Neuropsychological Rehabilitation
An International Journal

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/pnrh20

Validity of the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) in children and adolescents with pediatric acquired brain injury

Bendik Romundstad, Stian Solem, Anne Elisabeth Brandt, Ruth Elizabeth Hypher, Kari Risnes, Torstein Baade Rø, Jan Stubberud & Torun Gangaune Finnanger

To cite this article: Bendik Romundstad, Stian Solem, Anne Elisabeth Brandt, Ruth Elizabeth Hypher, Kari Risnes, Torstein Baade Rø, Jan Stubberud & Torun Gangaune Finnanger (2023) Validity of the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) in children and adolescents with pediatric acquired brain injury, Neuropsychological Rehabilitation, 33:4, 551-573, DOI: 10.1080/09602011.2022.2034649

To link to this article: https://doi.org/10.1080/09602011.2022.2034649

Published online: 19 Feb 2022.

NEUROPSYCHOLOGICAL REHABILITATION
2023, VOL. 33, NO. 4, 551–573
https://doi.org/10.1080/09602011.2022.2034649

Validity of the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) in children and adolescents with pediatric acquired brain injury

Bendik Romundstad[a], Stian Solem [a], Anne Elisabeth Brandt[b,c], Ruth Elizabeth Hypher[d], Kari Risnes[b,c], Torstein Baade Rø[b,c], Jan Stubberud[e,f] and Torun Gangaune Finnanger [b]

[a]Department of Psychology, Norwegian University of Science and Technology, Trondheim, Norway; [b]Children's Clinic, St. Olavs's Hospital, Trondheim, Norway; [c]Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, NTNU, Trondheim, Norway; [d]Department of Clinical Neurosciences for Children, Division of Paediatric and Adolescent Medicine, Oslo University Hospital, Oslo, Norway; [e]Department of Psychology, University of Oslo, Oslo, Norway; [f]Department of Research, Lovisenberg Diaconal Hospital, Oslo, Norway

ABSTRACT

The Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) was developed to address the need for a standardized ecologically valid test of executive function (EF) in the pediatric population. Our study aimed to investigate the discriminant, concurrent, and ecological validity of BADS-C in a sample with pediatric acquired brain injury (pABI). Seventy-four participants with pABI aged 10–17 years were included to a pre-registered randomized controlled trial, and baseline assessment was used for the current study. Controls consisted of 60 participants aged 10–17 years. Participants with pABI were assessed with neuropsychological tests and questionnaires of EF, and measurements of general intellectual ability (IQ). Results showed that all BADS-C subtests discriminated between participants with pABI and controls, except for the Playing Cards Test. Concurrent and ecological validity was demonstrated through associations between BADS-C total score, Key Search Test, and Zoo Map Test 1, and neuropsychological tests and teacher questionnaire ratings of EF. Key Search Test and Zoo Map Test 1 predicted teacher ratings of EF, beyond IQ and other neuropsychological test of EF. These findings provide support for BADS-C as a valid clinical assessment tool that can detect everyday executive dysfunction in the pABI population, and guide rehabilitation and treatment decisions.

ARTICLE HISTORY
Received 16 June 2021
Accepted 21 January 2022

KEYWORDS
Executive function; ecological validity; psychometric properties; pediatric acquired brain injury; cognitive outcome

Pediatric acquired brain injury (pABI) concerns childhood insult to the brain and includes traumatic brain injury (TBI), or atraumatic insults such as infections, brain tumours, or stroke. pABI is one of the leading causes of childhood

CONTACT Torun Gangaune Finnanger torun.gangaune.finnanger@stolav.no Children's Clinic, St Olav's Hospital, Trondheim University Hospital, Postbox 3250 Torgarden, 7006 Trondheim, Norway

acquired disability (Thurman, 2016) and mortality (WHO, 2009), affecting cognitive, social and behavioural functioning (Anderson et al., 2005; Ylvisaker & Feeney, 2007). While regenerative mechanisms (e.g., axonal sprouting) happen in the first months after brain injury and improvement may be seen years after the injury/disease onset, impairment usually stabilize (i.e., chronic phase) after the first year post injury (Christensen et al., 2008; Nudo, 2013). Patients with pABI often show impairments in executive function (EF), especially in the more severe cases (e.g., Babikian & Asarnow, 2009). EF is an umbrella term that involves a constellation of high-level cognitive functions that enable goal-related behaviour (Lezak et al., 2004), including skills such as effective problem solving, planning, and regulation of emotion and behaviour. These higher-order functions have a significant impact on children's and adolescents' everyday functioning, as well as cognitive, behavioural, social, and emotional development and adaptation (e.g., Austin et al., 2014; Moriguchi, 2014). Executive dysfunction in patients with pABI has been found to predict negative functional outcomes at a later developmental stage in childhood, adolescence, and into adulthood (e.g., Arnett et al., 2013; Nybo et al., 2004; O'Keeffe et al., 2014). Thus, accurate instruments for the assessment of EF in children and adolescents with pABI are important in order to detect executive dysfunction, and hence guide rehabilitation and treatment.

Due to the complexity of this domain, neuropsychological assessment of EF in the pediatric population is challenging. This purports especially to the ecological validity of EF measures, which typically refers to the degree to which test performance corresponds to real-world performance (Franzen & Wilhelm, 1996). The relationship between performance-based tests of EF and ratings of EF in daily life activities (questionnaires) is inconsistent and unclear, especially post pABI (e.g., Burgess et al., 1998; Chaytor & Schmitter-Edgecombe, 2003; Chevignard et al., 2012; Toplak et al., 2013). This has led to the notion that performance-based EF tests seem to lack ecological validity. It has been argued that the highly structured nature of tests, in addition to being examiner guided, masks potential executive difficulties, thus not representing the relatively unstructured situations of daily life (Chevignard et al., 2012; Jurado & Rosselli, 2007; Shallice & Burgess, 1991). As such, there is a need to develop EF tests with an effort to provide increased representativeness and generalizability to real-life situations (i.e., ecological validity; Burgess et al., 2006). Consequently, clinicians may become better able to detect executive dysfunction interfering with everyday life.

The Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C) was developed to provide a standardized ecologically valid EF test that would predict executive dysfunction present in everyday life for the pediatric population (Emslie et al., 2003). It consists of six developmentally appropriate subtests examining different aspects of EF and a questionnaire (i.e., Dysexecutive Questionnaire for Children; DEX-C). BADS-C aims to assess the so-called dysexecutive syndrome, a constellation of functional deficits related to different aspects of EF (Baddeley, 1986). It includes relatively open-ended tasks assessing flexibility, novel problem solving, impulsivity, planning, and the ability to use feedback in order to moderate behaviour. However, research investigating its validity is sparse, especially in pABI samples (Chevignard et al., 2012; Longaud-Valès et al., 2016).

Studies have demonstrated that BADS-C is able to discriminate between pediatric participants with typical development and various neurological and neurodevelopmental conditions (i.e., discriminant validity). These include TBI (Chevignard et al., 2010), brain tumours (Longaud-Valès et al., 2016; Ward et al., 2009), attention deficit hyper-activity disorder (ADHD; Shimoni et al., 2012; Siu & Zhou, 2014), and autism spectrum disorder (White et al., 2009). In all these studies, the clinical groups performed worse on BADS-C compared to a control group or compared to normative data. However, in the studies comparing participants with pABI with typically developing children and adolescents, only one individual subtest (i.e., the Six Parts Test; Chevignard et al., 2010) or the total score of BADS-C (Ward et al., 2009) was used. While the whole test battery was used in the study by Longaud-Valès et al. (2016), they assessed performance in a limited sample only including a specific type of pABI (i.e., frontal lobe tumour), which makes comparison across studies difficult. Taken together, there is no direct evidence of the discriminant validity of the whole battery of BADS-C in a heterogenous pABI sample.

