Journal of the International Neuropsychological Society (2015), **21**, 596–609.

doi:10.1017/S1355617715000636

Empirical Derivation and Validation of a Clinical Case Definition for Neuropsychological Impairment in Children and Adolescents

Miriam H. Beauchamp,[1,2] Brian L. Brooks,[3–6] Nick Barrowman,[6] Mary Aglipay,[7] Michelle Keightley,[8–10] Peter Anderson[11,12] Keith O. Yeates,[5,6,13] Martin H. Osmond,[12,14] AND Roger Zemek[12,14]

[1]Department of Psychology, University of Montreal, Montreal, Quebec, Canada
[2]Ste-Justine Hospital Research Center, Montreal, Quebec, Canada
[3]Neurosciences program (Brain Injury and Rehabilitation), Alberta Children's Hospital, Calgary, Alberta, Canada
[4]Departments of Pediatrics and Clinical Neurosciences, University of Calgary, Calgary, Alberta, Canada
[5]Alberta Children's Hospital Research Institute, University of Calgary, Calgary, Alberta, Canada
[6]Department of Psychology, University of Calgary, Calgary, Alberta, Canada
[7]Clinical Research Unit, Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada
[8]Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada
[9]Departments of Occupational Science and Occupational Therapy and Rehabilitation Sciences Institute, University of Toronto, Toronto, Ontario, Canada
[10]Toronto Rehabilitation Institute, Toronto, Ontario, Canada
[11]Behavioural Neurosciences & Consultation-Liaison program, Children's Hospital of Eastern Ontario, Ottawa, ON
[12]Children's Hospital of Eastern Ontario Research Institute, Ottawa, Ontario, Canada
[13]Hotchkiss Brain Institute, University of Calgary, Calgary, Alberta, Canada
[14]Departments of Pediatrics and Emergency Medicine, University of Ottawa, Ottawa, Ontario, Canada

(RECEIVED September 25, 2014; FINAL REVISION July 15, 2015; ACCEPTED July 16, 2015; FIRST PUBLISHED ONLINE August 26, 2015)

Abstract

Neuropsychological assessment aims to identify individual performance profiles in multiple domains of cognitive functioning; however, substantial variation exists in how deficits are defined and what cutoffs are used, and there is no universally accepted definition of neuropsychological impairment. The aim of this study was to derive and validate a clinical case definition rule to identify neuropsychological impairment in children and adolescents. An existing normative pediatric sample was used to calculate base rates of abnormal functioning on eight measures covering six domains of neuropsychological functioning. The dataset was analyzed by varying the range of cutoff levels [1, 1.5, and 2 standard deviations (*SD*s) below the mean] and number of indicators of impairment. The derived rule was evaluated by bootstrap, internal and external clinical validation (orthopedic and traumatic brain injury). Our neuropsychological impairment (NPI) rule was defined as "two or more test scores that fall 1.5 *SD*s below the mean." The rule identifies 5.1% of the total sample as impaired in the assessment battery and consistently targets between 3 and 7% of the population as impaired even when age, domains, and number of tests are varied. The NPI rate increases in groups known to exhibit cognitive deficits. The NPI rule provides a psychometrically derived method for interpreting performance across multiple tests and may be used in children 6–18 years. The rule may be useful to clinicians and scientists who wish to establish whether specific individuals or clinical populations present within expected norms *versus* impaired function across a battery of neuropsychological tests. (*JINS*, 2015, *21*, 596–609)

Keywords: Neuropsychology, Assessment, Pediatric, Cognition, Performance, Brain injury

INTRODUCTION

The broad aim of pediatric neuropsychological assessment is to provide a comprehensive evaluation of neurocognitive development and to establish profiles of strengths and weaknesses that quantify brain-behavior relationships. By definition, this suggests the inclusion of a broad number of domains of functioning in assessment protocols and their corresponding test batteries and measures. In some cases, multiple subtests measuring the same or similar functions are necessary to assess individuals' performances. Although such assessments are designed to provide a broad overview of a child's abilities, they also generate a large number of observations that ultimately need to be analyzed, interpreted, and integrated into a logical, evidence-based conclusion. Despite advances in the field and a certain commonality in the way

Correspondence and reprint requests to: Miriam Beauchamp, Department of Psychology, University of Montreal, C.P. Succursale Centre-Ville, Montréal, Québec, Canada, H3C 3J7. E-mail: miriam.beauchamp@umontreal.ca

neuropsychological assessments are performed, no universally accepted definition exists to identify neuropsychological impairment.

The absence of a clear definition may constitute a problem in clinical settings when lengthy, personalized test batteries are used. In such cases, conclusions about neuropsychological impairment may be based on the presence of limited and isolated low test scores. The lack of a quantitative definition also poses a challenge in research settings where scientists need to determine whether an individual is considered "impaired" or not, or when impairment needs to be defined in specific clinical populations (e.g., a dichotomized decision rule). This is especially true for pediatric brain injuries where there is substantial heterogeneity in the nature and degree of impairment (Beauchamp & Anderson, 2013; Hung et al., 2014; Yeates, 2010). Studies of neuropsychological outcome after pediatric brain injury typically report outcomes separately across a range of measures and domains without inferring whether individual profiles of performance reflect general neurocognitive impairment (Anderson et al., 2009; Conklin, Salorio, & Slomine, 2008; Levin et al., 2013; Rieger et al., 2013). Clinical *versus* control group comparisons provide evidence regarding the effect of brain injury on neurocognitive performance, but some research designs preclude such comparisons. Central to the purpose of the current study is the difficulty of defining neuropsychological impairment to identify a "case" (i.e., an individual with a particular clinical condition who demonstrates neuropsychological impairment) when no comparison group is present.

The existing literature suggests caution when interpreting multiple test scores to identify true cognitive impairment (Binder, Iverson, & Brooks, 2009; Brooks & Iverson, 2012; Crawford, Garthwaite, & Gault, 2007; Hurks, Hendriksen, Dek, & Kooij, 2013). When interpreting a single test score in isolation, Gaussian distributions (bell curve) apply and clinicians can be confident that the rate of impairment is defined by the individual test score cutoff (i.e., 10% impairment rate when cutoff is at the 10^{th} percentile). However, as mentioned above, neuropsychological assessments involve numerous observations or test scores that are not interpreted in isolation. Brooks and Iverson (2012) summarize several principles for consideration when interpreting multiple test scores to minimize the chances of a false positive, including: (1) test-score scatter or variability is common (i.e., it is uncommon to have all scores fall at the same level); (2) having some low scores is common (i.e., all cognitive batteries will have a "normal" number of low scores); (3) the number of low scores is related to the cutoff score used for interpretation (i.e., using a more liberal cutoff score will result in more low scores and a higher false positive rate); (4) the number of low scores depends on the number of tests administered (i.e., more tests administered and interpreted will result in more low scores and a higher false positive rate); and (5) the number of low scores varies by characteristics of the examinees (i.e., persons with lower intelligence or fewer years of education are likely to have more low scores than those with higher intelligence or more years of education).

Without adequate consideration of these principles, clinicians and researchers are at risk of higher-than-desired rates of suggested cognitive impairment when interpreting multiple neuropsychological scores in children and adolescents (Brooks, 2010; Brooks, Iverson, Feldman, & Holdnack, 2009; Brooks et al., 2013; Brooks, Iverson, Sherman, & Holdnack, 2009; Brooks, Iverson, Sherman, & Roberge, 2010; Brooks, Sherman, & Iverson, 2010; Crawford et al., 2007; Hurks et al., 2013). Although these principles add another layer of interpretation with multiple test scores, understanding and considering multivariate principles may provide an opportunity for developing empirically derived criteria for identifying abnormal neuropsychological profiles (see Brooks, Iverson, Feldman, & Holdnack, 2009, for an example of the development and application of criteria for identifying cognitive impairment).

To date, there is no universally accepted definition of neuropsychological impairment. Elaborating such a definition would assist in identifying individual cases of impairment. The objective of this study was to empirically derive a Neuropsychological Impairment (NPI) rule to provide a definition that includes a range of functional and cognitive domains and which may be applied to children between the ages of 6 and 18 years, regardless of etiology or nature of neurological, developmental or psychiatric disorders. The NPI rule was derived on the basis of existing normative pediatric neuropsychological data, collected in the National Institutes of Health Pediatric Database project (NIHPD, Evans & Brain Development Cooperative Group, 2006). Associations with parental income and education were explored to verify the hypothesis that impairment varies as a function of examinee characteristics.

