Inter-rater reliability of the Dysexecutive Questionnaire (DEX): comparative data from non-clinician respondents – all raters are not equal

BARKER, Lynne <http://orcid.org/0000-0002-5526-4148>, MORTON, Nicholas, MORRISON, Todd and MCGUIRE, Brian

Published version

BARKER, Lynne, MORTON, Nicholas, MORRISON, Todd and MCGUIRE, Brian (2011). Inter-rater reliability of the Dysexecutive Questionnaire (DEX): comparative data from non-clinician respondents – all raters are not equal. Brain Injury, 25 (10), 997-1004.

Copyright and re-use policy

Inter-rater reliability of the Dysexecutive Questionnaire (DEX):

Comparative data from non-clinician respondents – all raters are not equal

Key Words: executive function, reliability, brain injury.

Abstract

Primary objective: The Dysexecutive Questionnaire (DEX) is used to obtain information about executive and emotional problems after neuropathology. The DEX is self-completed by the patient (DEX-S) and an independent rater such as a family member (DEX-I). This study examined the level of inter-rater agreement between either two or three non-clinician raters on the DEX-I in order to establish the reliability of DEX-I ratings.

Methods and procedures: Family members and/or carers of 60 people with mixed neuropathology completed the DEX-I. For each patient, DEX-I ratings were obtained from either two or three raters who knew the person well prior to brain injury.

Main outcomes and results: We obtained two independent-ratings for 60 patients and three independent-ratings for 36 patients. Intra-class correlations revealed that there was only a modest level of agreement for items, subscale and total DEX scores between raters for their particular family member. Several individual DEX items had low reliability and ratings for the emotion sub-scale had the lowest level of agreement.

Conclusions: Independent DEX ratings completed by two or more non-clinician raters show only moderate correlation. Suggestions are made for improving the reliability of DEX-I ratings.

Introduction

Executive functions are cognitive operations that initiate, inhibit and integrate other functions, simultaneously termed supervisory, attentional or control processes [1-3]. Although novel and non-routinised situations make the greatest demand upon executive functions, executive abilities are also considered to play a crucial role in many aspects of behaviour and cognition, including goal planning and attainment, inhibition of no-longer-relevant responses, strategy initiation and implementation, attentional switching, self-regulation and self-monitoring [4-6]. Executive dysfunction is common after brain injury and can have grave consequences for the individual, diminishing decision making ability and adversely affecting employment prospects and relationship status [7-10]. However, measurement of executive ability is problematic given that executive function tests are not process-pure (i.e., they invariably recruit other non-executive functions that may be variously spared or impaired after neuropathology), and standard clinical neuropsychological tests are thought to lack ecological validity [11].

The Behavioural Assessment of the Dysexecutive Syndrome (BADS) [12] is considered an ecologically valid multidimensional measure of executive function comprising six subtests and a dysexecutive questionnaire (DEX). The DEX is a 20-item, three-factor questionnaire designed to assess everyday changes to cognition, emotion, and behaviour after neuropathology. The DEX is completed by the patient (self-rating: DEX-S) and by a person who knows the patient well (independent rater: DEX-I). A global measure of insight into post-morbid deficits can be obtained by subtracting patient's self-ratings from those of an independent rater, with a high discrepancy value representing low insight into behavioural problems [12, 13-16]. Research suggests that patients often rate themselves as having fewer and less severe problems compared to relative or carer evaluations, reflecting diminished awareness of deficits after neuropathology [17]. Consequently, DEX-I ratings are considered a more reliable post-morbid index of executive dysfunction than self-ratings and DEX-I ratings were used to quantify the presence and severity of executive deficits in head-injured patients in the BADS validation study [12]. The findings showed that DEX-I ratings correlated negatively with patients' performance scores on each BADS subtest, Action Program, Key Search, Modified Six Elements, Rule Shift Cards, Zoo Map and Temporal Judgement, *r*s = -.37, -.31, -.40, -.45, -.46, -.40 respectively) and overall test score *r* = -.62. These results indicate that DEX-I raters accurately identified 'real world' executive deficits in patients, and that 'real-world' problems measured by the DEX corresponded with subsequent performance on BADS executive subtests. Discrepancy scores (DEX-I minus DEX-S) were moderately correlated with three BADS subtests (Modified Six Elements, Zoo Map and Rule Shift Cards, *r*s = -.37, -.38 and -.34, respectively) and total profile score (*r* = .40), indicating that degree of behavioural insight as measured by the DEX discrepancy score, was associated with specific subtest and overall BADS scores. DEX-S ratings showed no relationship with any of the BADS scores, supporting the assumption that others' ratings of patient problems are more reliable than post-injury self-report.