Concurrent validity (i.e., to what degree a measure is similar to established measures of the same construct) has been demonstrated with strong correlations between several BADS-C subtests and performance-based EF tests (i.e., Trail Making Test from the Delis-Kaplan Executive Functioning System; D-KEFS; and Wisconsin Card Sorting Test; *r*'s between .49 and .76) for children and young adults treated for frontal lobe tumours (Longaud-Valès et al., 2016). However, a small ($n = 21$) and homogenous sample limits the generalizability of the findings. In contrast, weak correlations were found between BADS-C and traditional EF tests (i.e., Controlled Word Association Test, Test of Everyday Attention for Children, and Rivermead Behavioural Memory Test for Children) in the initial standardization of BADS-C (Emslie et al., 2003). The mixed evidence calls for more research, especially with larger and more heterogenous pABI samples. On the other hand, the concurrent validity of the DEX-C has been demonstrated in terms of a strong correlation ($r = .78$) with the Behavioural Rating Inventory of Executive Function (BRIEF) in typically developing children and adolescents (Roy et al., 2015). However, this is yet to be examined in a pABI sample.

While EF and general intellectual ability (IQ) are theoretically different concepts, it has been argued that parts of the constructs (e.g., metacognition of EF, performance IQ, and working memory) are related to each other (Ardila, 2013; Duncan et al., 1995; Wood & Liossi, 2007). Others have found evidence for associations between general EF and full scale IQ (Brydges et al., 2012). In contrast, some have found that other components of EF (i.e., inhibition and shifting) do not relate to intelligence (Friedman et al., 2006). The mixed evidence is also reflected in the link between BADS-C and IQ. In the initial standardization, moderate correlations between BADS-C and IQ were found in a mixed sample of healthy controls ($n = 259$) and children with developmental and neurological problems ($n = 114$), aged eight to 16 years (Emslie et al., 2003). Strong correlations between BADS-C and IQ were found in 16 patients aged eight to 27 years treated for childhood frontal lobe tumours (Longaud-Valès et al., 2016). However, a small sample size and substantial variability in age range and values for both EF and IQ scores, represent methodological limitations to these results. In contrast, another study found weak and mostly non-significant correlations between BADS-C and IQ in 120 neurologically healthy children aged seven to 12 years (Roy et al., 2015). Taken together, these studies may suggest that the correlations between BADS-C and IQ is stronger in clinical samples than non-clinical samples. However, more research with less methodological bias is needed.

Only one study has examined the ecological validity in a pABI sample (Longaud-Valès et al., 2016), and they found strong correlations between BADS-C and BRIEF teacher ratings (*r*'s between .60 and .81). Interestingly, Longaud-Valès et al. (2016) also found weak and non-significant correlations between BADS-C and BRIEF parent ratings. In non-ABI populations, research has shown mostly weak, moderate, or non-significant correlations between BADS-C and questionnaires of EF (Emslie et al., 2003; Roy et al., 2015; Shimoni et al., 2012; Siu & Zhou, 2014). Children with attention-deficit hyperactivity disorder (ADHD) were examined in both the studies of Shimoni et al. (2012; $n = 25$, aged eight to 11 years) and Siu and Zhou (2014; $n = 63$, ages not specified). However, these non-ABI studies only explored associations between parent reports and BADS-C. Thus, the associations between parent-reported EF (questionnaires) and BADS-C appear to be comparable to traditional tests of EF (i.e., not more ecologically valid). In contrast, the findings of Longaud-Valès et al. (2016) indicated ecological validity according to teacher perceptions. However, the small sample size for BRIEF teacher ($n = 11$) and parent ($n = 16$) ratings, and a heterogenous age range (children and young adults), limit the generalizability of the findings. Thus, these results need to be replicated in a larger sample, including other pABI conditions than frontal lobe tumours.

The aim of this study was to examine the discriminant, concurrent, and ecological validity of the BADS-C for children and adolescents with pABI. Indeed, the present study is the first to examine these aspects of validity in a relatively large and heterogenous pABI sample. Pursuing these aims may identify whether BADS-C can reveal executive dysfunction known to interfere with everyday life in children and adolescents with pABI, and hence guide rehabilitation and treatment at an earlier stage. The research hypotheses were as follows:

(1) The pABI group will perform worse on BADS-C compared to a group of typically developed children and adolescents.
(2) Performance on BADS-C will be associated with performance on neuropsychological EF tests in pABI participants. In addition, performance on BADS-C will be associated with IQ.
(3) DEX-C parent ratings will be associated with BRIEF parent ratings, and DEX-C teacher with BRIEF teacher ratings, both in pABI participants.
(4) BADS-C scores will be associated with parent, teacher, and self-report questionnaire measures of EF (DEX-C and BRIEF) in pABI participants.
(5) BADS-C will explain additional variance in EF questionnaires beyond IQ and performance-based tests of EF in pABI participants.

Materials and methods

Participants

The present study is a cross-sectional design, based on baseline data derived from a dual-center randomized controlled trial (see Hypher et al., 2019, for more details). Seventy-four children and adolescents in the chronic phase of pABI resulting from traumatic (TBI) and non-traumatic brain injuries (brain tumour, stroke, hypoxia/anoxia and brain infections/inflammations) were included. Inclusion required that they were aged between 10 and 17 years at time of inclusion; at least 12 months since injury/illness/finished cancer treatment; and parent-reported EF deficits in daily life determined by a semi-structured interview. Exclusion criteria included: brain injury acquired before 2 years of age; cognitive, physical, sensory or language impairment affecting the capacity to attend regular learning situations in the classroom setting; neurological disease pre-injury; recently detected relapse in brain tumour; unfit for outcome evaluation (evaluated independently by two investigators); and not fluent in Norwegian.

In addition, 60 children and adolescents aged 10–17 years constituted the control group. Exclusion criteria were: previous head injury with loss of consciousness surpassing five minutes, stroke, other brain injuries or brain diseases; severe (neuro) psychiatric disorder or substance abuse; cognitive, sensory, physical, or language impairment affecting the capacity to attend mainstream school; not fluent in Norwegian.

Materials

Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C)

BADS-C assesses a broad range of EFs in the pediatric population, with tests developed to reflect EF in everyday life (Emslie et al., 2003). It was developed as a more child friendly version of the adult Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al., 1997). Different aspects of EF are examined by six performance-based subtests (Playing Cards Test; Water Test; Zoo Map Tests 1 and 2; Key Search Test; and the Six Parts Test), and a questionnaire (DEX-C). Raw scores on BADS-C subtests are converted to scaled scores ($M = 10$, $SD = 3$), and these are used to obtain an overall standardized score for the BADS-C total score ($M = 100$, $SD = 15$). Low scores on BADS-C and high scores on DEX-C indicate executive difficulties. The norms were initially provided for British children and adolescents aged eight to 15 years and 11 months, based on a representative sample of 259 children balanced for sex, mean estimated IQ, and socio-economic background (Emslie et al., 2003). Later, the norms were extended to also include children from seven years of age (Emslie et al., 2013). The scaled scores can be adjusted for eight age groups and three IQ groups. In this study, scores were only adjusted for age. As there is mixed evidence in terms of the link between BADS-C/EF and IQ (e.g., Longaud-Valès et al., 2016), we chose not to adjust scores according to IQ. BADS-C was administered to both the pABI group and the control group.