METHODS

Reference Group

A normative reference group was created using participants selected from the NIHPD pertaining to the "first objective" of the MRI Study of Normal Brain Development (Evans & Brain Development Cooperative Group, 2006). This sample includes healthy children between the ages of 4 years, 6 months and 18 years, 3 months (at time of enrolment) recruited across multiple sites using community based sampling to reflect the demographic composition of the U.S. population with respect to sex, socioeconomic status (SES), and race/ethnicity and, including both males and females and right- and left-handed individuals. Participants were carefully screened for potential exclusionary criteria, which included any history or evidence of medical conditions that impact healthy brain development (see Evans & Brain Development Cooperative Group, 2006, for details). In total, 433 children were enrolled across six Pediatric Study Centers. In the current study, only children between 6 and 18 years were included and the neuropsychological results from the initial assessment were used, resulting in a sample of

370 individuals. See Table 1 for the participant demographic characteristics. This research was completed in accordance with the Helsinki Declaration and institutional research standards for human research.

Outcome Measures

In the NIHPD study, outcome measures were chosen to allow correlations with neuroimaging markers and selection was guided by three requirements: (1) Description of cognitive and sensory motor abilities; (2) Description of academic skills; (3) Quantification of memory and executive functions (Complete protocol details can be found at http://pediatricmri.nih.gov/nihpd/info/protocols.html, see also Waber et al., 2007). To meet the first requirement, standardized neuropsychological tests available for use by others were chosen and included the following: Wechsler Abbreviated Scale of Intelligence (WASI, Wechsler, 1999), Purdue Pegboard (Tiffin, 1968; Tiffin & Asher, 1948), and NEPSY verbal fluency (Korkman, Kirk, & Kemp, 1998). The second requirement was met by including three sub-tests from the Woodcock-Johnson-III Achievement Battery and the Pre-School Language Scale (Mather & Woodcock, 2001; Zimmerman, Steiner, & Pond, 2002). For the third requirement, verbal learning and memory was assessed using the California Verbal Learning Test (Delis, Kramer, Kaplan, & Ober, 1994), working memory was assessed using the Spatial Span and Spatial Working Memory subtests of the Cambridge Neuropsychological Test Automated Battery (CANTAB, Sahakian et al., 1988; Sahakian & Owen, 1992), and executive functions were quantified using the CANTAB Intra/Extra-dimensional (ID/ED) set-shifting subtest. No basic measure of attentional skills was included in the NIHPD study. Instead, the presence of attentional difficulties was indirectly measured using the Attention Problems subscale of the Child Behavior Checklist (Achenbach, 1991; Evans & Brain Development Cooperative Group, 2006).

The set of subtests used for derivation of the NPI rule was drawn from the original NIHPD test battery and established by a national group of Canadian experts with the following constraints: (1) Ability to evaluate outcome on a range of neuropsychological domains; (2) Length of assessment time limited to approximately 2 hr to ensure feasibility in both clinical and research settings; (3) Target functions that are known or suspected to be altered in a range of neuro-developmental and acquired conditions. Academic measures were not included in the initial derivation dataset because these skills are often secondarily affected by primary deficits in core cognitive functions. We chose not to include scores derived from parental questionnaires, such as the Child Behavior Checklist, because they are more subject to bias and are a reflection of parental perception rather than a direct observation of performance. We used the CANTAB Spatial Span subtest to represent the attentional domain in the current analyses. Although this subtest is considered to measure working memory, it overlaps with selective attention skills (Gazzaley & Nobre, 2012). The final list of subtests for the rule derivation covered six domains of neuropsychological functioning (IQ, Attention/Working memory, Memory,

Table 1. Characteristics of the NIHPD sample

	Total $N = 370$	Group 1 $N = 95$	Group 2 $N = 156$	Group 3 $N = 119$
No. of males (%)	176 (47.6)	45 (47.4)	68 (43.6)	63 (52.9)
Mean age, months (*SD*)	132.96 (42.53)	82.9 (6.3)	123.5 (16.7)	185.4 (18.7)
Mean IQ (*SD*)	110.7 (12.5)[a]	110.1 (14.1)	111.9 (12.4)	109.5 (11.3)
Education level (respondent), *n* (%)				
Graduate Level	93 (25.1)	24 (25.4)	38 (24.4)	31 (26.1)
College	221 (59.7)	61 (64.2)	92 (59.0)	68 (57.1)
High school	56 (15.1)	10 (10.5)	26 (16.7)	20 (16.8)
Race, *n* (%)				
African American or Black	34 (9.2)	11 (11.6)	10 (6.4)	13 (10.9)
American Indian or Alaskan Native	3 (0.8)	1 (1.1)	2 (1.3)	0 (0.0)
Asian	5 (4.4)	0 (0.0)	4 (2.6)	1 (0.8)
White	302 (81.6)	81 (85.3)	125 (80.1)	96 (80.7)
Not provided	26 (7.0)	2 (2.1)	15 (9.6)	9 (7.6)
Right handedness, *n* (%)	324 (87.6)	81 (85.3)	139 (89.1)	104 (87.4)
Household income level, *n* (%)				
$100 0001 to $150 000	81 (21.9)	20 (21.1)	29 (18.6)	32 (26.9)
$75 001 to $100 000	91 (24.6)	24 (25.3)	44 (28.2)	23 (19.3)
$50 001 to $75 000	92 (24.9)	26 (27.4)	37 (23.7)	29 (24.4)
$35 001 to $50 000	68 (18.4)	16 (16.8)	27 (17.3)	25 (21.0)
Zero to $35 000	34 (9.2)	9 (9.5)	16 (10.3)	9 (7.6)
Not provided	4 (1.1)	0 (0.0)	3 (1.9)	1 (0.8)

Note: Group 1 = 6–8 years, Group 2 = 9–12 years, Group 3 = 13–18 years.
[a]The average IQ level for the NIHPD sample ($M = 110.7$, $SD = 12.5$, 95% CI [109.4, 111.9]) was significantly higher than the normative mean of 100 ($SD = 10$) ($p < .001$ by a one-sample t-test).

Table 2. List of the NIHPD domains and subtests used for rule derivation and internal validation and comparable subtests used from Yeates et al. (2002) for external clinical validation

Domain	NIHPD subtest/variable	Type of score	Yeates et al. (2002)[c] subtest/variable	Type of score
Original rule derivation				
Intellectual functioning	WASI FSIQ	Standard	N/A	N/A
Attention/Working memory	Digit Span	Scaled	CPT-3 omissions	*Z*-score
	CANTAB spatial span[a]	Raw	CELF Recalling sentences	Standard
Memory	CVLT Total words recalled Trials 1–5	Raw	CVLT Total words recalled Trials 1–5	*Z*-score
Executive functions	CANTAB ID/ED stage errors	Raw	Contingency Naming Test Efficiency	*Z*-score
	NEPSY Letter Fluency[b] total words	Scaled	COWA Letter Fluency total words	*Z*-score
Processing speed	Coding	Scaled	Underlining Test	*Z*-score
Motor function	Purdue Pegboard number of pegs both hands	Raw	Grooved Pegboard number of pegs both hands	*Z*-score
Short internal validation (10 variables)				
– ~~Intellectual functioning~~	~~WASI FSIQ~~	~~Standard~~	—	—
+ Language	WASI Vocabulary	T-score	WISC Vocabulary	Scaled
+ Perceptual/Visuo-constructive	WASI Block Design	T-score	WISC Block Design	Scaled
+ Academic	WJ Letter-Word Identification	Standard	WJ Letter-Word Identification	Standard
Extended internal validation (16 variables)				
+ Attention/Working memory	CANTAB spatial working memory errors	Raw	—	—
+ Visuo-constructive/Perceptual reasoning	WASI Matrix Reasoning	T-score	—	—
+ Language	WASI Similarities	T-score	—	—
+ Academic	WJ Calculation	Standard	—	—
	WJ Passage Comprehension		—	—

[a]The CANTAB spatial span was used in the rule derivation as an indicator of attention.
[b]In the NIHPD sample verbal fluency was only administered to children 7 years and older, therefore the case definition for Group 1 (6–8 years) was based on 6 domains with 7 subtests, rather than 8 subtests in the older age groups.
[c]For external clinical validation, data was used with permission from : Yeates, K. O., Taylor, H. G., Wade, S., Drotar, D., Stancin, T., & Minich, N. (2002). A prospective study of short- and long-term neuropsychological outcomes after traumatic brain injury in children. *Neuropsychology, 16*(4), 514–523. Please see original manuscript for test descriptions and references.
CANTAB = Cambridge Neuropsychological Test Automated Battery; CELF = Clinical Evaluation of Language Fundamentals; CNT = Contingency Naming Test; CPT = Conners Performance Test; CVLT = California Verbal Learning Test; COWA = Controlled Oral Word Association Test; NEPSY = Developmental Neuropsychological Assessment; WASI = Wechsler Abbreviated Scale of Intelligence; WISC = Wechsler Intelligence Scale for Children-III; WJ = Woodcock-Johnson Tests of Cognitive Abilities.