The DEX is a widely used screening instrument designed to evaluate head-injured patients' awareness of their behavioural problems post-injury and is considered an ecologically valid way of informing rehabilitative approaches. Originally conceived as a qualitative instrument, it has become broadly adopted as a quantitative instrument [18]. In addition, recent work investigated the reliability of DEX-S ratings to discriminate mild, moderate or severe injury and found that severity of injury could be classified using quartiles [18]; the authors proposed that a similar analysis should be applied to DEX-I total scores. However, the reliability of DEX-I ratings is difficult to quantify in the absence of DEX inter-rater reliability data; inter-rater correlations were not conducted for DEX-I ratings at test standardisation [12].

Several studies have investigated the psychometric properties of the BADS and the relationship between DEX ratings and performance on executive subtests. Wood and Liossi [19] correlated independent DEX ratings with performance scores on the Hayling and Brixton Test [20] and the Zoo Map and Key Search subtests of the BADS [12] in a group of 59 severely head injured patients. DEX-I ratings correlated with some executive test scores, but not others. Evans and colleagues [21] found that DEX-S ratings correlated with BADS scores in frontally brain-injured patients indicating that DEX-S ratings provide a useful measure of behavioural insight, although there was no relationship between DEX-S and BADS in schizophrenic patients [21]. Similar equivocal findings have also been shown in other studies [22-24]. Considered together these findings might indicate that familial ratings lack reliability or that functions measured by executive tests do not reliably correspond with everyday impairments measured by the DEX, although findings are difficult to interpret in the absence of inter-rater reliability data for the DEX.

Chaytor and Schmitter-Edgecombe [26] compared proxy ratings on the DEX and the Brock Adaptive Functioning Questionnaire (BAFQ)[27] for 46 neurological patients with mixed aetiology and found high concordance, suggesting that both measures share similar global or factorial structure. Another interpretation of these findings is that raters scored consistently across the measures. DEX-I proxy ratings, although widely used, do not provide a direct measure of patients' functional deficits instead they depend upon supposedly accurate observation and categorization of patient deficits by a significant other or clinician. Therefore the validity of the DEX measure is inextricably associated with the efficacy of the rater (rather than the measure) in detecting post-insult behavioural change. This is specifically relevant to studies investigating the DEX factorial structure indexed through (DEX-I) proxy ratings [28], and comparisons of concordance of proxy ratings across measures [26] where DEX-I ratings are assumed to be reliable as standalone ratings or across different measures in the absence of patient scores on objective neuropsychological measures [26, 28]

Several studies have compared the ratings of family members with those of clinicians to investigate reliability of non-clinician raters on the DEX. Bennett, Ong and Ponsford [25] compared the ratings of clinical professionals (a neuropsychologist and occupational therapist) on the DEX-I, to one family member DEX-I rating, patient self-ratings (DEX-S) and patient BADS test scores. Moderate negative correlations were found for clinician ratings and the BADS composite score but this was not found for family member ratings, although family member and clinician total DEX-I ratings were moderately correlated. The authors cautioned against the use of family ratings without consideration of the specific patient population, particularly in acute settings where family members may lack familiarity with patient behavioural change. Other reports have also suggested that clinician ratings may be more accurate than family member ratings [23, 30-33]. However, the typical convention in rehabilitation and out-patient settings is that DEX-I ratings are completed by family members who knew the patient well pre-injury and have current close contact with them, rather than clinicians completing the DEX. Other data show that family member DEX-I ratings reliably capture deficits measured by executive function tests [12-13] indicating that in some circumstances familial ratings are reliable, although it is not known which attributes of the DEX-I rater increase the likelihood that they will generate reliable ratings. There are little data evaluating the inter-rater reliability of non-clinician DEX-I ratings as a standalone metric of cognitive, behavioural, and emotional outcome after neuropathology.