Playing Cards test assesses mental flexibility in terms of the ability to change an established pattern of responding to cards being presented, including two conditions with different rules (Emslie et al., 2003). Scoring is based on the number of uncorrected errors in the second condition. Water Test is a novel, practical task that assesses the ability to develop a plan of action to solve a problem, by manipulating different tools in order to extract a cork from a tube. There are five correct steps with two points awarded for each step (maximum 10 points), taking into account time and perseverations. Key Search Test examines the ability to organize an efficient, systematic, implementable plan of action, with a maximum score of 14 based on a set of criteria (e.g., efficient search, type of search, understanding task requirements). Zoo Map Test 1 is an open-ended test with little structure, examining the ability to plan an adequate route, with a maximum score of eight based on correct sequences weighed by errors score. Zoo Map Test 2 is a more structured version of the same task. Six Parts Test examines planning, task scheduling and performance monitoring. It includes three tasks of different colour (each with two parts), and the examinee is challenged to attempt something from each part over a five minute period, without doing two parts of the same colour after one another. A maximum score out of 16 is based on the number of parts attempted weighed by number of broken rules, in addition to qualitative strategies.

DEX-C is a 20-item questionnaire where the child's or adolescent's behaviour is rated by parents or teachers (Emslie et al., 2003). It is as such regarded as a measure of everyday EF at home or school. The items cover four broad areas of potential executive dysfunction: emotional/personality, motivational, behavioural, and cognitive. Each question is sensitive to a specific characteristic of the dysexecutive syndrome. The items are scored on a Likert scale ranging from 1 ("never") to 5 ("very often"). Examples of items are: "Has difficulty thinking ahead or planning when undertaking a task or activity"; "Acts without thinking." Raw total scores were used for analyses in this study. For the present study, DEX-C was rated by a parent and main teacher in the pABI group. For the control group, it was rated by one of the parents.

Behavior Rating Inventory of Executive Function (BRIEF)

The BRIEF is a standardized self- and informant questionnaire assessing EF in children and adolescents aged five to 18 years (11–18 years for self-report) in home and school environments (Gioia et al., 2000). It is widely used and has shown good psychometric properties (Roth et al., 2014). For this study, it was used as a criterion measure for ecological validity of BADS-C. The BRIEF Global Executive Composite (GEC) is an overarching summary score based on different executive component scales (inhibition, shifting, emotional control, initiation, working memory, planning, organization of materials, and monitoring). A total of 86 statements describes different behaviours which are rated by their frequency on a Likert scale ranging from 1 ("never") to 3 ("often"). The total scores are converted into standardized *T*-scores ($M = 50$, $SD = 10$). High scores indicate more EF difficulties in daily life. In the present study, BRIEF GEC was rated by parents, teachers, and as self-report for the pABI group. It was not administered to the control group.

Delis-Kaplan Executive Functioning System (D-KEFS)

The D-KEFS is a neuropsychological test battery, which includes nine individually administered tests, covering a range of EFs (Delis et al., 2001). It was used in this study as a criterion measure for concurrent validity of the BADS-C. Norms are provided for ages eight to 89 years of age. Raw scores are converted into scaled scores ($M = 10$, $SD = 3$). Higher scores indicate better performance. The D-KEFS is one of the most widely used performance-based test batteries of EF, and it has demonstrated good psychometric properties (Stephens, 2014). It is administered in a traditional pen-and-pencil manner, and it is therefore not explicitly made with an effort to provide ecological validity. In the present study, two of the subtests were administered to the pABI group and used for analyses: Color-Word Interference Test (conditions 3 and 4) and Trail Making Test (condition 4). The scaled scores were based on numbers of seconds used on each task.

Wechsler Intelligence Scale for Children – fifth edition (WISC-V)

The WISC-V (Wechsler, 2014) was used in this study for analyses of concurrent validity of the BADS-C. Subtests were administered to the pABI participants in order to produce Verbal Comprehension Index (Similarities and Vocabulary), Working Memory Index (Digit Span and Picture Span), and Processing Speed Index (Coding and Symbol Search). Also, the subtests Block Design and Matrix reasoning were administered, and an estimated Full Scale IQ were produced. For the control group, Verbal Comprehension Index was the only index administered, in addition to the Block Design and Matrix Reasoning subtests.

Procedure

The participants with pABI were recruited to a clinical trial (Hypher et al., 2019) from three university hospitals, based on hospital discharge diagnosis and record information. Invitation letters were sent to potential participants (*N* = 223), specifically soliciting patients experiencing EF dysfunction in daily life. Written informed consent was required from participants (>16 years) or primary caregivers (participants <16 years), followed by a semi-structured interview to assess eligibility. Of the 223 children and adolescents invited, 124 were excluded during the first telephone contact. Of these, 56 declined to participate, 46 did not meet inclusion criteria (i.e., 36 had insufficient EF complaints, seven were based on function, three for other reasons), or were excluded for other reasons (i.e., 15 for no contact, and seven were wrongly invited).

Of the 99 participants eligible for screening, 10 did not meet inclusion criteria and were excluded, and two individuals declined to participate. Eighty-seven participants were eligible for assessment. Pre-assessment attrition was 11; nine withdrew due to worsening of illness, initiation of medication or intensification of rehabilitation, and two were excluded due to violation of inclusion criteria not previously communicated. Of the 76 participants completing baseline assessment, data was missing for two participants, making the final sample of *N* = 74 for the present study (see also Brandt et al., 2021, for more details). Data were collected between 2018 and 2019 at two sites in Norway, namely St. Olavs Hospital in Trondheim and Oslo University Hospital.

The control group was recruited as a convenience sample invited from local public schools in Trondheim and Oslo. Written informed consent was required by participants (>16 years) or primary caregivers (participants <16 years). Schools with families from various backgrounds were targeted. An equal number of boys and girls were recruited, and with an age span evenly distributed across three age groups (i.e., 10–11 years, 12–13 years, and 14–17 years).

Participants in the control group were assessed during a 2–3-hour period, at daytime or evening, in the hospital or at school. These factors increased the variation in testing environment within the control group. However, the participants were assessed by the same test technician at each of the two sites. Participants in the pABI group were assessed at the hospital during one workday (7–8 h), and completed additional tests compared to the control group. Testing was limited to one day to prevent unnecessary use of the participants' time and unnecessary burden of additional travel for assessment, as we recruited from the whole country with long travel distances. To alleviate tiredness, participants were given frequent breaks and a one-hour lunch break. The participants with pABI were assessed at two sites within the hospital and by different test technicians. Nevertheless, a Standard Operating Procedure was made as a guide for the test technicians to minimize variation in the testing situation.

Statistical analyses

All variables used for statistical analyses were checked for normality by examining the ratio between skewness/kurtosis and standard error. Water Test, Zoo Map Test 2, DEX-C teacher, and D-KEFS Color-Word Interference condition 3, were transformed into variables with more acceptable normality distributions. *T*-tests compared results between the pABI group and the control group with respect to BADS-C, DEX-C parent ratings, and WISC-V. Effect sizes were calculated as Cohen's *d,* based on the pooled standard deviation between the pABI group and the control group. The strength of the effect sizes was interpreted based on Cohen's (1988) conventions, namely small (*d* between 0.2 and 0.5), medium (*d* between 0.5 and 0.8), and large ($d > 0.8$). Analysis of variance (ANOVA) tests compared BADS-C performance between pABI subgroups and the control group. Water test, Zoo Map Test 2, and DEX-C parent were reported with Welch F values because of lack of homogeneity of variance. Games-Howell post hoc test was used for these variables, and Bonferroni post hoc test was used for the other significant variables.

The relationships between BADS-C, DEX-C, BRIEF, WISC-V, and D-KEFS, were examined using Pearson correlations. Correlations between .10 and .30 were considered weak, .30–.50 moderate, and .50–1.0 strong (Cohen, 1988). Hierarchical multiple regression was used to investigate the predictive value of BADS-C on questionnaire ratings of EF, beyond established neuropsychological measures. BRIEF GEC and DEX-C were employed as dependent variables. FSIQ and the three D-KEFS measures were entered on step 1 and step 2, and BADS-C measures were entered on step 3. Only the BADS-C measures with significant correlations with the dependent variables were used. The variance inflation factor (VIF) and the tolerance statistic were examined, to assess multi-collinearity. Also, the Durbin-Watson test examined autocorrelation between residuals.