Executive functions, Processing speed, Motor skills) using eight measures (see Table 2).

Analyses

All statistical analyses were performed using SPSS version 21 (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.).

The NIHPD study included both experimental and normed tests, resulting in a mix of standard (M = 100; *SD* = 10), scaled (M = 10; *SD* = 3), and raw scores for the key variables. In light of this, individuals were separated into three age groups to ensure greater comparability of performance. The three groups (early school period, Group 1: 6–8 years; middle school, Group 2: 9–12 years; adolescence, Group 3: 13–18 years) were formed using a developmental rationale reflecting cerebral growth spurts (Casey, Jones, & Hare, 2008; Huttenlocher, 1979; Huttenlocher & Dabholkar, 1997; Kolb, Pellis, & Robinson, 2004) and trajectories of memory and executive function maturation (Gathercole, 1998). Similar age groups have also been used in clinical research on brain injury (Anderson et al., 2009; Sady, Vaughan, & Gioia, 2014; Zemek, Osmond, Barrowman, & Pediatric Emergency Research Canada Concussion Team, 2013).

Standard and scaled scores were used for subtests when available, otherwise raw scores were used. The distributional properties of scores were assessed graphically using qq-plots, histograms, and boxplots. The distributions of the raw scores were found to deviate from normality (Shapiro-Wilk test, CANTAB ID/ED: $W = .93$; $p < .001$; CANTAB spatial span: $W = .93$; $p < .001$; CVLT trials 1–5: $W = .96$; $p < .001$; Pegboard: $W = .97$; $p < .001$), therefore, cutoffs were based on percentiles rather than means and standard deviations (*SD*s) for these scores. Data distributions for all subtests were checked for outliers with an interest primarily in underperforming individuals. No individuals were excluded on this basis. Frequency distributions were generated and the number of individuals performing at three different cutoff levels (1, 1.5, and 2 *SD*s below the mean, or

https://doi.org/10.1017/S1355617715000636 Published online by Cambridge University Press

Table 3. Characteristics of the samples from Yeates et al. (2002).

	Orthopedic injury $N = 62$	Traumatic brain injury (TBI)[c] $N = 98$	TBI mild/moderate $N = 53$	TBI severe $N = 45$
No. of males (%)	37 (60)	74 (76)	38 (72)	36 (80)
Mean age at testing, years (*SD*)	10.00 (1.89)	10.31 (2.03)	10.52 (1.91)	10.07 (2.16)
Mean age at injury, years (*SD*)	9.39 (1.88)	9.72 (2.02)	9.92 (1.90)	9.48 (2.16)
Mean IQ[a] (*SD*)	102.32 (16.28)	99.57 (18.30)	100.21 (17.62)	98.79 (19.28)
Education level[b] (mother)				
Graduate level, *n* (%)	5 (8.1)	8 (8.4)	5 (9.8)	3 (6.8)
College, *n* (%)	19 (30.6)	34 (35.8)	16 (31.4)	18 (40.9)
High school, *n* (%)	38 (61.3)	53 (55.8)	30 (58.8)	23 (52.3)
Race – number white vs minority (%)	38 (61)	77 (79)	41 (77)	36 (80)

[a]IQ level based on performance on the Wechsler Intelligence Scale for Children Third Edition (WISC-III, Wechsler, 1991)
[b]Education level categories were defined as Graduate degree, College = 1–4 years of college education, High school = high school graduate or fewer years of education.
[c]N = 45 severe TBI and 53 mild/moderate TBI; lowest GCS for TBI group: M = 9.79 (4.9), range 3–15.

the equivalent percentiles based on a normal distribution, namely the 15.9th, 6.7th, and 2.3rd percentiles) were identified by age group for each subtest. Our goal was to classify roughly 95% of the population as "typically functioning," as this arbitrary definition offers a relatively strict specification of "caseness."

Following rule derivation, the performance of the Neuropsychological Impairment Rule (NPI rule) was studied in three ways. First, the proportion of individuals impaired as a function of measure and domain was reported and these proportions were compared using McNemar's tests with Holm-Bonferroni adjustments. Second, to explore putative links between socioeconomic factors and neuropsychological performance, the distribution of number of indicators of impairment was generated according to level of parental education and income and associations between these variables were assessed using the linear-by-linear χ^2 test. The size of the effect of these comparisons was assessed *via* spearman rank correlation. Third, validation of the rule was conducted as follows: (1) *Bootstrap validation:* Bootstrap analysis was performed with 1000 samples to demonstrate the expected variability in classification of impairment; (2) *Internal validation:* Additional subtests were added to the battery from the NIHPD dataset to test the performance of the rule when the number of functional domains and variables increases. A dataset containing 10 variables covering eight domains of functioning was created by removing the WASI full scale IQ variable and replacing it by two of its components pertaining to language (Vocabulary) and visuo-constructive skills (Block Design), as well as adding in a test of academic functioning (Woodcock-Johnson Letter-Word Identification, Mather & Woodcock, 2001) (see Table 2 for details of all datasets and versions used in the derivation and validation of the NPI rule). An extended battery was then created by increasing the number of variables, and therefore the number of correlations, but keeping the number of domains of functioning stable (see Table 2). Pearson correlations were calculated between all the variables included from the NIHPD dataset. (3) *External clinical validation:* The NPI rule was applied to two novel datasets from a previous study (Yeates et al., 2002) of children known to have higher rates of neuropsychological impairment than normative or community control samples: a control group of children with orthopedic injuries (OI) and children who sustained mild/moderate or severe traumatic brain injury (TBI). The sample characteristics for these groups are presented in Table 3. In comparison to the NIHPD sample, children from the orthopedic control group were not screened for pre-existing learning or attentional problems (Yeates et al., 2002), had significantly lower levels of intellectual functioning ($t = 3.84$; $p < .001$), were from families with lower levels of education ($\chi^2 = 68.80$; $p < .001$), and were more often of minority status (white vs other race, $\chi^2 = 11.91$; $p < .001$) than those in the NIHPD sample. Ten variables were chosen to mirror the variables and eight domains used from the NIHPD dataset (see Table 2). As the choice of measures differs substantially between research projects and in clinical practice, the goal of this external validation was to test the performance of the rule when different tests are used and to assess the performance of the rule in populations suspected to exhibit higher levels of impairment. Rates of impairment were then calculated using the NPI rule for the OI and TBI groups. Pearson correlations were calculated between all the variables selected from the Yeates et al. (2002) study.

RESULTS AND RULE DERIVATION

The results for the NIHPD participants on each of the neuropsychological tests included in the derivation dataset are presented according to age group in Table 4, and correlations between variables included in the NIHPD dataset (all versions) are presented in Table 5. The number of individuals with impairments on any of the eight measures (indicators of impairment) in the original dataset is presented in Table 6 for each age group and the total sample as a function of cutoff score.

Table 4. Mean performance (SD) of individuals in the NIHPD sample on the eight neuropsychological measures used in the initial derivation by age

Domain	Measures	Group 1 (*n* = 95)	Group 2 (*n* = 156)	Group 3 (*n* = 119)	Total (*n* = 370)
Intellectual functioning	WASI FSIQ	110.1 (14.1)	111.9 (12.4)	109.5 (11.3)	110.7 (12.5)
Attention/Working memory	Digit Span	10.7 (2.5)	10.5 (2.7)	10.8 (2.8)	10.7 (2.7)
	CANTAB spatial span	4.3 (1.0)	5.5 (1.2)	7.2 (1.6)	5.8 (1.7)
Memory	CVLT Trials 1–5	37.8 (10.3)	50.5 (7.7)	55.1 (7.3)	48.6 (10.7)
Executive functions	CANTAB ID/ED	25.0 (17.0)	26.2 (12.7)	18.7 (12.9)	23.5 (14.4)
	NEPSY letter fluency	N/A	21.5 (8.2)	30.8 (9.4)	23. 9 (10.5)
Processing speed	Coding	10.3 (3.1)	10.5 (2.9)	10.7 (3.2)	10.5 (3.0)
Motor function	Purdue Pegboard	7.6 (1.6)	9.9 (1.6)	11.0 (1.6)	9.6 (2.1)

NPI Rule Derivation

Indicators of impairment on each of the individual measures were considered at three cutoff scores, i.e., 1, 1.5, or 2 *SD*s below the mean. Frequency distributions are presented for all three levels, as there is not universal agreement on cutoffs for the interpretation of cognitive impairment. Higher cutoff scores have improved sensitivity for detecting cognitive problems, but have reduced specificity (Brooks & Iverson, 2012). A relatively liberal definition of impairment was first considered by the research group using a −1 *SD* cutoff. Although this cutoff may be ill-suited for defining impairment in normative populations because it would rarely be considered as clinically meaningful in a typically developing child, it may be relevant to use in settings where researchers or clinicians seek to identify more subtle deviations from average performance or where the accumulation of minor difficulties may impact on functioning. A more conservative definition was considered by using a −2 *SD* cutoff. Although this cutoff for defining impairment has been used in clinical and research settings, it is considered to be a striking deficit and strict cutoff. Arguably, performances at 1.5 *SD* below the average may also be clinically and functionally significant.

