester to confirm understanding).

There are, however, some exceptions to this. For example, ZM requires the respondent to change pen each time he or she has 'visited' a location on the map. This introduces an extraneous demand for attentional switching that is likely to be particularly difficult for those with executive difficulties. Furthermore, some of the administrative instructions lack precision. For example, in 6E, participants are not told to divide their time equally between the tasks, and yet they are penalised for not doing so. TJ also contains questions that may make it culturally biased (this is acknowledged by the authors who suggest that it be an optional test). There are also no practice questions by which to confirm understanding. These deficiencies cannot be remedied by the tester without compromising the validity of this standardised test.

Some tests are also awkward to administer. For example, administration of RS requires the tester to turn cards

quickly with one hand, write responses with the other and have the question sheet, manual and stopwatch in view. However, smooth delivery can be achieved with a degree of practice.

Most tests are straightforward to score, with the exception of KS, which takes some thought and practice.

COST OF MATERIALS

- £275 for the BADS complete pack (containing stimulus cards, manual, stimulus books, materials for action programme test, 25 each of self- and other-rated questionnaires and 25 scoring sheets)
- £15 for BADS scoring sheets (pack of 25)
- £15 for DEX questionnaires (pack of 2 x 25)

Therefore, initial outlay is £11 per respondent, dropping to £1.20 per respondent after the first 25 (or, if the DEX is not administered, the price drops to £0.60 per respondent).

All prices are excluding VAT.

Extra equipment is required, such as a tape recorder, a range of coloured pens or pencils, an eraser, water and access to a photocopier.

TIMINGS

All tests are untimed, bar 6E. Most take 5–10 minutes to administer and the time taken to administer the entire battery is approximately 30–45 minutes without the questionnaire. Time for scoring is approximately 15 minutes.

RELEVANT OCCUPATIONS

This test is designed for use with individuals with brain injury and therefore is not aimed at a particular occupational group. However, effective executive functioning is important in any job in which the incumbent is not closely and continually supervised and that requires self-monitoring (for accuracy, for example). It is particularly important in work that entails planning and organisation, such as managerial or supervisory roles.

TECHNICAL INFORMATION

Norms
The control group consisted of 216 non-brain-injured participants with a range of abilities and ages (from 16 to

over 64). It contained an equal number of men and women. Socio-economic status also seems to have been considered but details of this are not provided in the manual.

The criterion group consisted of 92 participants with a variety of neurological disorders (most with closed head injury, others with brain injury resulting from encephalitis or stroke and some with dementia). Only 78 completed all 6 tests (reasons for this are not given). They were not age-matched with the control group (being generally older), but were matched in terms of ability level. The number of individuals who completed the tests is therefore small and heterogeneous, potentially affecting the reliability of the results.

Reliability

To assess inter-rater reliability, two raters scored the performance of a small sample from the control group. Impressive correlations were found between ratings (0.88–1.00), some of the lower correlations being for timings.

To assess test-retest reliability, a small sample from the control group was re-tested 6–12 months after the first session. A substantial practice effect was found (correlations between scores ranging between –0.08 and 0.71 and only three correlations were significant (AP, KS, TJ). The authors attribute the poorest correlations to the effects of outliers in the small sample. They also argue that test-retest reliability is an inappropriate measure for tests of executive functioning which aim to assess ability to solve novel problems. If the respondent takes the test twice, this essential element of novelty is removed and so a practice effect would be expected. This is a valid argument that raises important questions about how to test stability in any test of problem-solving. In apparent support of this assertion, the test-retest reliability was comparable to the performance of the same sample on three established tests of executive functioning (Cognitive Estimates Test, Modified Card Sorting Test and Verbal Fluency Test).

Reliability for the brain-injured participants was not reported in the manual and, given that memory difficulties might attenuate the practice effect, this would have been a useful addition.

It is worth noting that the methods used to assess reliability were surprising. For example, test-retest reliability was confirmed by quoting the percentage of absolute agreement between the two sets of scores; the more usual practice being to examine the degree of correlation between the two sets. There also appears to have been an element of 'data-mining', with the consequent risks of Type I errors. This is neither corrected for nor acknowledged and undermines confidence in the reliability results reported.

Validity

Face validity: The BADS tests involve apparently real-life problems and these generally have a more 'practical' feel than comparable tests. In my experience, individuals generally engage well with the tasks.

Construct validity: The authors found that the brain-injured group achieved significantly lower scores on the test than did the control group. This remained when the effect of the age difference between the two samples was controlled for. A further small-scale study confirmed this, finding significant group differences on AP, ZM and 6E (Norris and Tate, 2000). Furthermore, regression analysis in this study produced a hit rate of 74% in correctly identifying group membership from test scores (although 6E was the only test that contributed significantly to this). This percentage was comparable to that achieved by a group of established tests of executive functioning. This suggests that the construct validity of these tests is as good as that of established tests.

However, it is not made clear in either study which (if any) in the criterion group had frontal lobe damage or acknowledged executive functioning difficulties. These results therefore only suggest that the tests are adequately sensitive to brain injury, rather than executive difficulties. It should also be noted that there is a sizeable, positive skew in the population of test results and a low ceiling for some tests. This will compromise the tests' ability to detect mild executive difficulties.