Missing data were not imputed or replaced. Thus, the specific sample sizes for the different measures with missing values were as follows for the pABI group: BADS-C total score and Six Parts Test ($n = 73$), DEX-C teacher ($n = 69$), BRIEF parent ($n = 73$) and teacher ($n = 69$), WISC-V Block Design and Matrix Reasoning ($n = 73$), Verbal Comprehension Index ($n = 72$), Working Memory Index ($n = 71$), and Processing Speed Index and Full Scale IQ ($n = 70$). For the control group, the sample sizes were: BADS-C ($n = 59$), DEX-C parent ($n = 58$), and WISC-V Verbal Comprehension Index ($n = 59$).

Results

Preliminary analyses

Correlations between the six BADS-C subtests were mostly weak, both for the control group (r's ranging from −.28 to .38) and the pABI group (r's ranging from −.06 to .30). For the control group, there were, however, significant correlations between Playing Cards Test and Six Parts Test ($r = .38$, $p = .003$), and between Water Test and Zoo Map Test 2 ($r = -.28$, $p = .03$). All other correlations were non-significant. For the pABI group, only the correlations between Key Search Test and Zoo Map Test 1 were significant ($r = .30$, $p = .01$). There were no significant sex differences in BADS-C scores. DEX-C items were found to have a strong internal consistency, both for the parent ratings ($\alpha = .94$) and the teacher ratings ($\alpha = .97$). Non-significant data not shown.

Demographics, test scores, and group comparisons

The groups did not differ significantly in terms of age at testing or in sex distribution (Table 1). However, the parents of the control group had significantly higher education compared to the parents in the pABI group. The control group had significantly higher scores compared to the pABI group on the IQ measures Verbal Comprehension, Block Design, and Matrix Reasoning. Scores were slightly above the normative average for the control group on these tests, and slightly below the normative average for the pABI group. Still, both

Table 1. Demographic variables for the pABI group and the control group.

Measure	*n (%)*		*p*
	pABI ($n = 74$)	Control group ($n = 60$)	
Sex			.435
Female	42 (56.8)	30 (50)	
Male	32 (43.2)	30 (50)	
Maternal education			<.001
Completed High School or less	27 (38.1)	2 (3.3)	
One or more years of university	44 (61.9)	58 (96.7)	
Father's education			<.001
Completed High School or less	33 (51.6)	10 (17.2)	
One or more years of university	31 (48.4)	48 (82.8)	
pABI aetiologies			
TBI	18 (24.3)		
Tumour	27 (36.5)		
Stroke	17 (22.9)		
Brain infection	7 (9.5)		
Anoxia	5 (6.8)		
	M (SD)		
Age	13.43 (2.31)	12.75 (1.94)	.065
WISC-V			
Verbal Comprehension Index	96.36 (11.95)	108.73 (12.11)	<.001
Block Design	8.55 (3.06)	10.90 (2.34)	<.001
Matrix Reasoning	9.42 (2.90)	11.42 (2.82)	<.001

Note: pABI = pediatric acquired brain injury; TBI = traumatic brain injury; WISC-V = Wechsler Intelligence Scale for Children – fifth edition.

groups scored within what can be considered the normal range relative to the standardization samples (i.e., within 1.5 *SD* from the normative mean).

Table 2 presents test and questionnaire scores for the pABI group, except BADS-C and DEX-C parent scores, which are presented separately in the next section. All the D-KEFS and WISC-V scores were below the normative average, and BRIEF GEC teacher and parent scores were higher than normative average. Although in the normal range, some of the scores approached 1 *SD* below normative average. BRIEF GEC self-report scores were significantly lower than both parent ($p = .001$; $d = .44$) and teacher ($p = .009$; $d = .40$) ratings.

The control group performed significantly better than the pABI group on all of the measures, except for the Playing Cards Test (Table 3). The effect sizes ranged from medium to large.

Correlations between BADS-C, IQ, and EF tests and questionnaires

BADS-C total score, Key Search Test, and Zoo Map Test 1 had significant negative correlations with both DEX-C and BRIEF teacher ratings for the pABI group (Table 4). Correlations were moderate (r's ranging from −.37 to −.46). The BADS-C tests were mostly weakly and non-significantly correlated with DEX parent, and BRIEF parent and self-report ratings. There were significant strong correlations between DEX-C and BRIEF parent ratings, and DEX-C and BRIEF teacher ratings.

There were significantly moderate to strong positive correlations between BADS-C total score, Key Search Test, and Zoo Map Test 1, and all of the D-KEFS and WISC-V scores in the pABI group (Table 5). Some of the DEX-C parent and teacher ratings were significantly correlated with some of the D-KEFS and WISC-V variables, but most were not significant, and the correlations were weak.

Table 2. Test and questionnaire scores for the pABI group ($n = 74$).

Measure	*M (SD)*
WISC-V	
Full Scale IQ	92.60 (13.46)
Working Memory Index	94.04 (14.11)
Processing speed Index	89.57 (17.62)
D-KEFS	
CWIT 3	7.84 (3.23)
CWIT 4	7.69 (3.61)
TMT 4	8.01 (3.72)
BRIEF GEC	
Parent	59.53 (10.59)
Teacher	59.75 (14.04)
Self	54.36 (12.84)
DEX-C Teacher	17.75 (13.46)

Note: pABI = pediatric acquired brain injury; BRIEF = Behavioral Rating Inventory of Executive Function; GEC = Global Executive Composite; WISC-V = Wechsler Intelligence Scale for Children – fifth edition; D-KEFS = Delis-Kaplan Executive Functioning System; CWIT 3 = Color-Word Interference Test condition 3; CWIT 4 = Color-Word Interference Test condition 4; TMT 4 = Trail Making Test condition 4; DEX-C = Dysexecutive Questionnaire for Children.

Table 3. Comparisons between the pABI group and the control group on BADS-C and DEX-C.

Measure	*M (SD)*		*p*	d_{pooled}
	pABI	Control group		
BADS-C Total	83.89 (19.94)	102.08 (16.33)	<.001	1.00
Playing Cards	8.15 (3.68)	9.24 (3.39)	.082	0.31
Water	10.01 (3.18)	11.42 (2.39)	.037	0.50
Key Search	10.47 (4.28)	12.27 (3.71)	.012	0.45
Zoo Map 1	8.31 (4.11)	9.85 (3.55)	.025	0.40
Zoo Map 2	8.68 (3.5)	10.25 (2.40)	.003	0.52
Six Parts	7.14 (3.08)	9.20 (2.63)	<.001	0.72
DEX-C Parent	26.53 (12.37)	12.66 (8.99)	<.001	1.28

Note: $n_{pABI} = 73–74$. $n_{Control\ Group} = 58–59$. pABI = pediatric acquired brain injury; BADS-C = Behavioural Assessment of the Dysexecutive Syndrome for Children; DEX-C = Dysexecutive Questionnaire for Children.

Predictive capacity of BADS-C on questionnaire ratings of EF

Two regression models were tested to assess the predictive value of BADS-C on questionnaire measures of EF (DEX-C and BRIEF GEC). In the first regression model, Zoo Map Test 1 and Key Search Test were used as predictor variables in step 3. The second model was analogous to the first, with the exception of having BADS-C total score in step 3 instead of Zoo Map Test 1 and Key Search Test. In all of the regression analyses performed, multicollinearity was not an issue. VIF values ranged between 1 and 2.52, and no tolerance level was below 0.2. Also, the Durbin-Watson values were all close to 2, ranging between 1.78 and 2.33. This indicated no issue with autocorrelation in the residuals.

