SYSTEMS: School-years screening test for the evaluation of mental status

Robert. A. Ouvrier, MD
The Children's Hospital at Westmead

Julie Hendy
The Children's Hospital at Westmead

Laurel. J. Bornholt, PhD
University of Sydney

Fiona. H. Black, PhD (now Spencer)
The Children's Hospital at Westmead

Ouvrier, RA., Hendy, R., Bornholt, L., & Black, FH. (1999). The SYSTEMS - School-Years Screening Test for the Evaluation of Mental Status. Journal of Child Neurology, 14, 772 - 780.

Publishers/Journal Web Page
http://www.bcdecker.com/default.aspx
http://www.bcdecker.com/productDetail.aspx?BJID=69%20.

Keywords
SYSTEMS screening test, MMSE, test development, children's cognitive functioning

ABSTRACT
This paper introduces the School-Years Screening Test for Evaluation of Mental Status (SYSTEMS). It was designed to be used by neurologists, pediatricians, and other health professionals assessing children with suspected cognitive problems or changes. SYSTEMS was initially based on the adult Mini-Mental State Examination developed by Folstein, Folstein, and McHugh in 1975. SYSTEMS is a 7- to 12-minute, one-on-one interview test containing 46 items for use in children between 5 and 12 years of age. Although a full diagnosis cannot be made, the results do provide an indication of whether to send a child for further detailed cognitive assessment. The development of SYSTEMS comprised seven studies with a total of 1207 children involved from Sydney primary schools and neurology clinics of the New Children's Hospital, Westmead, New South Wales, Australia. All children were administered the SYSTEMS. Some of the children also were administered the Stanford-Binet Intelligence Test, 4th edition, or the Differential Ability Scales. Results showed that the SYSTEMS was internally consistent, unbiased by sex, socioeconomic indicators, or language groups; discriminated well by age; and strongly correlated (r = 0.88) with mental age. No significant differences in results obtained by two trained administrators were evident and no indication of apparent practice effect was found. The SYSTEMS was found to have desirable levels of sensitivity (83% and 92%), specificity (76% and 95%), and likelihood ratio for cognitive impairment (3.63 and 17.5) when compared with neurologic judgments and the Differential Ability Scales, respectively.

This paper presents and evaluates the School-Years Screening Test for the Evaluation of Mental Status (SYSTEMS). The test was designed to assess the cognitive state of a child when he or she first presents to a neurologist or pediatrician. Low scores on the SYSTEMS would suggest cognitive impairment or cognitive deterioration and would indicate the need for a more detailed cognitive assessment.

A quick and reliable screening test of higher mental function is of benefit to an initial neurologic examination. It can provide quick and worthwhile information about a child's cognitive state. Any such screening test needs to be brief, to have results available quickly, to be able to be administered in the office or at the bedside, and to have high sensitivity with respect to cognitive impairment. The MiniMental State Examination is one such test used with adult patients. It was developed by Folstein, Folstein, and McHugh in 1975 to evaluate psychiatric patients, but has been found most effective in the evaluation of dementia. 1,2,3 The MiniMental State Examination, with minor modifications at times, is used worldwide in adult neurology and psychiatry practice to document the level of cognitive function in a reproducible and simple way, with scores that vary systematically with age and education level.4,5 The test takes 5 to 10 minutes to administer and covers a range of cognitive functions including orientation, attention-calculation, memory@ and language. Decisions about whether a patient needs further cognitive assessment are supported by Mini-Mental State Examination scores below a given cut-off value. The Mini-Mental State Examination is highly reliable and valid as a cognitive teSt,6 and has only a small practice effect.' In an early study of adult medical inpatients, the Mini-Mental State Examination had a sensitivity of 87% and a specificity of 82% for impairment of cognitive function.3 A more recent (1992) study by Feher et al, however, determined a sensitivity of 69% and a specificity of 90%. A low score on the MiniMental State Examination does not constitute a diagnosis, and cannot be used for creating cognitive profiles,5 but it does indicate a need for further testing. 1,4

Although screening tests that assist in the decision about further cognitive testing are in common use for adults,1,5 there is no widely accepted screening test for children.7 Yet such a test would be of value to pediatric medicine.7 The changing nature of cognitive development in children makes it critical that any such test be devised with an appreciation of the developmental progression in children so as to allow the incorporation of age-appropriate cut-off values.