In the NIHPD sample, the −2 *SD* cutoff lead to a very conservative interpretation of performance, with only 0.5% of the sample showing a deficit on two or more subtests and 8.9% showing a deficit on one or more subtests (see Table 6). The −1.5 *SD* cutoff offered a less restrictive definition of impairment, with 5.1% of individuals defined as impaired on two or more subtests and 24.1% on one or more subtests . The −1 *SD* cutoff identified a larger number of typically developing children as being impaired, with 25.5% of children impaired on two or more subtests and 53.8% on one or more subtests.

A further consideration in deriving a rule for neuropsychological impairment is the number of indicators of impairment that are necessary. The results in Table 6 indicate that very few individuals had three or more indicators of impairment, regardless of the cutoff used. At a −1 *SD* cutoff, using three indicators would identify 9.8% of the population as impaired, while using a −1.5 *SD* cutoff would identify impairment in only 0.8% of the sample. No individuals would be identified as impaired when using three indicators at a −2 *SD* cutoff. Most individuals had no deficits on any subtests or only on one subtest for that cutoff.

Thus, based on the frequency distributions presented in Table 6, the definition of neuropsychological impairment that best fits the NIHPD data and identifies approximately 95% of the population as "typically developing" is the following: *"A neuropsychological impairment is present when an individual performs 1.5 standard deviations below the mean on two or more measures."* Our derived NPI rule identifies 5.1% of the total sample as impaired on two or more of the eight subtests in the assessment battery, which covers six domains of neuropsychological functioning. Applying this rule to individual age groups identifies between 3.0 and 7.2% of the population as impaired, suggesting that the case definition is appropriate for children between the ages of 6 and 18 years.

Working with this definition, we sought to explore which tests and domains of functioning were most frequently impaired and whether there was any systematic pattern of impairment (see Table 7). The highest percentages of impairment were found on the CANTAB ID/ED (6.5%) and the CVLT (6.1%), while the lowest percentage was found on the WASI FSIQ, with only one individual showing a deficit (0.3%). When the proportion of individuals impaired on each test was compared (see Table 8), most tests did not have a significantly different proportion of individuals with impairment. However, six tests had a significantly higher proportion of impairment compared to the WASI FSIQ. As noted previously, the NIHPD sample had above average intellectual functioning. This may have contributed to the difference between the IQ measure and other subtests.

Associations with Income and Education

Significant negative associations were found between number of indicators of neuropsychological impairment and parental level of education ($\chi^2(1,368) = 9.19$; $p = .002$) and income ($\chi^2(1,366) = 12.97$; $p < .001$) (see Table 9), with those children from families with greater income and higher levels of education demonstrating fewer indicators of neuropsychological impairment. When the number of

https://doi.org/10.1017/S1355617715000636 Published online by Cambridge University Press

Table 5. Pearson correlations between tests from the NIHPD database

Test battery	Subtest	FSIQ	Block Design	Matrix Reasoning	Similarities	Vocabulary	Digit Span	Coding	IED	Spatial Span	Spatial Working Memory	Verbal Fluency	CVLT	Calculation	Letter-Word Identification	Passage Comprehension	Purdue Both Hands	Purdue Left Hand	Purdue Right Hand
WASI	FSIQ	1	.68**	.70**	.73**	.77**	.32**	.26**	−.32**	.15**	−.13*	.17**	.25**	.49**	.46**	.55**	.15**	.15**	.13*
WASI	Block Design	.68**	1	.38**	.25**	.30**	.18**	.22**	−.25**	.22**	−.14**	.11*	.13*	.34**	.17**	.29**	.20**	.20**	.12*
WASI	Matrix Reasoning	.70**	.38**	1	.31**	.41**	.16**	.16**	−.19**	.07	−.14**	−.01	.09	.39**	.33**	.36**	.03	<.01	−.02
WASI	Similarities	.73**	.25**	.31**	1	.54**	.26**	.15**	−.23**	<.01	.01	.11	.20**	.31**	.39**	.43**	.06	.06	.08
WASI	Vocabulary	.78**	.30**	.41**	.54**	1	.32**	.24**	−.28**	.15**	−.10	.25**	.32**	.39**	.42**	.50**	.18**	.18**	.21**
WISC	Digit Span	.32**	.18**	.16**	.26**	.32**	1	.22**	−.09	.12*	−.11*	.16**	.13*	.23**	.37**	.33**	.11*	.10*	.11*
WISC	Coding	.26**	.22**	.16**	.15**	.24**	.22**	1	−.23**	.15**	−.011	.15**	.23**	.30**	.16**	.24**	.20**	.17**	.22**
CANTAB	ID/ED	−.34**	−.25**	−.19**	−.23**	−.28**	−.09	−.23**	1	−.21**	.09	−.23**	−.20**	−.20**	−.16**	−.26**	−.23**	−.19**	−.20**
CANTAB	Spatial Span	.15**	.22**	.07	<−.01	.15**	.12*	.15**	−.21**	1	−.24**	.44**	.45**	.13*	−.09	.13*	.47**	.48**	.50**
CANTAB	Spatial Working Memory	−.13*	−.14**	−.14**	.01	−.10	−.11*	−.01	.087	−.24**	1	−.21**	−.14**	−.07	−.03	−.13*	−.19**	−.19**	−.18**
NEPSY	Verbal Fluency	.17**	.11*	−.01	.11	.25**	.16**	.15**	−.23**	.44**	−.21**	1	.50**	.23**	.09	.28**	.36**	.42**	.45**
CVLT	Trials 1 – 5	.25**	.13*	.09	.20**	.32**	.13*	.23**	−.20**	.45**	−.14**	.50**	1	.16**	−.05	.16**	.55**	.54**	.60**
WJ	Calculation	.49**	.34**	.39**	.31**	.39**	.23**	.30**	−.20**	.13*	−.07	.23**	.16**	1	.50**	.54**	.13*	.04	.08
WJ	Letter – Word Identification	.46**	.17**	.33**	.39**	.42**	.37**	.16**	−.16**	−.09	−.03	.09	− .05	.50**	1	.64**	−.05	− .11*	−.07
WJ	Passage Comprehension	.55**	.29**	.36**	.43**	.50**	.33**	.24**	−.26**	.13*	−.13*	.28**	.16**	.54**	.64**	1	.16**	.10	.09
Purdue	Both Hands	.15**	.20**	.03	.06	.18**	.11*	.20**	−.23**	.47**	−.19**	.36**	.55**	.13*	−.05	.16**	1	.66**	.68**
Purdue	Left Hand	.15**	.20**	<.01	.06	.18**	.10*	.17**	−.19**	.48**	−.19**	.42**	.54**	.04	−.11*	.10	.66**	1	.70**
Purdue	Right Hand	.13*	.12*	−.02	.08	.21**	.11*	.22**	−.20**	.50**	−.18**	.45**	.60**	.08	−.07	.09	.68**	.70**	1

CANTAB = Cambridge Neuropsychological Test Automated Battery; CVLT = California Verbal Learning Test; FSIQ = Full Scale IQ; ID/ED Intra-Extradimensional Set-shifting; NEPSY = Developmental Neuropsychological Assessment; Purdue = Purdue Pegboard; WASI = Wechsler Abbreviated Scale of Intelligence; WJ = Woodcock-Johnson Tests of Cognitive Abilities.
*$p < .05$.
**$p < .01$.