As shown in Table 6, Zoo Map Test 1 and Key Search Test explained significant amounts of the variance of both DEX-C (23%) and BRIEF GEC (12%) teacher ratings, beyond FSIQ and D-KEFS. Both Zoo Map Test 1 and Key Search Test were significant predictors of DEX-C teacher ratings. Only Zoo Map Test 1 significantly predicted BRIEF GEC teacher ratings. BADS-C total score did not predict DEX-C or BRIEF GEC teacher ratings, beyond FSIQ and

Table 4. Correlations between BADS-C, DEX-C, and BRIEF for the pABI group.

Measure	DEX-C		BRIEF GEC		
	Parent	Teacher	Parent	Teacher	Self
BADS-C Total	−.17	−.38***	−.24*	−.42***	−.10
Playing Cards	−.01	.06	.04	.08	.05
Water	−.12	−.25*	−.19	−.07	−.09
Key Search	−.14	−.38***	−.29*	−.37**	−.09
Zoo Map 1	−.14	−.46***	−.09	−.44***	−.10
Zoo Map 2	−.15	.05	.01	−.14	.14
Six Parts	−.06	−.04	−.08	−.19	−.13
DEX-C					
Parent	–	.17	.75***	.13	.39***
Teacher	.17	–	.35**	.77***	.13

Note: $n_{BADS.C} = 73–74$. $n_{DEX\text{-}C} = 74$ (Parent); 69 (Teacher). $n_{BRIEF} = 73$ (Parent); 69 (Teacher); 74 (Self). pABI = pediatric acquired brain injury; BADS-C = Behavioural Assessment of the Dysexecutive Syndrome for Children; DEX-C = Dysexecutive Questionnaire for Children; BRIEF = Behavioral Rating Inventory of Executive Function; GEC = Global Executive Composite. *p <.05. **p <.01. ***p <.001.

Table 5. Correlations between BADS-C, DEX-C, D-KEFS, and WISC-V for the pABI group.

Measure	D-KEFS			WISC-V Full scale IQ and Index				WISC-V subtests	
	CWIT 3	CWIT 4	TMT 4	FSIQ	WMI	PSI	VCI	BD	MR
BADS-C Total	.45***	.54***	.56***	.60***	.45***	.51***	.31**	.45***	.55***
Playing Cards	.16	.21	.31**	.21	.03	.17	.07	.26*	.25*
Water	.19	.27*	.29*	.15	.03	.18	.04	.11	.30*
Key Search	.31**	.41***	.36**	.57***	.45***	.37**	.26*	.34**	.58***
Zoo Map 1	.36**	.32**	.28*	.43***	.35**	.27*	.29*	.37**	.31**
Zoo Map 2	.10	.10	.01	.07	.08	.03	.12	.05	.06
Six Parts	.15	.30*	.28*	.20	.14	.36**	.14	.15	.10
DEX-C									
Parent	−.16	−.18	−.19	−.21	.01	−.03	−.25*	−.11	−.18
Teacher	−.20	−.25*	−.22	−.31*	−.26*	−.22	−.02	−.23	−.37**

Note: $n_{\text{BADS-C}} = 73–74$. $n_{\text{DEX-C}} = 74$ (Parent); 69 (Teacher). $n_{\text{D-KEFS}} = 74$. $n_{\text{WISC-V}} = 70$ (FSIQ and PSI); 71 (WMI); 72 (VCI); 73 (BD and MR). pABI = pediatric acquired brain injury; BADS-C = Behavioural Assessment of the Dysexecutive Syndrome for Children; DEX-C = Dysexecutive Questionnaire for Children; D-KEFS = Delis-Kaplan Executive Functioning System; CWIT 3 = Color-Word Interference Test condition 3; CWIT 4 = Color-Word Interference Test condition 4; TMT 4 = Trail Making Test condition 4; WISC-V = Wechsler Intelligence Scale for Children – fifth edition; FSIQ = Full Scale Intelligence Quotient; WMI = Working Memory Index; PSI = Processing Speed Index; VCI = Verbal Comprehension Index; BD = Block Design; MR = Matrix Reasoning. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 6. Regression analyses predicting DEX-C and BRIEF GEC teacher scores for the pABI group.

	DEX-C Teacher				
Step	R^2	R^2_{Adj}	R^2 *cha*		
1. FSIQ	.09	.08	.09		
2. D-KEFS	.10	.04	.01		
3. Zoo Map 1 & Key Search	.33	.26	.23***		
Final step	*B (SE)*	*LL*	*UL*	*β*	*t*
FSIQ	0.16 (0.03)	−0.034	0.067	.12	0.65
CWIT3	−0.39 (0.47)	−1.330	0.547	−.13	−0.83
CWIT4	−0.06 (0.08)	−0.227	0.107	−.12	−0.72
TMT4	−0.63 (0.07)	−0.210	0.083	−.13	−0.87
Zoo Map 1	−0.20 (0.05)	−0.308	−0.092	−.45	−3.72***
Key Search	−0.12 (0.05)	−0.223	−0.008	−.28	−2.15*
	BRIEF GEC Teacher				
Step	R^2	R^2_{Adj}	R^2 *cha*		
1. FSIQ	.15	.13	.15**		
2. D-KEFS	.21	.16	.06		
3. Zoo Map 1 & Key Search	.32	.25	.12*		
Final step	*B (SE)*	*LL*	*UL*	*β*	*t*
FSIQ	0.08 (0.20)	−0.323	0.489	.08	0.41
CWIT3	−0.75 (3.82)	−8.405	6.901	−.03	−0.20
CWIT4	−1.28 (0.68)	−2.643	0.094	−.31	−1.87
TMT4	−0.09 (0.59)	−1.270	1.082	−.02	−0.16
Zoo Map 1	−1.22 (0.44)	−2.106	−0.331	−.34	−2.75**
Key Search	−0.62 (0.44)	−1.501	0.264	−.19	−1.40

Note: pABI = pediatric acquired brain injury; BRIEF = Behavioral Rating Inventory of Executive Function; GEC = Global Executive Composite; FSIQ = Full Scale Intelligence Quotient; D-KEFS = Delis-Kaplan Executive Functioning System; CWIT 3 = Color Word Interference Test condition 3; CWIT 4 = Color Word Interference Test condition 4; TMT 4 = Trail Making Test condition 4; BADS-C = Behavioural Assessment of the Dysexecutive Syndrome for Children. *p <.05. **p <.01. ***p <.001.

D-KEFS (not included in table). Analogous regression analyses were also performed with BRIEF GEC parent ratings as dependent variable, with none of the BADS-C tests predicting this questionnaire measure of EF.

Discussion

The present study is the first to examine the validity of the BADS-C in a relatively large and heterogenous pABI sample. Discriminant validity was demonstrated for all of the subtests, except for the Playing Cards Test. Also, the study demonstrated concurrent validity between some of the BADS-C measures (i.e., total score, Key Search Test, and Zoo Map Test 1), EF tests and IQ. Ecological validity was indicated between the same BADS-C measures and teacher questionnaires assessing EF, but not with parent- or self-report. Furthermore, ecological validity was demonstrated, by showing that the subtests Key Search Test and Zoo Map Test 1 predicted teacher questionnaires of EF, beyond IQ and other EF tests. Overall, our hypotheses were partly supported as the validity of BADS-C varied according to different subtests, being stronger for the Key Search Test and Zoo Map Test 1 compared to the others.

Discriminant validity

The weaker performance on the BADS-C for the pABI group compared to the control group, suggests that the BADS-C is sensitive to detecting executive dysfunction in children and adolescents with pABI. These results are in line with previous research on discriminant validity of BADS-C for several clinical groups (e.g., Chevignard et al., 2010; Shimoni et al., 2012; Siu & Zhou, 2014; Ward et al., 2009; White et al., 2009). However, we did not find a significant difference in performance between the pABI sample and the control group on the Playing Cards Test, although the pABI group performed worse. The control group scored below normative mean, which may have contributed to why the difference did not meet the level of significance. This may also indicate a reduced representativeness of our control group, or cultural differences between Norwegian and British children (on which the norms are based).