Responses on the DEX were factor analysed to produce three factors; cognitive (involving statements relating to perseveration, distractibility, memory and decision-making), behavioural (involving statements relating to impulsivity and insight) and emotional.

No explicit information on the factor structure of the other tests in the battery is given. However, it is reported that AP, KS and TJ achieved moderate, significant correlations to the behavioural factor and 6E, RS and ZM also achieved similar correlations with the cognitive factor.

Concurrent validity: This was not assessed in the standardisation study. However, the Norris and Tate (2000) study found that all bar one of the BADS tests (TJ) correlated significantly with the results of at least one established test of executive functioning. Most precise was 6E, which correlated moderately, but significantly with only the tests purporting to measure the same abilities. AP correlated with all the established tests and RS correlated significantly with all but the most relevant (the Wisconsin Card Sorting Test).

For the standardisation study, the authors took scores on the DEX as their index of real-life functioning. Correlating test scores against other's ratings on the DEX showed a moderate but significant negative correlation ($r = -0.62$ overall profile scores). So, the more severe the significant other rated the respondent's difficulties, the worse the respondent performed on the tests. The relationship between self-ratings and test scores, on the other hand, was not significant. The authors attribute this to the lack of insight, which is a common characteristic of dysexecutive syndrome. In fact, they urge testers to use the discrepancy between self and other's ratings as a guide to the degree of insight possessed by the respondent.

The BADS overall profile scores, along with the scores from two established tests of executive functioning (Wisconsin Card Sorting Test and Cognitive Estimates Test) and two tests of general ability (WAIS and NART) were also regressed against scores on the DEX. Only BADS scores were found to account for a significant amount of the variance. This is taken by the authors to suggest that the BADS tests have superior ecological validity to the established tests. However, there is a major methodological problem in attempting to validate one new measure (i.e. the BADS tests) by comparing it with another (i.e. the DEX). To be cautious then, these results suggest that the DEX and the BADS are largely measuring the same properties.

Fortunately, support for the ecological validity of the BADS comes from the Norris and Tate study. They found significant positive correlations between scores achieved on the BADS and those achieved on an established behavioural rating scale (the Role Functioning Scale: McPheeters, 1984) for a brain-injured group. AP, ZM and 6E together were able to predict 16.2% of the variance in ratings. This was superior to the established tests, which were unable to predict a significant amount of variance.

Restrictions in usage
The test is accessible to Chartered Psychologists or those 'eligible for Chartered Psychologist status'. Other users, such as occupational therapists, need to attend a Thames Valley Test Company-accredited BADS training workshop to qualify.

EVALUATION

Potential contribution of the instrument to employment assessment of disabled people
Failures of executive functioning are arguably among the most damaging cognitive consequences of brain injury and among the most difficult to ameliorate via remedial training. A test of executive functioning with high ecological validity would therefore be useful for predicting the likely level of independence and higher-level functioning of individuals with brain injury (and possibly schizophrenia, although a discussion of this application is outside the scope of the current article). It could therefore guide rehabilitation programmes, vocational counselling and job and workplace redesign. As such, it could be relevant to psychologists and occupational therapists working in these areas. That it lacks precision and has not been validated against a group of people with known executive difficulties makes it less useful for diagnostic purposes.

SUMMARY AND RECOMMENDATIONS

As claimed by the authors, these tests do appear to present demands that are more naturalistic than those of traditional measures while retaining a format that enables standardised scoring. The tests have also achieved a respectable degree of construct and concurrent validity. The most impressive finding is that, while still modest, the BADS tests have a higher ecological validity than established tests. This is most relevant when assessment questions relate to predictions of outcome for treatment planning or vocational guidance, for example.

However, there are issues that need to be borne in mind if using this battery. First, the tests do not overcome the difficulty inherent in all standardised tests; that of requiring tests to be administered in the 'ideal' environment of the test room, with no distractions and a

high degree of imposed structure. This removes the very factors that often lead to failures in naturalistic situations.

Secondly, and more importantly, no work has yet been done to identify the actual abilities underlying performance on each test. It is therefore difficult to interpret test scores and accurately identify the sorts of real-life tasks that might prove difficult for the respondent.

Given these short-comings, one way in which to usefully incorporate this battery into an assessment programme is to triangulate BADS test results with other measures of executive difficulties. These would be more precise measures of specific executive skills (such as attention and flexible thinking) and naturalistic measures (such as planning and completing a project within a set time limit). It is also viable to use the BADS in a more informal way and administer only the most psychometrically promising tests to provide a quick snapshot of important executive skills such as planning, behavioural control and self-monitoring. 6E, AP and ZM would be the candidates for this approach.

REFERENCES

Norris, G. and Tate, R.L. (2000). The behavioural assessment of the dysexecutive syndrome (BADS): ecological, concurrent and construct validity,. *Neuropsychological Rehabilitation,* 10 (1), 33-45.

McPheeters, H.L. (1984). Statewide mental health outcome evaluation: a perspective of two southern states. *Community Mental Health Journal,* 20, 44-55.