The need for such a brief screening test in childhood is set against a background of numerous, rather lengthy, cognitive and neurologic test batteries. Several tests have been designed for infants and preschool children. For example, the Bayley Infant Neurodevelopmental Screen, intended for infants 3- to 24-months old, is used to identify developmental delay or neurologic impairment. It contains 78 items and is completed in 10 minutes.8 The Denver Developmental Screening Test, Revised identifies developmental delay in children aged less than 6 years.8 Some tests are designed to screen for specific deficits. The Dallas Pre-School Screening is a 30-item test that identifies learning disabilities in 3to 6-year-olds.8 Many such tests are time consuming to administer (eg, Halstead-Reitan Neurological Battery for Older Children with 168 items, the Halstead Category Test for Younger Children with 80 items, and the Luria-Nebraska Neuropsychological Battery-Children's Revision with 149 items),' or need special training (eg, Griffiths Mental Development Scales and the Quick Neurological Screen).8

The Mental Measurements Yearbook identifies other quite lengthy test batteries used to identify cognitive dysfunction. One of the screening tests, Quick Neurological Screening Test-Revised, which contains 15 questions and requires 20 minutes to complete, is suitable for school-age children but covers domains other than cognition, such as motor development, spatial organization, perceptual skills, balance, attention, and rhythm.9 Another such test, the Brief Assessment Examination, was developed for the evaluation of severely handicapped children with subacute encephalitis and is therefore inappropriate for initial or regular clinical use.10

This paper is by no means a criticism of the existing screening tests used for the assessment of various neurologic deficits. Rather, it points out that the existing tests are either for very young children, test specific deficits, do not include tests of cognitive function, are lengthy, or require special training for administration. At present a screen for cognitive functioning in children aged 5 to 11 years is nonexistent. There does appear to be a need for a short initial screening test of cognitive function in school-age children, which has age-based norms and cut-off scores, is suitable for use in clinical settings, and can be administered by a variety of health professionals.

Ouvrier et all modified the Mini-Mental State Examination slightly, adapting it for children (Table 1). The adapted version (MMSE-VAR) takes 5 to 10 minutes to administer and incorporates the items from the Mini-Mental State Examination with several minor modifications. The findings of Ouvrier et al indicated that the MMSE-VAR could be used with children from the age of 4 to 12 years. Their results suggested that scores obtained (out of a maximum of 35) reach a plateau at a mental age of about 9 or 10 years, presumably because by that age the child's performance on the test corresponds with normal adult performance. For children older than 10 years, MMSE-VAR values below 27 suggested impaired cognitive functioning.7

Table 1. Summary of Measures
Mini-Mental State Examination
- Developed by Folstein, Folstein, and McHugh, 1975.
- Used to screen for dementia in the adult clinical population.
- Scored out of 30 with dichotomous and mixed scoring.
- Takes 5 to 10 minutes to administer.

Modified Mini-Mental State Examination
- Used by Ouvrier, Goldsmith, Ouvrier, and Williams, 1993.
- Applicable, but not totally accurate in a pediatric setting with children from the age of 4 years and older.
- Total score out of 35 including a maximum of five points for spelling *world* backward, five points for serial sevens subtractions, and three points for the command to take a piece of paper in your right hand, fold the paper in half, and put it on the floor.
- Takes 5 to 10 minutes to administer.

- Incorporates the items from Mini-Mental State Examination with several minor modifications. These include the use of a practice word (*cat*) for younger or retarded children in the backward spelling test. Older children were asked to spell *world* forward and then backward. Items to be memorized were repeated twice by the examiner before the patient was asked to repeat them. The intersecting pentagon was a simpler design than in the original Mini-Mental State Examination.

Pediatric Mini-Mental State Examination
- Extended version of the Modified Mini-Mental State Examination.
- Developed by Ouvrier, Hendy, Bornholt and Black.
- Contains 98 items from MMSE-VAR plus extra items.
- Dichotomous scoring for each item; that is, one point for a correct answer and zero points for an incorrect answer.
- Administered in approximately 15 minutes.

School-Years Screening Test for Evaluation of Mental Status
- Final version of the screening test developed by Ouvrier, Hendy, Bornholt, and Black.
- Considered appropriate for screening children.
- Contains 46 items scored dichotomously.
- Takes 7 to 12 minutes to administer and score.