Given the usual clinical practice of relying on familial DEX-I ratings, and the recommendation of other researchers that such ratings are less accurate than clinical ratings, it is important to determine the inter-rater reliability of non-clinician independent raters (i.e. family, friends and carers) on the DEX questionnaire.

In the current study we investigated the reliability of non-clinician DEX-I ratings by item and by subscale (cognitive, behavioural, emotional) across either two or three patient–selected raters for 60 neurological patients with mixed aetiology. DEX-I raters were heterogeneous family members and /or carers in line with other studies where rater category data are provided [14-16,19, 26], and broadly consistent with the independent raters used in validation studies of the BADS [12-13]. We also investigated whether DEX-I raters agreed on severity classifications for patients in line with the suggestion that the measure can be used to reliably quantify injury severity as mild, moderate or severe and based on the recommendation of other authors [18]. This additional analysis provides data on the reliability of DEX-I ratings for assessing severity of deficit in addition to 'type' of functional deficit. Several studies examining psychometric properties of the DEX rely upon DEX-I and in some cases DEX-S ratings, so reliability data should aid interpretation of these findings [18, 26, 28]. Similarly, reliability data should inform evaluation of equivocal findings shown in a number of studies examining the relationship between DEX-I ratings and patient executive function test scores [19, 21-24]. In conclusion, results of the current study should aid selection of reliable measures to accurately identify executive deficits post-brain injury in clinical and academic work.

Method

Participants

Research was conducted in accordance with the declaration of Helsinki and all participants gave informed consent. Sixty participants of mixed neurological aetiology and their relatives or carers were recruited through clinical contact and postal recruitment via a head injury charity.

Executive deficits are often manifested after neurological injury, depending on lesion size, location and severity of injury. The present cohort was expected to provide the degree of heterogeneity recommended for correlational analyses [34]. Of the total patient cohort, 45 were male and 15 were female; 57 had sustained a traumatic brain injury through road traffic accidents, assaults, falls, or sporting accidents and three had sustained other neurological insult (cerebrovascular accident, virus, neurosurgical procedures). Mean age at time of test was 35.6 years (*SD* = 11.3, range: 18-60 years), and mean duration of time since injury was 57.5 months (*SD* = 44.2, range: 10-168 months). The latter information was not available for five cases. We collected data from two independent raters for all 60 participants and from a third independent-rater for 36 of the participants. Family raters were; parent (*n* = 41), spouse (*n* = 26), sibling (*n* = 24), adult child (*n* = 10), other family members (*n* = 12), and we obtained data in some cases from a close friend (*n* = 20) and from paid carers (*n* = 20). Raters were ≥ 18 years of age at the time of the study.

Materials

The DEX is a 20-item questionnaire designed to assess a range of cognitive, behavioural, emotional and motivational problems associated with executive dysfunction in neuropathological groups. Sample statements from the DEX-S are: ‘I act without thinking, doing the first thing that comes to mind,’ with the DEX-I version merely substituting ‘he or she’ in place of the personal pronoun. Items are scored on a 5-point Likert scale (0-4) ranging from 'Never' to 'Very often', with higher scores indicating more problems.

Procedure

Participants were provided with a pack containing instructions, an information sheet and consent form, a demographic questionnaire, one DEX-Self rating scale and three DEX-Independent scales. Participants were instructed to complete the DEX-Self measure and distribute the other measures to close relatives (spouse, parents, siblings, adult children) or friends who knew the participant well both currently and premorbidly, and/or paid carers who had regular current contact with them. Instructions specified that raters should complete the measures independently and confidentially without conferring or comparing ratings.

Results

Table 1 presents descriptive data for mean total DEX scores for DEX-S and DEX-I raters and rater frequency information. Cronbach's alpha coefficients for the DEX-S and DEX-I were excellent, with lower bound estimates for the 95% confidence intervals suggesting satisfactory levels of scale score reliability. Large alpha values also indicate that patients and independent raters scored DEX items similarly (i.e. consistently mild, moderate or severe). Mean inter-item correlations ranged from .58 (DEX-S) to .66 (DEX-I) suggesting moderate to good interrelation between DEX items. Results also show that DEX-I mean total ratings were greater than mean DEX-S ratings, in line with expectations that relatives/significant others rate the patient as having greater and more severe deficits compared to the patients self-rating. DEX-I total ratings were similar in magnitude for independent raters, although note the smaller sample size for the DEX-I ‘rater 3’ group. Results of a one-way ANOVA showed no main effect of group (R1, R2, R3) on ratings, F (2, 153) = .26, p = ns.