A non-significant difference for the Playing Cards Test has also been observed in a study by Shimoni et al. (2012), comparing children with ADHD and healthy controls. This result was attributed to ceiling effects, which does not apply to the results of our study. In contrast, another study found the Playing Card Test to be the subtest from BADS-C that best differentiated between children with ADHD and healthy controls (Siu & Zhou, 2014). However, as both studies included participants with ADHD, comparison to the findings in our study with a pABI sample is limited. Thus, while the discriminant validity of the BADS-C test battery as a whole seems to be strong, it is more unclear with respect to the subtest Playing Cards Test. Whether this can be explained by properties of the test (e.g., inclination for ceiling effect), sample characteristics, or methodological limitations (e.g., biased control group), merit further research.

Concurrent validity

As expected, concurrent validity was indicated for BADS-C total score, Key Search Test, and Zoo Map Test 1. Nevertheless, this did not apply for the other BADS-C subtests. The results of the present study are in accordance with the results of Longaud-Valès et al. (2016), who found strong significant correlations between most of the BADS-C measures, IQ, and D-KEFS Trail Making Test condition 4. In our study, Trail Making Test condition 4 was the only measure significantly correlated with all of the BADS-C tests (except Zoo Map Test 2). The high similarity between these tests may indicate that BADS-C is particularly suitable at capturing non-verbal mental flexibility (a central component of EF; Diamond, 2013; Miyake et al., 2000), as this is the predominant EF process examined by the Trail Making Test (Delis et al., 2001).

WISC-V does not primarily measure EF, but rather general intellectual ability. Nevertheless, the significant correlations between BADS-C and the Working Memory Index of WISC-V suggest concurrent validity, as working memory is considered one of the core components of EF (Diamond, 2013; Miyake & Friedman, 2012). Strong correlations between EF and working memory has also been shown in previous research (Friedman et al., 2006; McCabe et al., 2010). However, only some measures of BADS-C showed such a relationship in our study (i.e., BADS-C total score, Zoo Map Test 1, and Key Search Test). Although conjectural, it may be that these tests challenge working memory to a larger extent than the others. For instance, the Zoo Map Test 1 demands monitoring of current, previous, and future locations of a route through the zoo. However, one would also expect a high demand of working memory for some of the other subtests (e.g., the Six Parts Test, which requires a high degree of task scheduling and performance monitoring), of which the correlations were surprisingly weak. We may speculate that the high demands of the Six Parts test on other cognitive abilities (e.g., multi-tasking, control abilities) may explain the weak correlation with working memory.

The link between EF and IQ is unclear, which is reflected in a discussion of whether the BADS-C scaled scores should be calculated according to IQ or not (Roy et al., 2015), as is the case in the current manual (Emslie et al., 2003). The decision to scale BADS-C scores according to IQ bands was based on moderate correlations (not specified in the manual) between four of the sub-tests and IQ estimated by The Basic Reading Test of the Weschler Objective Reading Dimensions Test (WORD), using a mixed clinical and healthy sample (Emslie et al., 2003). Although strong correlations between BADS-C and IQ were found in another study on patients treated for childhood frontal lobe tumours (Longaud-Valès et al., 2016), the results are limited by factors such as low sample size ($n = 16$), and highly variable age range and values for both EF and IQ scores. In contrast, our study found significant low to moderate correlations only between Key Search Test and Zoo Map Test 1 and IQ measured by WISC-V, while the remaining subtests showed mostly non-significant and weak correlations. Almost the exact same pattern was found by Roy et al. (2015) using a healthy sample, arguing that BADS-C scores should not be adjusted according to IQ. Thus, evidence using different clinical and healthy samples show differential links between BADS-C subtests and IQ. Also, the decision to estimate BADS-C according to IQ in the manual was based on a rather limited IQ measure. Hence, we support the conclusions from Roy et al. (2015), and will as such recommend clinicians to choose not to adjust BADS-C scores according to IQ.

Our findings further showed strong concurrent validity for the DEX-C, both for parent and teacher ratings, which is in accordance with findings from Roy et al. (2015). However, the correlations between parent and teacher ratings were only low to moderate. Moderate inter-rater agreement between parents and teachers in BRIEF has also been shown in other studies, which has been attributed to differences in expectations between the home and school settings (Gioia et al., 2000).

Ecological validity

We found support for ecological validity for parts of BADS-C (i.e., total score, Zoo Map Test 1, and Key Search Test). However, this only applied to teacher perceptions of the participants' EF. A similar pattern with strong teacher and weak parent correlations between BADS-C and EF questionnaires was also demonstrated in the study of Longaud-Valès et al. (2016). Thus, our study replicated this finding in a larger and more heterogenous pABI sample. Furthermore, ecological validity of the subtests Key Search Test and Zoo Map Test 1 was demonstrated beyond the mere correlations with DEX-C and BRIEF teacher ratings, as they explained variation in these questionnaires even after controlling for conventional neuropsychological tests. However, this did not apply to the BADS-C total score. This can be explained by the fact that the total score is based on the results from all of the BADS-C subtests, including those that did not correlate well with teacher EF questionnaires. Based on face value, the Zoo Map and Key Search subtests appear to be more similar to real-life situations than the other subtests, as well as school tasks, which may explain the higher ecological validity of these measures. However, more exact explanations merit further research.

Although our findings suggest that parts of BADS-C may be more ecologically valid than traditional EF tests, the associations with teacher questionnaires of EF were far from perfect. This may suggest that BADS-C does not fully replicate everyday EF performance, even though it is developed in a slightly more open-ended and unstructured manner (Emslie et al., 2003). It is still a standardized paper-and-pencil task, with instructions, which may mask potential EF difficulties, however, possibly to a smaller degree than traditional EF tests (Chevignard et al., 2012; Jurado & Rosselli, 2007; Shallice & Burgess, 1991). Ecological tests with fewer behavioural constraints, such as the Children's Cooking Task (Chevignard et al., 2010), have been developed with the intention of simulating real-life situations. Although these tests may be more ecologically valid than traditional EF tests, the lack of control over key variables is a disadvantage. Thus, the BADS-C may serve as a middle-ground between the highly structured traditional EF tests and the highly open-ended ecological tasks.

The weak correlations between BADS-C and self-report ratings may indicate that self-report is a weak measure of cognitive performance compared to clinician or informant ratings, also suggested by previous research (e.g., Chaytor & Schmitter-Edgecombe, 2003). This can be related to numerous factors, such as reduced self-awareness following brain injury (Hart et al., 2005), or other common self-report biases, such as social desirability, recall bias, or context effects (Demetriou et al., 2014; Van de Mortel, 2008). Others have argued that tests and questionnaires measure different aspects of cognition (i.e., cognitive efficiency and success in goal-pursuit, respectively; Toplak et al., 2013), which might be another factor explaining their discrepancy. Nevertheless, self- and informant reports of everyday EF may still contribute with important additional outcome data, as they can assess elements of daily life EF functioning not captured by the structured nature of performance-based tests (e.g., social participation; Cicerone, 2004). Also, self-report may provide information on patients' awareness and subjective experience of their impairment, and hence guide clinicians in intervention propositions.

As most studies of the ecological validity of the BADS-C (Emslie et al., 2003; Roy et al., 2015; Shimoni et al., 2012; Siu & Zhou, 2014) and other EF tests (Toplak et al., 2013) have used parent-reports, our study adds an important perspective by including teacher-reports. The present study converges with the majority of evidence suggesting weak ecological validity for BADS-C and other EF tests, based on parent perceptions. However, the results are notably different when including teacher ratings. Taken together, this may suggest that performance on BADS-C should be considered in terms of context, and that it has more resemblance to structured and demanding school settings. On the other hand, the results may suggest that some teachers are better at evaluating the EF of their pupils compared to some parents. Teachers are able to compare EF performance with peers of the same age, while observing the child in an environment which may reveal EF dysfunction to a larger degree compared to the home context (e.g., expectations of efficient problem-solving on academic tasks, compliance in the classroom, and social interactions with other pupils). Also, we believe some parents may have more variability in reporting because personal and affective factors may be involved, in contrast to teacher's report being based on a more neutral relationship with the child. Thus, our study underpins the importance of including both parent and teacher perspectives on child executive functioning, as addressed by previous studies (Longaud-Valès et al., 2016).