Table 6. Distribution (percentages) of individuals in the NIHPD study with impairments on any of the eight neuropsychological measures (i.e., number of indicators of impairment) at three levels of functioning: 1, 1.5, 2 standard deviations below the mean by age group

	Group 1 (n = 95)			Group 2 (n = 156)			Group 3 (n = 119)			Total (n = 370)		
Indicators	1 *SD*	1.5 *SD*	2 *SD*	1 *SD*	1.5 *SD*	2 *SD*	1 *SD*	1.5 *SD*	2 *SD*	1 *SD*	1.5 *SD*	2 *SD*
0	47 (50)	81 (85)	87 (92)	75 (48)	113 (72)	143 (92)	49 (41)	87 (73)	107 (90)	171 (46)	281 (76)	337 (91)
1	25 (26)	11 (12)	7 (7)	44 (28)	32 (21)	12 (8)	36 (30)	27 (23)	12 (10)	105 (28)	70 (19)	31 (8)
2	16 (17)	3 (3)	1 (1)	25 (16)	9 (6)	1 (.6)	17 (14)	4 (3)	—	58 (16)	16 (4)	2 (.5)
3	5 (5)	—	—	6 (4)	1 (.6)	—	12 (10)	1 (.8)	—	23 (6)	2 (.5)	—
4	2 (2)	—	—	1 (.6)	1 (.6)	—	4 (3)	—	—	7 (2)	1 (.3)	—
5	—	—	—	4 (3)	—	—	1 (.8)	—	—	5 (1)	—	—
6	—	—	—	1 (.6)	—	—	—	—	—	1 (.3)	—	—

Note. Shading represents the number of individuals considered for the NPI rule as having impairments (1.5 *SD* below the mean) on two or more indicators.

indicators of neuropsychological impairment were broken down by level of education, NPI rates were significantly higher for the group with parental education of high school or less (11.1%) than those with college (4.4%) or graduate (3.5%) education. The college and graduate groups were therefore combined and a significant difference was also found when the "college or above" and "high school" groups were compared on number of indicators of impairment, $\chi^2(1,368) = 11.22$; $p = .001$. However, the magnitude of these associations was small, as denoted by weak spearman rank correlations (income $\rho = -.15$, education $\rho = -.13$). Nonetheless, an education-adjusted cutoff was explored in the current sample resulting in a −1.4 *SD* cutoff for college and −1.6 *SD* cutoff for high school.

NPI Rule Validation

Bootstrap validation

A bootstrap analysis with 1000 samples was performed to validate the rule. The bootstrap estimate of the proportion below the 1.5 *SD* criterion was 5.1% (the same as the observed estimate). The 95% bootstrap confidence interval for this proportion was 3.0 to 7.6%. This suggests that, had the NIHPD data been different (but sampled from the same population), the estimated proportion identified as impaired using our rule could have been as low as 3.0% or as high as 7.6%.

Internal and external clinical validation

The number of individuals with impairments on neuropsychological measures (indicators of impairment) is presented in Table 10 for each validation database (NIHPD/ Yeates et al., 2002) and for each version (10 variables/16 variables) as a function of the 1.5 *SD* below the mean cutoff. When the NPI rule (performance 1.5 *SD*s below the mean on two or more subtests) was applied to a more comprehensive version of the NIHPD data that included 8 domains and 10 variables (intellectual functioning removed, language, visuo-constructive, and academic functions added), the rule identified 6.2% of the total sample as impaired. The rate of impairment rose slightly to 7.8% in the 16 variable dataset; however, neither of these changes was significantly different from the proportion of individuals identified as impaired in the original rule (5.1%, 10 variable version $z = -.64$; $p = .52$; 16 variable version $z = -.86$; $p = .39$).

When the NPI rule was applied to a novel dataset of children with orthopedic controls and children with TBI (10 variables, 8 domains), the rule identified 17.7% and 25.5% of the samples as impaired, respectively. Correlations between variables included in the Yeates et al. (2002) study are presented in Table 11. When rates of impairment were compared across TBI severity levels, 16/45 (35.6%) of children with severe TBI were impaired according to our definition compared to 9/53 (17.0%) of those with

Table 7. Frequency of impairment in the NIHPD sample by subtest and domain

Domain	NIHPD subtest	Responses (n)	Subtest impaired n (%)	Domain impaired n (%)
Intellectual functioning	WASI FSIQ	370	1 (0.3)	1 (0.3)
Attention/Working memory	Digit Span	370	8 (2.2)	19 (5.2)
	CANTAB spatial span	368	11 (3.0)	
Memory	CVLT	361	22 (6.1)	22 (6.1)
Executive functions	CANTAB ID/ED	368	24 (6.5)	40 (12.7)
	NEPSY letter fluency	312	16 (5.1)	
Processing speed	Coding	368	13 (3.5)	13 (3.5)
Motor function	Purdue Pegboard	368	17 (4.6)	17 (4.6)

https://doi.org/10.1017/S1355617715000636 Published online by Cambridge University Press

Table 8. Holm-Bonferroni adjusted *p*-values for McNemar tests comparing the proportion of individuals impaired between neuropsycholgical tests

	WASI FSIQ	Digit Span	CANTAB spatial span	CVLT	CANTAB ID/ED	NEPSY letter fluency	Coding	Purdue Pegboard
WASI FSIQ	*N/A*	0.78	0.138	<0.001*	<0.001*	0.006*	0.01*	0.003*
Digit Span	—	*N/A*	1.000	0.138	0.147	1.000	1.000	1.000
CANTAB spatial span	—	—	*N/A*	1.000	0.780	1.000	1.000	1.000
CVLT	—	—	—	*N/A*	1.000	1.000	1.000	1.000
CANTAB ID/ED	—	—	—	—	*N/A*	1.000	1.000	1.000
NEPSY letter fluency	—	—	—	—	—	*N/A*	1.000	1.000
Coding	—	—	—	—	—	—	*N/A*	1.000
Purdue Pegboard	—	—	—	—	—	—	—	*N/A*

*Significant differences in the proportion of individuals impaired between tests.

mild/moderate injuries, a statistically significant difference (z = –2.10; $p = .04$). The rate of impairment in children with mild/moderate TBI was comparable to the orthopedic control sample (z = –.01; $p = .91$).

DISCUSSION

We derived a clinical case definition rule for defining neurocognitive impairment using a large representative U.S. sample of children 6 to 18 years who completed a comprehensive neuropsychological battery. After considering a variety of cutoff scores and number of indicators of impairment, we concluded that the best rule for identifying neurocognitive impairment in 5% of the population would require two or more performances 1.5 *SD*s below the mean when using a test battery covering six domains of functioning and eight subtests (IQ, attention/working memory, memory, executive functions, processing speed, fine motor skills).

The cutoff used for the rule derivation results in 95% of the population being defined as typically developing (therefore, a 5% rate of identifying cognitive impairment). While this cutoff is somewhat arbitrary, it was chosen to establish a relatively conservative NPI definition. Some may prefer a less conservative definition, but this would increase the identification rate of abnormal performances in presumed healthy children. This rule is in accord with the principles outlined by Brooks and Iverson (2012). That is, many individuals had low scores on at least one of the eight subtests and this proportion decreased when a more conservative cutoff was used. No individuals in the current sample had deficits on more than six of the eight subtests.

There was no clear pattern of subtests or domains that contributed more or less to the NPI rule. However, lower rates of impairment were found on the IQ estimate compared to measures of memory, executive functions, processing speed, and fine motor skills. While these differences are likely to be in part related to the fact that the level of intellectual functioning in the NIHPD sample was higher than expected norms, they may also reflect the relative insensitivity of IQ to variations in performance in normative groups and some clinical samples (e.g., traumatic brain injury, Crowe, Catroppa, Babl, & Anderson, 2012; Ewing-Cobbs, Barnes, & Fletcher, 2003).