[Insert Table 1 about here]

Intraclass correlations (ICC) and their 95% confidence intervals were used to determine inter-rater reliability for DEX-I items assessed by two and three independent raters (ns = 60 and 36, respectively) for each patient (Table 2). Absolute agreement ICC’s were estimated using a one-way random effects model. For interpretative purposes, ICC values > .74 denote excellent agreement; values between .60 and .74 reflect good agreement; values between .40 and .59 represent fair agreement; and values < .40 are poor [35]. Thus, higher ICC values for DEX-I items suggest greater similarity in independent rater scores assessing the same individual for item, subscale and DEX total score.

[Insert Table 2 about here]

ICC values also were computed for total scores on the DEX as well as three DEX subscales: 1) Behaviour (8 items); 2) Cognition (5 items); and 3) Emotion (3 items) (See Table 3).

[Insert Table 3 about here]

Tables 2 and 3 indicate modest levels of inter-rater agreement (i.e., single measure ICC values tended to be < .60). In addition lower-bound estimates for the 95% confidence intervals reveal that poor levels of agreement among raters are possible for individual items, subscales and total scores. We investigated effect of time since injury on inter-rater reliability to establish whether raters with relatives who had long-standing injuries showed greater inter-rater reliability (due to longer exposure to behavioural change in the patient) compared to raters whose relatives had more recent injuries. Time of injury data was not available for 5 cases. Two subgroups were created based on time of injury (group 1 = $<$ 5 years; group 2 = $\geq$ 5 years). ICC values and their 95% confidence intervals were calculated for DEX-I ratings provided by 2 or 3 independent raters and compared to the values reported in Table 2. There was no obvious trend for group 1 (date of injury = $<$ 5 years, n = 36) and for some items the ICC values for group 1 were *lower* than the values obtained for the entire sample, for example item 10, ICC values = .40 for 2 raters and .37 for 3 raters compared to .45 for 2 raters and .44 for 3 raters for the whole sample for item 10. The same analysis was conducted for group 2 (date of injury = $\geq$ 5 years, n = 19). Again, no discernible trend emerged for group 2 (ICC values for item 10 = .43 for 2 raters and .44 for 3 raters) indicating that time since injury did not facilitate inter-rater reliability in the present sample. In addition, the 95% confidence intervals for group 2 were broad (often including zero) that may reflect the small sample size of the group.

Finally, while inter-rater agreement on individual items may be modest, raters' overall classification of participants' degree of impairment, as determined by total DEX scores, may evidence greater levels of convergence. To investigate this possibility, DEX-I total scores for rater one and two were divided into tertiles corresponding (approximately) to mild, moderate, and severe impairment (see Bodenburg & Dopslaff [18] for a similar analysis with DEX-S ratings). Rater group three was omitted from this analysis due to the small number of participants. For raters one and two, congruent classifications (e.g., both raters' total scores fell into the mild impairment category) were assigned a value of 1 and non-congruent classifications (e.g., one rater score fell within the mild category and the other rater score fell within the moderate category) were assigned a value of 0. Results showed that 60% of the total sample (36 rater dyads) agreed on a classification of mild or moderate impairment and 40% (24 rater dyads) disagreed. Among the 24 rater dyads who did not agree on a mild/moderate classification, (4 dyads) differed in excess of 1 category (i.e., mild impairment was designated by one rater and severe impairment was designated by the other rater).