Limitations

There are limitations in the present study that should be addressed. As there is no consensus concerning categorization of injury severity in the pABI population, we were unable to examine whether BADS-C is sensitive to capturing different degrees of impairment. As EF difficulties tend to increase with injury severity (Babikian & Asarnow, 2009), examining whether BADS-C is sensitive to this differentiation might be an important validity indicator, which is currently unexplored. Another issue is that parents of the participants in the control group had a higher level of education compared to parents of the pABI group. Also, the IQ levels of the controls were slightly above normative average. Thus, the control group seems to be biased towards being more resourceful compared to the normal population. This might have caused exaggerated differences between the pABI group and the control group.

Furthermore, as only participants with reported executive difficulties, motivated for participation in cognitive rehabilitation were included, our sample was most likely not representative of the entire pABI population. Also, our sample did not include younger children below 10 years of age. In addition, as executive dysfunction may cause fatigue, this may have contributed to a sample with a higher load of fatigue symptoms. Also, we cannot rule out potential sub-optimal performance caused by cognitive fatigue on tests administered late during the long test days (7–8 h).

The present study used two EF tests from D-KEFS as criteria for concurrent validity of the BADS-C. However, these tests only represent parts of EF, which is a complex construct, involving a wide array of cognitive functions. Previous research has found concurrent validity for BADS-C with some EF tests (i.e., Trail Making Test and Color-Word Interference Test) but not with others (i.e., Tower of London; Longaud-Valès et al., 2016). Thus, the limited number of EF tests included in our study does not demonstrate a full range of concurrent validity for the BADS-C, which may vary according to different tests. Also, as DEX-C only use raw scores, these were analysed in relation to scaled scores of other measures, leading to potential bias. Thus, results involving DEX-C should be interpreted with caution. Finally, there is no objective target measure of ecological validity, which means that it is unclear to what degree the measures of comparison to BADS-C are biased.

Conclusions

The findings of the present study indicate that BADS-C is a robust assessment tool in terms of discriminant validity, and partly of concurrent, and ecological validity for children and adolescents with pABI. While the results showed discriminant validity for the battery as a whole (except Playing Cards Test), Key Search Test and Zoo Map Test 1 stand out as the most valid measures with respect to concurrent and ecological validity. These measures seem to be more ecologically valid than traditional EF tests and IQ when assessed in relation to teacher perceptions of the child's or adolescent's EF. Thus, the findings indicate that BADS-C is able to capture executive dysfunction in patients with pABI, which, to some degree, is generalizable to the school context of everyday life. However, as it does not perfectly replicate everyday function, a comprehensive assessment of EF in the pediatric population should also include other methods, such as self- and informant reports.

Acknowledgements

We would like to thank the participants and their families, Head of Clinics, patient advisers, and the staff at St. Olavs Hospital and Oslo University Hospital for their participation and support. We would especially thank our research nurses, occupational therapists Lene Bobakk and Gøril Okkenhaug Johansen, test-technicians, neuropsychologists and pediatricians involved in this study. We also would like to thank pediatrician Bård Forsdal for help recruiting patients from at the University Hospital of North Norway, the Unit for Applied Clinical Research, Norwegian University of Science and Technology (NTNU) responsible for the monitoring and developing the webCRF, with special thanks to Anne Caroline Wiik.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Research Council of Norway (grant number 260680/H10); and grants from the Children's Clinic, St Olavs Hospital, Trondheim University Hospital.

ORCID

Stian Solem http://orcid.org/0000-0002-6942-2645
Torun Gangaune Finnanger http://orcid.org/0000-0003-4158-5642

Data availability statement

The raw data supporting the findings of the manuscript can be found at the Children's Clinic, St. Olavs Hospital, Trondheim, Norway. Due to regulations, the anonymity of the informants must be secured. In the raw data, it is possible to identify the informants, and restrictions therefore apply to the availability of these data. Reasonable requests concerning the data can be sent to the corresponding author.

References

Anderson, V., Catroppa, C., Morse, S., Haritou, F., & Rosenfeld, J. (2005). Functional plasticity or vulnerability after early brain injury? *Pediatrics*, *116*(6), 1374–1382. https://doi.org/10.1542/peds.2004-1728

Ardila, A. (2013). Development of metacognitive and emotional executive functions in children. *Applied Neuropsychology: Child*, *2*(2), 82–87. https://doi.org/10.1080/21622965.2013.748388

Arnett, A. B., Peterson, R. L., Kirkwood, M. W., Taylor, H. G., Stancin, T., Brown, T. M., & Wade, S. L. (2013). Behavioral and cognitive predictors of educational outcomes in pediatric traumatic brain injury. *Journal of the International Neuropsychological Society*, *19*(8), 881–889. https://doi.org/10.1017/S1355617713000635

Austin, G., Groppe, K., & Elsner, B. (2014). The reciprocal relationship between executive function and theory of mind in middle childhood: A 1-year longitudinal perspective. *Frontiers in Psychology*, *5*, 655. https://doi.org/10.3389/fpsyg.2014.00655

Babikian, T., & Asarnow, R. (2009). Neurocognitive outcomes and recovery after pediatric TBI: Meta-analytic review of the literature. *Neuropsychology*, *23*(3), 283–296. https://doi.org/10.1037/a0015268

Baddeley, A. D. (1986). *Working memory*. Clarendom Press.

Brandt, A. E., Finnanger, T. G., Hypher, R. E., Rø, T. B., Skovlund, E., Andersson, S., Risnes, K., & Stubberud, J. (2021). Rehabilitation of executive function in chronic paediatric brain injury:

A randomized controlled trial. *BMC Medicine*, *19*(1), 1–15. https://doi.org/10.1186/s12916-021-02129-8

Brydges, C. R., Reid, C. L., Fox, A. M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, *40*(5), 458–469.

Burgess, P. W., Alderman, N., Evans, J. O. N., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*(6), 547–558. https://doi.org/10.1017/S1355617798466037

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. A., Dawson, D. R., Anderson, N. D., Gilbert, S. J., Dumontheil, I., & Channon, S. (2006). 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*, *12*(2), 194–209. https://doi.org/10.1017/S1355617706060310

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, *13*(4), 181–197. https://doi.org/10.1023/B:NERV.0000009483.91468.fb

Chevignard, M. P., Catroppa, C., Galvin, J., & Anderson, V. (2010). Development and evaluation of an ecological task to assess executive functioning post childhood TBI: The children's cooking task. *Brain Impairment*, *11*(2), 125–143. https://doi.org/10.1375/brim.11.2.125

Chevignard, M. P., Soo, C., Galvin, J., Catroppa, C., & Eren, S. (2012). Ecological assessment of cognitive functions in children with acquired brain injury: A systematic review. *Brain Injury*, *26*(9), 1033–1057. https://doi.org/10.3109/02699052.2012.666366

Christensen, B. K., Colella, B., Inness, E., Hebert, D., Monette, G., Bayley, M., & Green, R. E. (2008). Recovery of cognitive function after traumatic brain injury: A multilevel modeling analysis of Canadian outcomes. *Archives of Physical Medicine and Rehabilitation*, *89*(12), S3–S15. https://doi.org/10.1016/j.apmr.2008.10.002

Cicerone, K. D. (2004). Participation as an outcome of traumatic brain injury rehabilitation. *Journal of Head Trauma Rehabilitation*, *19*(6), 494–501. https://doi.org/10.1097/00001199-200411000-00006

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Academic Press. https://doi.org/10.4324/9780203771587

Delis, D. C., Kaplan, E., & Kramer, J. (2001). *Delis Kaplan executive function system*. Psychological Corporation.