Table 9. Number of indicators of impairment at 1.5 *SD* below the mean, by mothers' education level and income

Total number of indicators *versus* education and income						
	Total number of indicators (1.5 *SD*)					
	0	1	2	3	4	*p*-Value
Mother's education level (*n*, %)						.002
High school or less	34 (63.0)	14 (25.9)	3 (5.6)	2 (3.7)	1 (1.9)	
College	176 (76.9)	43 (18.8)	10 (4.4)	0 (0.0)	0 (0.0)	
Graduate	70 (82.4)	12 (14.1)	3 (3.5)	0 (0.0)	0 (0.0)	
Household income (*n*, %)						<.001
Zero to $35 000	19 (55.9)	9 (26.5)	4 (11.8)	1 (2.9)	1 (2.9)	
$35 001–$50 000	48 (70.6)	16 (23.5)	3 (4.4)	1 (1.5)	0 (0.0)	
$50 001–$75 000	70 (76.1)	18 (19.6)	4 (4.3)	0 (0.0)	0 (0.0)	
$75 001–$100 000	78 (85.7)	10 (11.0)	3 (3.3)	0 (0.0)	0 (0.0)	
$100 001–$150 000	63 (77.8)	16 (19.8)	2 (2.5)	0 (0.0)	0 (0.0)	

https://doi.org/10.1017/S1355617715000636 Published online by Cambridge University Press

Table 10. Testing the NPI: Distribution (percentages) of individuals in the NIHPD and Yeates et al. (2002) studies with impairments on the neuropsychological measures (i.e., number of indicators of impairment) at 1.5 standard deviations below the mean

	Dataset	NIHPD	NIHPD	Yeates et al.	Yeates et al.	Yeates et al.	Yeates et al.
	Version/group	Short	Extended	Orthopedic injury	TBI	TBI mild/moderate	TBI Severe
	No. of variables	10	16	10	10	10	10
	Domains	8	8	8	8	8	8
	Sample size (n)	370	370	62	98	53	45
Indicators	0	275 (74.3)	258 (69.7)	37 (59.7)	50 (51.0)	31 (58.5)	19 (42.2)
	1	72 (19.5)	83 (22.4)	14 (22.6)	23 (23.5)	13 (24.5)	10 (22.2)
	2	20 (5.4)	21 (5.7)	9 (14.5)	10 (10.2)	2 (3.8)	8 (17.8)
	3	2 (0.5)	5 (1.4)	0 (0.0)	5 (5.1)	1 (1.9)	4 (8.9)
	4	1 (0.3)	2 (0.5)	1 (1.6)	1 (1.0)	1 (1.9)	0 (0.0)
	5	—	0 (0.0)	1 (1.6)	4 (4.1)	2 (3.8)	2 (4.4)
	6	—	1 (0.3)	—	4 (4.1)	3 (5.7)	1 (2.2)
	7	—	—	—	0 (0.0)	0 (0.0)	0 (0.0)
	8	—	—	—	0 (0.0)	0 (0.0)	0 (0.0)
	9				1 (1.0)	0 (0.0)	1 (2.2)

Note. Shading represents the number of individuals considered for the NPI rule as having impairments (1.5 SD below the mean) on two or more indicators. TBI = traumatic brain injury.

The number of indicators of impairment was significantly associated with parental income and education level. In the NIHPD study, income was also found to be significantly associated with neuropsychological outcome. However, when its predictive ability was tested, the association was significant for broad indictors such as general academic achievement (e.g., reading comprehension and calculation) and IQ subtests, but specific neurocognitive measures were not significantly related to income (e.g., working memory, verbal recall) (Waber et al., 2007). There is literature to suggest that more low scores are found in individuals with fewer years of parent education (Schoenberg, Lange, Brickell, & Saklofske, 2007; van der Sluis, Willemse, Boomsma, & Posthuma, 2008). This means the impairment identification rate increases with fewer years of education in parents and decreases in those with parents with more education. Our NPI rule is validated by the expected association found between education level and indicators of impairment, a finding that is also consistent with recent literature describing the effects of socioeconomic status (SES) on brain and cognitive development, suggesting that SES affects neuropsychological functioning and that higher rates of impairment should be found in children from lower SES homes (see, Hackman, Farah, & Meaney, 2014). Despite this, the magnitude of the association between the number of indicators of impairment and education was weak. While those wishing to apply a more conservative definition of impairment that takes parental education into consideration may want to consider an education-adjusted rule, doing so does not result in particularly meaningful cutoff scores (i.e., 1.4 *SD* cutoff for college and −1.6 *SD* cutoff for high school level education) and is technically and clinically problematic.

The performance of the NPI rule was tested when domains and variables are increased. First, the rule was applied to an

Table 11. Pearson correlations between tests from the Yeates et al. (2002) database (orthopedic injury group)

Test Battery	Subtest/Variable	CPT-3 Omissions	CELF Recalling Sentences	CVLT Trials 1–5	CNT Efficiency	Grooved Pegboard	COWA Verbal Fluency	WISC Block Design	WISC Vocabulary	WJ Letter-Word Identification
CPT-3	Omissions	1	−.24**	−.34**	−.28**	.24**	−.22**	−.34**	−.19*	−.29**
CELF	Recalling Sentences	−.24**	1	.28**	.35**	−.08	.23**	.29**	.49**	.46**
CVLT	Trials 1−5	−.34**	.28**	1	.37**	−.05	.28**	.28**	.42**	.21**
CNT	Efficiency	−.28**	.35**	.37**	1	−.35**	.37**	.47**	.40**	.30**
Grooved Pegboard	Both hands	.24**	−.08	−.05	−.35**	1	−.23**	−.32**	−.22**	−.14
COWA	Verbal fluency	−.22**	.23**	.28**	.37**	−.23**	1	.32**	.39**	.40**
WISC	Block Design	−.34**	.29**	.28**	.47**	−.32**	.32**	1	.45**	.43**
WISC	Vocabulary	−.19*	.49**	.42**	.40**	−.22**	.39**	.45**	1	.57**
WJ	Letter-Word Identification	−.29**	.46**	.21**	.30**	−.14	.40**	.43**	.57**	1

CELF = Clinical Evaluation of Language Fundamentals; CNT = Contingency Naming Test ; CPT = Conners Performance Test; CVLT = California Verbal Learning Test; COWA = Controlled Oral Word Association Test; WISC = Wechsler Intelligence Scale for Children-III; WJ = Woodcock-Johnson Tests of Cognitive Abilities.
*$p < .05$.
**$p < .01$.

alternate version of the NIHPD dataset. Full scale IQ was removed because it was the only variable that had a very low proportion of impairment (only one individual) and because it consists of several different functions and is therefore more heavily weighted. Additional domains of functioning including language and visuo-spatial construction were added to verify the performance of the rule for a more comprehensive neuropsychological battery. Doing so resulted in a small increase in the rate of impairment identified (from 5.1% to 6.2%), suggesting that the rule is relatively robust to increases in the number of indicators. We expected that further increasing the number of subtests would result in an even greater number of deficits; however, a 16 variable version was associated with only a modest increase in the impairment rate (7.8%), which remained comparable to the range of impairment detected in the rule derivation when age groups were considered (3.0–7.2%). It is probable that raising the number of subtests further would result in additional increases in the number of deficits on the test battery (Brooks & Iverson, 2012). This should be considered when applying the rule to even longer test batteries. If the number of subtests or variables were considerably greater than 16, the NPI derivation approach outlined here could be applied to examine impairment distributions, provided adequate standardized norms or a normative group is available.

The NPI rule was also validated by applying it to novel datasets with comparable domains of functioning. The rule was applied to the performance of two groups of children, those with OI and those who sustained TBI (Yeates et al., 2002). As expected, doing so resulted in an increase in the rate of impairment in the OI group (10 variable NIHPD data set 6.2% *vs*. 10 variable OI dataset 17.7% rate of impairment). This change suggests that the NPI rule is sensitive to changes in sample characteristics, such as disparities in race and SES, and differences in cognitive functioning. Children with OI also typically present with greater pre-morbid intellectual, attentional and learning difficulties than children from the community, and as such are thought to represent a good match for children with brain injuries who have similar baseline characteristics (Babikian et al., 2011; Yeates, 2010).

Logically, the rate of impairment was even higher (25.5%) when the rule was applied to a clinical population of children with TBI, which included severe injuries known to produce neuropsychological deficits post-injury (Babikian & Asarnow, 2009; Beauchamp & Anderson, 2013), further demonstrating the sensitivity of the rule to changes in cognitive performance. It is noteworthy that although 17% of the children in the OI group were impaired according to the NPI rule, the majority of these children (82%) had impairments on only two indicators and only two individuals (18%) had impairments on four or five indicators. In contrast, the majority of the TBI group had more appreciable rates of impairment (60% impaired on three or more indicators). Of interest, children with mild/moderate injuries had a rate of impairment comparable to that of the orthopedic control group (17.0% and 17.7%, respectively), while the severely injured group had a 35.6% rate of impairment. This finding accords with a previous publication that showed equivalent rates of neuropsychological impairment in the OI and mild/moderate TBI groups over the first 4 years post-injury (Fay et al., 2009).