Discussion

The present study investigated the inter-rater reliability and severity classification concordance of non-clinician raters on the DEX-I measure of the Dysexecutive Questionnaire. The typical pattern of lower DEX-Self ratings compared to DEX-Independent ratings was observed in this study, supporting other findings that neurological patients under-report cognitive, emotional and behavioural deficits [8, 14-16]. Our data indicate that whilst non-clinician raters identified a greater degree of impairment across items than self-raters, the agreement between raters on type of deficit exhibited by a particular patient was only weak to moderate, with poor lower-bound confidence intervals. This finding was consistent for individual items, subscale scores and total scores, although there was fair agreement for items 3 and 5 across raters. The DEX 'emotion' subscale yielded the lowest levels of agreement, perhaps reflecting difficulty among raters in identifying problems of this nature, variable exposure to problems of this type, and the nature of the rater’s relationship with the patient. Severity classification was more consistent across raters than rating of type of deficit exhibited by the patient, and only a small subgroup of the total cohort showed a large discrepancy between raters for severity classification.

The findings suggest that use of this measure in clinical settings may be problematic unless non-clinician raters are carefully selected. However the caveat must be made that rater accuracy in identifying problems was not measured in the present study and future work could include patient performance measures of executive and emotional functioning and DEX-I ratings to establish the predictive validity of non-clinician compared to clinician ratings on this scale. The important finding of the present study is that all raters are not equal and thus it would be erroneous for clinicians to assume that DEX-I ratings are always accurate. Future studies might extend our research by comparing a number of carer ratings with an objective measure of cognitive and behavioural functioning to elucidate the characteristics of accurate independent raters.

In the present study, each rater group (R1, R2 and R3) was relatively heterogeneous in frequency of rater (parents predominated for R1, paid carers for R2 and friend for R3), and ANOVA results revealed no significant effect of group on DEX-I ratings. Although it could be argued that paid carers were not familiar with patients' pre-morbid behaviour and so may have provided less accurate ratings than family members, our data do not support this assumption. The duration of time elapsed since injury (< five years or > five years) did not influence the level of agreement between raters. Factors determining accuracy of rater response are likely to include the raters' closeness to the patient, perceptivity and sensitivity to behavioural change, and personal expectations of the functional capacity of a significant other after illness/injury. Future work might investigate the specific attributes that increase the likelihood of reliable DEX-I ratings, for example, familiarity with patient, amount of patient contact, and other psychological and emotional variables compared with patient neuropsychological test scores to better elucidate factors determining reliable DEX-I ratings.

Non-clinical raters showed some convergence for classification of mild, moderate and severe injury on the basis of total DEX scores [18]. Sixty percent of the total cohort agreed on a 'mild' injury severity category, 33% disagreed on the classification of either 'mild' or 'moderate' injury and approximately 7% were highly discrepant. This finding could be seen as encouraging because mild/moderate head injuries share some symptom/feature overlap and there is continued debate among experts about how to differentiate these injury severity categories even at the level of neurological variables [36]. Additionally, the DEX is not typically used for severity classification and it may be unrealistic to assume that either self-raters [18] shown to typically under-report symptom severity [8, 14-16], or non-clinician independent-raters have the precision or skill to reliably categorise functional severity in family members post-neuropathy. Considered overall, only modest classification concordance was found in the present study and caution is advised for use of the measure in this way, although more data are needed to establish the usefulness of total DEX-I scores in approximating deficit severity.

Several studies have examined the factorial structure of the DEX, focussing on DEX-S ratings [18, 36]. Researchers have also compared the association between DEX-I ratings (either by clinicians or family member) to performance on executive tests but provide no data on inter-rater reliability across different categories of clinician or non-clinician raters [25, 30]. Beyond the present study, there are little data on the inter-rater reliability of DEX-I ratings. This represents a lacuna in knowledge about the clinical utility of the DEX and our findings of poor to moderate inter-rater reliability provide a possible explanation for inconsistent associations between DEX-I ratings and some, but not other measures of executive function [19].

An alternative way of using the measure is to calculate a DEX-Discrepancy score by subtracting DEX-S from DEX-I to provide a general metric of behavioural awareness across cognitive, behavioural and emotional factors, a method broadly adopted in the awareness literature [15-16]. Although this method relies upon total DEX-S and DEX-I scores and provides only a coarse-grained measure of functional outcome, it circumvents the reliability problems of utilising either measure alone to quantify functional deficits and for comparison with performance-based measures of executive function. The caveat must be made that the person completing the DEX-I measure should be carefully selected and possibly given training/guidance on measure completion.