Demetriou, C., Ozer, B. U., & Essau, C. A. (2014). Self-report questionnaires. *The Encyclopedia of Clinical Psychology*, 1–6. https://doi.org/10.1002/9781118625392.wbecp507

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia*, *33*(3), 261–268. https://doi.org/10.1016/0028-3932(94)00124-8

Emslie, H., Burden, V., Nimmo-Smith, I., Wilson, B. A., & Wilson, C. (2003). *Behavioural Assessment of the Dysexecutive Syndrome for Children*. Thames Valley Test Company.

Emslie, H., Wilson, F. C., Burden, V., Nimmo-Smith, I., & Wilson, B. A. (2013). *Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C). Supplement to the professional manual*. Thames Valley Test Company.

Franzen, M. D., & Wilhelm, K. L. (1996). Conceptual foundations of ecological validity in neuropsychological assessment. In R. J. Sbordone & C. J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 91–112). Gr Press/St Lucie Press, Inc.

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179. https://doi.org/10.1111/j.1467-9280.2006.01681.x

Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Test Review: Behavior Rating Inventory of Executive function. *Child Neuropsychology*, *6*(3), 235–238. https://doi.org/10.1076/chin.6.3.235.3152

Hart, T., Whyte, J., Kim, J., & Vaccaro, M. (2005). Executive function and self-awareness of "real-world" behavior and attention deficits following traumatic brain injury. *Journal of Head Trauma Rehabilitation*, *20*(4), 333–347. https://doi.org/10.1097/00001199-200507000-00005

Hypher, R. E., Brandt, A. E., Risnes, K., Rø, T. B., Skovlund, E., Andersson, S., Finnanger, T. G., & Stubberud, J. (2019). Paediatric goal management training in patients with acquired brain injury: Study protocol for a randomised controlled trial. *BMJ Open*, *9*(8), e029273. https://doi.org/10.1136/bmjopen-2019-029273

Jurado, M. B., & Rosselli, M. (2007). The elusive nature of executive functions:a review of our current understanding. *Neuropsychol. Rev*, *17*, 213–233.

Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. Oxford University Press.

Longaud-Valès, A., Chevignard, M., Dufour, C., Grill, J., Puget, S., Sainte-Rose, C., Valteau-Couanet, D., & Dellatolas, G. (2016). Assessment of executive functioning in children and young adults treated for frontal lobe tumours using ecologically valid tests. *Neuropsychological Rehabilitation*, *26*(4), 558–583. https://doi.org/10.1080/09602011.2015.1048253

McCabe, D. P., Roediger III, H. L., McDaniel, M. A., Balota, D. A., & Hambrick, D. Z. (2010). The relationship between working memory capacity and executive functioning: Evidence for a common executive attention construct. *Neuropsychology*, *24*(2), 222–243. https://doi.org/10.1037/a0017619

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Moriguchi, Y. (2014). The early development of executive function and its relation to social interaction: A brief review. *Frontiers in Psychology*, *5*, 388. https://doi.org/10.3389/fpsyg.2014.00388

Nudo, R. J. (2013). Recovery after brain injury: Mechanisms and principles. *Frontiers in Human Neuroscience*, *7*, 887. https://doi.org/10.3389/fnhum.2013.00887

Nybo, T., Sainio, M., & Müller, K. (2004). Stability of vocational outcome in adulthood after moderate to severe preschool brain injury. *Journal of the International Neuropsychological Society*, *10*(5), 719–723. https://doi.org/10.1017/S1355617704105109

O'Keeffe, F., Liégeois, F., Eve, M., Ganesan, V., King, J., & Murphy, T. (2014). Neuropsychological and neurobehavioral outcome following childhood arterial ischemic stroke: Attention deficits, emotional dysregulation, and executive dysfunction. *Child Neuropsychology*, *20*(5), 557–582. https://doi.org/10.1080/09297049.2013.832740

Roth, R. M., Isquith, P. K., & Gioia, G. A. (2014). Assessment of executive functioning using the Behavior Rating Inventory of Executive Function (BRIEF). In *S. Goldstein & J.A. Naglieri (eds.), Handbook of executive functioning* (pp. 301–331). Springer Science + Business Media. https://doi.org/10.1007/978-1-4614-8106-5_18

Roy, A., Allain, P., Roulin, J. L., Fournet, N., & Le Gall, D. (2015). Ecological approach of executive functions using the Behavioural Assessment of the Dysexecutive Syndrome for Children (BADS-C): Developmental and validity study. *Journal of Clinical and Experimental Neuropsychology*, *37*(9), 956–971. https://doi.org/10.1080/13803395.2015.1072138

Shallice, T. I. M., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(2), 727–741. https://doi.org/10.1093/brain/114.2.727

Shimoni, M. A., Engel-Yeger, B., & Tirosh, E. (2012). Executive dysfunctions among boys with Attention Deficit Hyperactivity Disorder (ADHD): Performance-based test and parents report. *Research in Developmental Disabilities*, *33*(3), 858–865. https://doi.org/10.1016/j.ridd.2011.12.014

Siu, A. F., & Zhou, Y. (2014). Behavioral assessment of the dysexecutive syndrome for children: An examination of clinical utility for children with attention-deficit hyperactivity disorder (ADHD). *Journal of Child Neurology*, *29*(5), 608–616. https://doi.org/10.1177/2F0883073813516191

Stephens, T. L. (2014). The assessment of executive functioning using the delis-kaplan executive functions system (D-KEFS). In *Handbook of executive functioning* (pp. 209–222). Springer. https://doi.org/10.1007/978-1-4614-8106-5_13

Thurman, D. J. (2016). The epidemiology of traumatic brain injury in children and youths: A review of research since 1990. *Journal of Child Neurology*, *31*(1), 20–27. https://doi.org/10.1177/0883073814544363

Toplak, M. E., West, R. F., & Stanovich, K. E. (2013). Practitioner review: Do performance- based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*(2), 131–143. https://doi.org/10.1111/jcpp.12001

Van de Mortel, T. F. (2008). Faking it: Social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, The*, *25*(4), 40–48. https://search.informit.org/doi/10.3316/informit.210155003844269

Ward, C., Phipps, K., de Sousa, C., Butler, S., & Gumley, D. (2009). Treatment factors associated with outcomes in children less than 3 years of age with CNS tumours. *Child's Nervous System*, *25*(6), 663–668. https://doi.org/10.1007/s00381-009-0832-8

Wechsler, D. (2014). *Wechsler intelligence scale for children-fifth edition*. Pearson.

White, S. J., Burgess, P. W., & Hill, E. L. (2009). Impairments on "open-ended" executive function tests in autism. *Autism Research*, *2*(3), 138–147. https://doi.org/10.1002/aur.78

Wilson, B. A., Evans, J. J., Alderman, N., Burgess, P. W., & Emslie, H. (1997). Behavioural assessment of the dysexecutive syndrome. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 232–239). Psychology Press.

Wood, R. L., & Liossi, C. (2007). The relationship between general intellectual ability and performance on ecologically valid executive tests in a severe brain injury sample. *Journal of the International Neuropsychological Society*, *13*(1), 90. https://doi.org/10.1017/s1355617707070129

World Health Organization (WHO). (2009). *World health statistics. Geneva*.

Ylvisaker, M., & Feeney, T. (2007). Pediatric brain injury: Social, behavioral, and communication disability. *Physical Medicine and Rehabilitation Clinics of North America*, *18*(1), 133–144. https://doi.org/10.1016/j.pmr.2006.11.007