Although the NPI rule was derived based on the specific tests included in the NIHPD study and not on *a priori* considerations of test battery choice, they provide a reasonable semblance of the type of neuropsychological assessment carried out both in clinical and research settings, as they cover six neurocognitive domains and include eight measures in a streamlined battery. The measures used in the NIHPD study are for the most part standardized and well-validated tests that are known to neuropsychologists. Only the two subtests from the CANTAB were not standardized, but were from a well-validated cognitive test battery thought to be useful for identifying brain-behavior correlations (Waber et al., 2007). Domains that were not included in the original derivation, but are typical of neuropsychological assessment, include language, academic skills, and visuo-spatial functioning (Brooks & Iverson, 2012). These domains were, however, included in an alternate version of the data set, and the NPI rule was shown to perform similarly when these changes were made. Neuropsychologists often use questionnaires to obtain parental or self-reports on behavioral adjustment and social competence, but these were not included in the case definition because they are not direct measures of performance and are subject to social desirability and parental bias. Social cognition is an area of assessment that was also absent from the test battery, as there were no sociocognitive measures included at the time the NIHPD study was conducted. The availability of standardized, validated measures for the pediatric population is still limited in this domain, but is becoming increasingly relevant to both clinical and research assessments (Beauchamp & Anderson, 2010).

LIMITATIONS

Our rule derivation is somewhat limited by the sample population and databases used. First, the NIHPD study was conducted in the United States and may not be representative of the ethnic and socioeconomic distribution of other countries, although it is reported to be an adequate reflection of the 2000 U.S. census (Waber et al., 2007). Demographic details are reported and can be compared to local characteristics of other groups and clinical samples. Second, the level of intellectual functioning in the NIHPD study population was above the normative mean. Waber and colleagues (2007) suggest that the higher scores obtained in their sample are due to the fact that sources of morbidity were screened out by the exclusionary criteria. This is likely to be an adequate reflection of the level of functioning in typical comparison and control groups in research studies, who are also subject to medical and developmental exclusion criteria and tend to come from average to above average functioning families. Consistent with expectations, a higher rate of impairment was identified when the rule was applied to an orthopedically injured group of children who were not screened on the basis of pre-morbid cognitive difficulties and had a lower mean level of IQ, lower SES, and were more likely to be from minority groups. Third, a further limitation is that we were bound by the measures selected as part of the NIHPD and Yeates et al. (2002) studies. As a result, a combination of local and population norms were used in the rule derivation and validation because some tests were not standardized. Fourth, the rule's applicability to the performance of children under 6 years of age is unclear. When three age groups spanning 6 to 18 years were considered, the same definition of clinical impairment held true for all ages, with the exception that it was derived based on seven rather than eight measures in children 6–8 years of age. The rapid cognitive development that occurs in the preschool years and the substantial difference in neuropsychological tests used in younger age groups may require the derivation of a case definition rule specific to children under 6 years (Diamond, 2000; Isquith, Crawford, Espy, & Gioia, 2005; Lezak, 1995; Tsujimoto, 2008). Finally, the rule does not take into account pre-morbid levels of functioning, but this reflects the typical situation of clinical interpretations, as baseline data is seldom available other than third party report on academic performance (e.g., grades).

CONCLUSION

The NPI rule is based on psychometrically derived base rates of neurocognitive functioning in typically developing children and adolescents. The rule is expected to be useful to scientists and clinicians attempting to determine whether particular individuals or clinical populations demonstrate overall neuropsychological impairment when typical domains of neuropsychological functioning are assessed. Future studies should aim to prospectively validate the NPI rule in other normative populations and cross-validate in additional clinical groups. Other work could compare its utility in identifying neuropsychological impairment to other approaches, such as the Global Deficit Score method proposed by Carey and colleagues (2004). Further refinement of the rule may consider the addition of younger age groups, examination of the performance of the rule in terms of abnormal scatter of scores rather than number of low scores, exploration of supplementary domains such as somatosensory functioning and social cognition, and different ethnic, and socioeconomic distributions.

ACKNOWLEDGMENTS

This work was supported by the Canadian Institutes of Health Research (R.Z. 293380), (M.B. 231125). We would like to thank Dr. Gerry Taylor for his willingness to provide the external validation data and for facilitating access to it. B.L.B. reports that he receives royalties for a pediatric memory test (Sherman and Brooks, 2015, Child and Adolescent Memory Profile [ChAMP™], PAR Inc.), past research support from a computerized test publisher (CNS Vital Signs) through in-kind test credits, and book royalties from Oxford University Press. None of the other authors report any conflicts of interest. Data used in the preparation of this article were obtained from the Pediatric MRI Data Repository created by the NIH MRI Study of Normal Brain Development. This is a multi-site, longitudinal study of typically developing children, from ages newborn through young adulthood, conducted by the Brain Development Cooperative Group and supported by the National Institute of Child Health and Human Development, the National Institute on Drug Abuse, the National Institute of Mental Health, and the National Institute of Neurological Disorders and Stroke (Contract #s N01-HD02-3343, N01-MH9-0002, and N01-NS-9-2314, -2315, -2316, -2317, -2319, and -2320). A listing of the participating sites and a complete listing of the study investigators can be found at http://www.bic.mni.mcgill.ca/nihpd/info/participating_centers.html.

REFERENCES

Achenbach, T.M. (1991). *Integrative guide for the 1991 CBCL/4-18, YSR, and TRF profiles*. Burlington, VT: Department of Psychiatry, University of Vermont.

Anderson, V., Spencer-Smith, M., Leventer, R., Coleman, L., Anderson, P., Williams, J., … Jacobs, R. (2009). Childhood brain insult: Can age at insult help us predict outcome? *Brain*, *132*(1), 45–56. doi:10.1093/brain/awn293

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

Babikian, T., Satz, P., Zaucha, K., Light, R., Lewis, R.S., & Asarnow, R.F. (2011). The UCLA longitudinal study of neurocognitive outcomes following mild pediatric traumatic brain injury. *Journal of the International Neuropsychological Society*, *17*(5), 886–895. doi:10.1017/s1355617711000907

Beauchamp, M.H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. *Psychological Bulletin*, *136*(1), 39–64. doi:10.1037/a0017768

Beauchamp, M.H., & Anderson, V. (2013). Cognitive and psychopathological sequelae of pediatric traumatic brain injury. In O. Dulac, S. Di Mauro & M. Lassonde (Eds.), *Handbook of clinical pediatric neurology*, *Vol. 112*, (pp. 913–920). Amsterdam: Elsevier.

Binder, L.M., Iverson, G.L., & Brooks, B.L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology*, *24*(1), 31–46. doi:10.1093/arclin/acn001

Brooks, B.L. (2010). Seeing the forest for the trees: Prevalence of low scores on the Wechsler Intelligence Scale for Children, fourth edition (WISC-IV). *Psychological Assessment*, *22*(3), 650–656. doi:10.1037/a0019781

Brooks, B.L., & Iverson, G.L. (2012). Improving accuracy when identifying cognitive impairment in pediatric neuropsychological assessments. In E.M.S. Sherman & B.L. Brooks (Eds.), *Pediatric forensic neuropsychology* (pp. 66–88). New York, NY: Oxford University Press.

Brooks, B.L., Iverson, G.L., Feldman, H.H., & Holdnack, J.A. (2009). Minimizing misdiagnosis: Psychometric criteria for possible or probable memory impairment. *Dementia and Geriatric Cognitive Disorders*, *27*(5), 439–450. doi:10.1159/000215390

Brooks, B.L., Iverson, G.L., Koushik, N.S., Mazur-Mosiewicz, A., Horton, A., & Reynolds, C.R. (2013). Prevalence of low scores in children and adolescents on the test of verbal conceptualization and fluency. *Applied Neuropsychology; Child*, *2*(1), 70–77. doi:10.1080/21622965.2012.742792

Brooks, B.L., Iverson, G.L., Sherman, E.M., & Holdnack, J.A. (2009). Healthy children and adolescents obtain some low scores across a battery of memory tests. *Journal of the International*

Neuropsychological Society, *15*(4), 613–617. doi:10.1017/S1355617709090651

Brooks, B.L., Iverson, G.L., Sherman, E.M., & Roberge, M.-C. (2010). Identifying cognitive problems in children and adolescents with depression using computerized neuropsychological testing. *Applied Neuropsychology*, *17*(1), 37–43. doi:10.1080/09084280903526083

Brooks, B.L., Sherman, E.M., & Iverson, G.L. (2010). Healthy children get low scores too: Prevalence of low scores on the NEPSY-II in preschoolers, children, and adolescents. *Archives of Clinical Neuropsychology*, *25*(3), 182–190. doi:10.1093/arclin/acq005