To conclude, our findings suggest that clinicians should take care in selecting family members to complete DEX-I ratings, perhaps providing additional instruction or guidance, and should interpret responses with caution. Alternative methods for use of the DEX include calculating an insight score that incorporates both DEX-S and DEX-I ratings, as recommended by the test authors [13], and using DEX-I ratings to quantify post-injury functional severity, although more date are needed to assess the validity of these methods. Alternatively the measure could be used qualitatively as a semi-structured interview administered by a clinician to maximise the clinical utility of ratings. The DEX continues to serve a useful function as a proxy measure of insight by comparing self and other ratings with performance on tests of executive function and emotion.

Future work should replicate our findings with a larger sample and more equal frequencies of raters (our sub-sample of adult children was quite small), collect additional data on DEX-I raters, for example, length of relationship with patient, measure of understanding/knowledge of the neuropsychological effects of brain injury, intelligence, emotional state, and confidence in their own ratings. Finally, other studies might obtain several DEX-I clinician and non-clinician ratings together with patient performance data on neuropsychological tests to further elucidate the reliability and validity of specific rater 'types' in assessing functional outcome using the DEX.

Acknowledgements

The authors gratefully acknowledge the patients and their families who took part in this study.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

1. Miyake A, Friedman NP, Emerson MJ, Witski AH, Howerter A, Wager, TD. The unity and diversity of executive functions and their contribution to complex 'frontal lobe' tasks: A latent variable analysis. Cognitive Psychology 2000; 40: 49-100.

2. Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. Brain 1991; 114: 727-741.

3. Stuss DT, Alexander MP. Is there a dysexecutive syndrome? Philosophical Transactions of the Royal Society B 2007; 362: 901-915.

4. Burgess PW, Shallice T. Response suppression, initiation and strategy use following frontal lobe lesions. Neuropsychologia 1996; 34: 263-272.

5. Robbins TW. Shifting and stopping: Frontostriatal substrates, neurochemical modulation and clinical implications. Philosophical Transactions of the Royal Society B 2007; 362: 917-932.

6. Stuss DT, Levine B. Adult clinical neuropsychology: Lessons from studies of the frontal lobes. Annual Review of Psychology 2002; 53: 401–433.

7. Barker LA, Andrade J, Romanowski CAJ. Impaired implicit cognition with intact executive function after extensive bilateral prefrontal pathology: A case study. Neurocase 2004; 10: 233-248.

8. Barker LA, Andrade J, Morton N, Romanowski CAJ, Bowles D. Investigating the 'latent' deficit hypothesis: Age at time of head injury, implicit and executive functions and behavioral insight. Neuropsychologia 2010; 48: 2550-2563.

9. Eslinger PJ, Flaherty-Craig CV, Benton AL. Developmental outcomes after early prefrontal cortex damage. Brain and Cognition 2004; 55: 84-103.

10. Kentridge R W, Heywood CA. Metacognition and awareness. Consciousness and Cognition 2000; 9: 308-312.

11. Burgess PW, Alderman N, Forbes C, Costello A, Coates LMA, Dawson DR, Anderson ND, Gilbert SJ, Dumontheil I, Channon S. The case for the development and use of "ecologically

valid" measures of executive function in experimental and clinical neuropsychology. Journal of the International Neuropsychological Society 2006; 12: 194-209.

12. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural Assessment of the Dysexecutive Syndrome. London: Thames Valley Test Company 1996.

13. Burgess PW, Alderman N, Evans J, Emslie H, Wilson BA. The ecological validity of tests of executive function. Journal of the International Neuropsychological Society 1998; 4: 547- 558.

14. Barker LA, Andrade J, Romanowski CAJ, Morton N, Wasti A. Implicit cognition is impaired and dissociable in a head-injured group with executive deficits. Neuropsychologia 2006; 44: 1413-1424.

15. Morton N, Barker LA. Top down or bottom up? Executive and implicit function contributions to awareness after Traumatic Brain Injury. Brain Injury 2010; 24: 38.

16. Morton N, Barker, L A. The contribution of executive and implicit processes to awareness of deficits post head injury. Journal of the International Neuropsychological Society 2010, 16:1-10.

17. McBrinn, JM, Wilson, CF, Caldwell S, Carton S, Delargy M, McCann J, Walsh J, McGuire BE. Emotional distress and awareness following acquired brain injury: An exploratory analysis. Brain Injury 2008; 22: 765-772.

18. Bodenburg S, Dopslaff N. The dysexecutive questionnaire advanced: Item and test score

characteristics, 4-factor solution, and severity classification. The Journal of Nervous and Mental Disease 2008; 196: 75-78.

19. Wood RLl, Liossi C. The ecological validity of executive tests in a severely brain injured sample. Archives of Clinical Neuropsychology 2006; 21: 429-437.

20. Burgess PW, Shallice T. The Hayling and Brixton Tests. Bury St Edmunds, England: Thames Valley Test Company Ltd 1997.

21. Evans JJ, Chua SE, McKenna PJ, Wilson BA. Assessment of the dysexecutive syndrome in schizophrenia. Psychological Medicine 1997; 27: 635-646.

22. Stokes N, Bajo A. The relationship between general intelligence, performance on executive functioning tests and everyday executive function difficulties. Brain Injury 2003; 17: 174.

23. Norris G, Tate RL. The behavioural assessment of the dysexecutive syndrome (BADS): Ecological, concurrent and construct validity. Neuropsychological Rehabilitation 2000; 10: 33-45.

24. Wood RLI, Rutterford NA. Relationships between measured cognitive ability and reported psychosocial activity after bilateral frontal lobe injury: An 18-year follow-up. Neuropsychological Rehabilitation 2004; 14: 329–350.

25. Bennett PC, Ong B, Ponsford J. Assessment of executive function following traumatic brain injury: Comparison of the BADS with other clinical neuropsychological measures. Journal of the International Neuropsychological Society 2005; 11: 376-385.

26. Chaytor N, Schmitter-Edgecombe M. Fractionation of the dysexecutive syndrome in a heterogeneous neurological sample: Comparing the Dysexecutive Questionnaire and the Brock Adaptive Functioning Questionnaire. Brain Injury 2007; 21: 615-621.

27. Dywan J, Segalowitz S. Self and family ratings of adaptive behaviour after traumatic brain injury: Psychometric scores and frontally generated ERP's. Journal of Head Trauma Rehabilitation 1996; 11: 79-95.

28. Chan RCK, Bode RK. Analysis of patient and proxy ratings on the Dysexecutive Questionnaire: An application of Rasch analysis. Journal of Neurology, Neurosurgery and Psychiatry 2008; 79: 86-88.

29. Malloy P, Grace J. A review of rating scales for measuring behavior change due to frontal systems damage. Cognitive Behavioral Neurology 2005; 18: 18-27.

30. Chan RCK, Manly T. The application of 'dysexecutive syndrome' measures across cultures. Performance and checklist assessment in neurologically healthy and traumatically brain-injured Hong Kong Chinese volunteers. Journal of the International Neuropsychological Society 2002; 8:

778-780.

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

32. Knight C, Alderman N, Burgess PW. Development of a simplified version of the multiple errands test for use in hospital settings. Neuropsychological Rehabilitation 2002; 12: 231-255.

33. Riccio CA, Hall J, Morgan A, Hynd GW. Executive function and the Wisconsin Card Sorting Test: Relationship with behavioural ratings and cognitive ability. Developmental Neuropsychology 1994; 10: 215-229.

34. Tabachnik BG, Fidell LS. Using multivariate statistics. California: Harper Collins 1996.

35. Meyer GJ, Mihurra JL, Smith BL. The interclinician reliability of Rorschach interpretation in four data sets. Journal of Personality Assessment 2005, 84, 296-314.

36. Laalo JP, Kurki TJ, Sonninen PH, Tenovuo OS. Reliability of diagnosis of traumatic brain injury by computed tomography in the acute phase. Journal of Neurotrauma 2009; 26: 2169-2178.

37. Mooney B, Walmsley C, McFarland K. Factor analysis of the self-report (DEX-S) Questionnaire. Applied Neuropsychology 2006; 13: 12-18.

Table 1.