Carey, C.L., Woods, S.P., Gonzalez, R., Conover, E., Marcotte, T.D., Grant, I., & Heaton, R.K. (2004). Predictive validity of global deficit scores in detecting neuropsychological impairment in HIV infection. *Journal of Clinical and Experimental Psychology*, *26*(3), 307–319. doi:10.1080/13803390490510031

Casey, B.J., Jones, R.M., & Hare, T.A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*(1), 111–126. doi:10.1196/annals.1440.010

Conklin, H.M., Salorio, C.F., & Slomine, B.S. (2008). Working memory performance following paediatric traumatic brain injury. *Brain Injury*, *22*(11), 847–857. doi:10.1080/02699050802403565

Crawford, J.R., Garthwaite, P.H., & Gault, C.B. (2007). Estimating the percentage of the population with abnormally low scores (or abnormally large score differences) on standardized neuropsychological test batteries: A generic method with applications. *Neuropsychology*, *21*(4), 419–430. doi:10.1037/0894-4105.21.4.419

Crowe, L.M., Catroppa, C., Babl, F.E., & Anderson, V. (2012). Intellectual, behavioral, and social outcomes of accidental traumatic brain injury in early childhood. *Pediatrics*, *129*(2), e262–e268. doi:10.1542/peds.2011-0438

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1994). *CVLT-C: California Verbal Learning Test*. San Antonio, TX: The Psychological Corporation.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. *Child Development*, *71*(1), 44–56.

Evans, A.C., Brain Development Cooperative Group. (2006). The NIH MRI study of normal brain development. *Neuroimage*, *30*(1), 184–202.

Ewing-Cobbs, L., Barnes, M.A., & Fletcher, J.M. (2003). Early brain injury in children: Development and reorganization of cognitive function. *Developmental Neuropsychology*, *24*(2-3), 669–704. doi:10.1080/87565641.2003.9651915

Fay, T.B., Yeates, K.O., Wade, S.L., Drotar, D., Stancin, T., & Taylor, H.G. (2009). Predicting longitudinal patterns of functional deficits in children with traumatic brain injury. *Neuropsychology*, *23*(3), 271–282. doi:10.1037/a0014936

Gathercole, S.E. (1998). The development of memory. *Journal of Child Psychology and Psychiatry*, *39*(1), 3–27.

Gazzaley, A., & Nobre, A.C. (2012). Top-down modulation: Bridging selective attention and working memory. *Trends in Cognitive Science*, *16*(2), 129–135. doi:10.1016/j.tics.2011.11.014

Hackman, D.A., Farah, M.J., & Meaney, M.J. (2014). Socioeconomic status and the brain: Mechanistic insights from human and animal research. *Nature Reviews Neuroscience*, *11*(9), 651–659. doi:10.1038/nrn2897

Hung, R., Carroll, L.J., Cancelliere, C., Cote, P., Rumney, P., Keightley, M., … Cassidy, J.D. (2014). Systematic review of the clinical course, natural history, and prognosis for pediatric mild traumatic brain injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Archives of Physical Medicine and Rehabilitation*, *95*(3 Suppl.), S174–S191. doi:10.1016/j.apmr.2013.08.301

Hurks, P., Hendriksen, J., Dek, J., & Kooij, A. (2013). Normal variability of children's scaled scores on subtests of the Dutch Wechsler Preschool and Primary Scale of Intelligence-Third Edition. *The Clinical Neuropsychologist*, *27*(6), 988–1003. doi:10.1080/13854046.2013.797502

Huttenlocher, P.R. (1979). Synaptic density in human frontal cortex - Developmental changes and effects of aging. *Brain Research*, *163*(2), 195–205.

Huttenlocher, P.R., & Dabholkar, A. (1997). Developmental anatomy of the prefrontal cortex. In N. Krasnegor, G. Reid, & P. Goldman-Rakic (Eds.), *Development of the prefrontal cortex: Evolution, neurobiology and behaviour* (pp. 69–83). Baltimore: Paul H. Brookes Publishing Company.

Isquith, P.K., Crawford, J.S., Espy, K.A., & Gioia, G.A. (2005). Assessment of executive function in preschool-aged children. *Mental Retardation and Developmental Disabilities Research Reviews*, *11*(3), 209–215. doi:10.1002/mrdd.20075

Kolb, B., Pellis, S., & Robinson, T.E. (2004). Plasticity and functions of the orbital frontal cortex. *Brain and Cognition*, *55*(1), 104–115. doi:10.1016/S0278-2626(03)00278-1

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY: A developmental neuropsychological assessment*. San Antonio, TX: The Psychological Corporation.

Levin, H.S., Li, X., McCauley, S.R., Hanten, G., Wilde, E.A., & Swank, P. (2013). Neuropsychological outcome of mTBI: A principal component analysis approach. *Journal of Neurotrauma*, *30*(8), 625–632. doi:10.1089/neu.2012.2627

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

Mather, N., & Woodcock, R.W. (2001). *Woodcock-Johnson III Tests of Achievement: Examiner's manual*. Rolling Meadows, IL: Riverside Publishing.

Rieger, B.P., Lewandowski, L.J., Callahan, J.M., Spenceley, L., Truckenmiller, A., Gathje, R., & Miller, L.A. (2013). A prospective study of symptoms and neurocognitive outcomes in youth with concussion vs orthopaedic injuries. *Brain Injury*, *27*(2), 169–178. doi:10.3109/02699052.2012.729290

Sady, M.D., Vaughan, C.G., & Gioia, G.A. (2014). Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. *Archives of Clinical Neuropsychology*, *29*(4), 348–363. doi:10.1093/arclin/acu014

Sahakian, B.J., Morris, R.G., Evenden, J.L., Heald, A., Levy, R., Philpot, M., & Robbins, T.W. (1988). A comparative study of visuospatial memory and learning in Alzheimer-type dementia and Parkinson's disease. *Brain*, *111*(3), 695–718.

Sahakian, B.J., & Owen, A.M. (1992). Computerized assessment in neuropsychiatry using CANTAB: Discussion paper. *Journal of the Royal Society of Medicine*, *85*(7), 399–402.

Schoenberg, M.R., Lange, R.T., Brickell, T.A., & Saklofske, D.H. (2007). Estimating premorbid general cognitive functioning for children and adolescents using the American Wechsler Intelligence Scale for Children-Fourth Edition: Demographic and current performance approaches. *Child Neurology*, *22*(4), 379–388.

Tiffin, J. (1968). *Purdue pegboard: Examiner manual*. Chicago: Science Research Associates.

Tiffin, J., & Asher, E.J. (1948). The Purdue Pegboard: Norms and studies of reliability and validity. *Journal of Applied Psychology*, *32*(3), 234–247.

Tsujimoto, S. (2008). The prefrontal cortex: Functional neural development during early childhood. *Neuroscientist*, *14*(4), 345–358. doi:10.1177/1073858408316002

van der Sluis, S., Willemsen, G., de Geus, E.J., Boomsma, D.I., & Posthuma, D. (2008). Gene-environment interaction in adults' IQ scores: Measures of past and present environment. *Behavioral Genetics*, *38*(4), 348–360. doi:10.1007/s10519-008-9212-5

Waber, D.P., De Moor, C., Forbes, P.W., Almli, C.R., Botteron, K.N., Leonard, G., … Group, B.D.C. (2007). The NIH MRI study of normal brain development: Performance of a population based sample of healthy children aged 6 to 18 years on a neuropsychological battery. *Journal of the International Neuropsychological Society*, *13*(5), 729–746. doi:10.1017/S1355617707070841

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. New York, NY: The Psychological Corporation.

Yeates, K.O. (2010). Mild traumatic brain injury and postconcussive symptoms in children and adolescents. *Journal of the International Neuropsychological Society*, *16*(6), 953–960. doi:10.1017/s1355617710000986

Yeates, K.O., Taylor, H.G., Wade, S., Drotar, D., Stancin, T., & Minich, N. (2002). A prospective study of short- and long-term neuropsychological outcomes after traumatic brain injury in children. *Neuropsychology*, *16*(4), 514–523.

Zemek, R., Osmond, M.H., & Barrowman, N., Pediatric Emergency Research Canada Concussion Team. (2013). Predicting and preventing postconcussive problems in paediatrics (5P) study: Protocol for a prospective multicentre clinical prediction rule derivation study in children with concussion. *British Medical Journal Open*, *3*(8), 1–10. doi:10.1136/bmjopen-2013-003550

Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (2002). *Preschool language scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.