Questionnaire category	Total DEX score M (SD); α (95% CI)	Rater category: frequency and percentage
DEX-Self (*n* = 60)	33.8 (16.7); .92 (.89-.95)	
DEX-I Rater 1 (*n* =60)	41.3 (17.4); .93 (.91-.96)	Parent (*n* = 23: 38.3%), spouse (*n* = 19: 31.7%), sibling (*n* = 5: 8.3%), friend (*n* = 5: 8.3%), paid carer (*n* = 4: 6.7%), other family (*n* = 2: 3.3%), child (*n* = 2: 3.3%)
DEX-I Rater 2 (*n* = 60)	39.5 (18.9); .95 (.92-.96)	Paid carer (*n* = 16: 27.1%), sibling (*n* = 12: 20.3%), parent (*n* = 11: 18.6%), friend (*n* = 7: 11.9%), spouse (*n* = 5: 8.5%), other family (*n* = 4: 6.8%), child (*n* = 4: 6.8%)
DEX-I Rater 3 (*n* =36)	38.9 (17.7); .94 (.90-.96)	Friend (*n* =8, 23.5%), sibling (*n* = 7, 20.6%), parent (*n* = 7, 20.6%), other family (*n* = 6, 17.6%), child (*n* = 4, 11.8%), spouse (*n* = 2, 5.9%)

Table 2:

DEX-I Item	*2 Raters* (*n* = 60)	*3 Raters* (*n* = 36)
1. Has problems understanding what other people mean unless they keep things simple and straightforward	.36 (.12–.56)	.50 (.30–.68)
2. Acts without thinking, doing the first thing that comes to mind	.41 (.18–.60)	.41 (.20–.61)
3. Sometimes talks about events or details that never actually happened, but s/he believes did happen	.65 (.47–.77)	.63 (.46–.77)
4. Has difficulty thinking ahead or planning for the future	.42 (.19–.61)	.56 (.37–.72)
5. Sometimes gets over-excited about things and can be a bit 'over-the-top' at these times	.60 (.42–.74)	.59 (.40–.74)
6. Gets events mixed up with each other, and gets confused about the correct order of events	.23 (-.02–.46)	.36 (.15–.56)
7. Has difficulty realising the extent of his/her problems and is unrealistic about the future	.32 (.08–.53)	.36 (.15–.56)
8. Seems lethargic, or unenthusiastic about things	.34 (.09–.54)	.44 (.23–.63)
9. Does or says embarrassing things when in the company of others	.51 (.30–.68)	.66 (.49–.79)
10. Really wants to do something one minute, but couldn't care less about it the next	.45 (.23–.63)	.44 (.24–.63)
11. Has difficulty showing emotion	.28 (.03–.50)	.35 (.15–.56)
12. Loses his/her temper at the slightest thing	.44 (.21–.62)	.47 (.27–.66)
13. Seems unconcerned about how s/he should behave in certain situations	.52 (.31–.68)	.56 (.37–.72)
14. Finds it hard to stop repeating saying or doing things once started	.47 (.25–..64)	.48 (.28–.66)
15. Tends to be very restless and 'can't sit still' for any length of time	.24 (-.02–.46)	.39 (.18–.59)
16. Finds it difficult to stop doing something even if s/he knows s/he shouldn't	.52 (.31–.68)	.51 (.32–.69)
17. Will say one thing, but do something different	.36 (.12–.56)	.41 (.21–.61)
18. Finds it difficult to keep his/her mind on something, and is easily distracted	.37 (.13–.56)	.42 (.22–.62)
19. Has trouble making decisions, or deciding what s/he wants to do	.42 (.19–.61)	.48 (.28–.66)
20. Is unaware of, or unconcerned about, how others feel about his/her behaviour	.36 (.12–.56)	.30 (.10–.52)

Table 3.

DEX total scale and subscales	*2 Raters* (*n* = 60)	*3 Raters* (*n* = 36)
DEX-Total	.47 (.24–.64)	.52 (.33–.70)
DEX-Behaviour	.47 (.24–.64)	.52 (.32–.69)
DEX-Cognition	.50 (.29–.67)	.56 (.37–.72)
DEX-Emotion	.35 (.11–.55)	.44 (.24–.64)

Caption for table 1

Table 1: Mean total DEX score for Self- and Independent-raters and rater category frequencies for DEX-I ratings.

Caption for table 2

Table 2: Intraclass Correlations (95% Confidence Intervals) for DEX-Item ratings by rater group.

Caption for table 3.

Table 3: Intraclass Correlations (95% Confidence Intervals) for DEX-I Total and DEX-I Subscale Scores