The investigators identified certain deficiencies in the MMSE-VAR. In particular, the screening test appeared to be difficult for younger children. Also, information about the level of performance for normal children at differing ages was unavailable. The researchers suggested that the screen undergo rigorous validation and that cut-off values for each age group be established.7

The aim of the present research was to introduce and evaluate a new screening test that incorporates items suitable for children of varying ages and that conforms to appropriate psychometric methodology. To account for the complexities of developing a new test, seven studies were undertaken to develop the School-Years Screening Test for the Evaluation of Mental Status (SYSTEMS). Table 2 gives an overview of the studies, with more detailed numbers outlined in Table 3.

Table 2. Research Studies of the SYSTEMS Project

Study	*No. of Children*	*Sample*	*Assessment*
First school	614	Sydney primary school children	Pediatric Mini-Mental State Examination; 98 items; 42 subjects were administered the Stanford-Binet Intelligence Test, 4th edition
Second School	399	Sydney primary school children	School-Years Screening Test for the Evaluation of Mental Status; 46 items
First clinical	78	New patients for neurology clinics at the New Children's Hospital	Pediatric Mini-Mental State Examination; neurologic judgment of cognitive functioning
Second clinical	76	New and follow-up patients from the neurology clinics at the New Children's Hospital	SYSTEMS; Differential Ability Scales; neurologic judgment of cognitive functioning
Interrater reliability	69	Sydney primary school children	SYSTEMS; administered on two occasions by two different researchers at each occasion
Test-retest reliability	135	Sydney primary school children	SYSTEMS administered on two occasions at varying time intervals (2, 4, and 12 weeks)
Cultural language	20	Sydney primary school children from Arabic-and Vietnamese-speaking backgrounds	SYSTEMS administered twice, once in English and once in the background language (Arabic or Vietnamese)

Table 3. Details of the Studies

	Subject's Age, years							
	5	*6*	*7*	*8*	*9*	*10*	*11*	*Total Children Tested*
Study 1: First school study								
High SEIFA								
Boys	15	15	14	15	15	15	15	104
Girls	15	15	15	14	15	14	15	103
Medium SEIFA								
Boys	15	15	15	15	13	15	14	102
Girls	15	15	15	14	14	15	15	103
Low SEIFA								
Boys	15	15	13	13	15	14	14	99
Girls	15	15	15	15	15	14	14	103
Total	90	90	87	86	87	87	87	614
Stanford-Binet Intelligence Test	6	6	6	6	6	6	6	42
Study 2: Second school study								
Boys	31	37	33	27	26	29	21	204
Girls	27	37	31	33	18	24	25	195
Total	58	74	64	60	44	53	46	399
Study 3; First clinical study								
Boys	5	12	10	6	7	8	6	54
Girls	6	3	5	2	2	2	4	24
Total	11	15	15	8	9	10	10	78
Study 4: Second clinical study								
Boys	7	3	7	9	6	10	7	49
Girls	5	3	7	6	2	2	2	27
Total	12	6	14	15	8	12	9	76
Study 5: Interrater reliability study								
Researcher 1	4	7	3	6	4	5	5	34
Researcher 2	6	6	5	3	5	5	5	35
Total	10	13	8	9	9	10	10	69
Study 6: Test-retest reliability study								
Time A	30	34	35	36				135
Time B1 (2 weeks)	10	11	11	13				45
Time B2 (4 weeks)	10	13	13	12				48
Time B3 (12 weeks)	10	10	11	11				42
Study 7: Cultural language study								
Arabic								11
Vietnamese								9
Total								20

SEIFA = socioeconomic indices for areas

METHOD

Assessment Measures

The first assessment measure incorporated into the research was the Pediatric Mini-Mental State Examination. This measure was established by the present research team as an extended version of the MMSE-VAR, which was used by Ouvrier et al.7 It contained 98 items, took 15 to 20 minutes to complete, and consisted of the MMSE-VAR items plus extra questions chosen to cluster around the same themes used in the Mini-Mental State Examination and the MMSE-VAR (orientation, registration, attention and calculation, recall, language, repetition, commands, reading, writing, and copying). The extra items were chosen because they were clinically appropriate and relatively easy for younger children. Each item of the Pediatric

Mini-Mental State Examination was scored on a dichotomous scale with one point for a correct answer and zero points for an incorrect answer. All correct items were summed to give a final score.

The second assessment measure used in the research was the SYSTEMS. Items selected for inclusion in the SYSTEMS were taken from the Pediatric Mini-Mental State Examination based on results from the first school and clinical studies. All items in the Pediatric Mini-Mental State Examination were individually scrutinized and a decision was made regarding inclusion in the final test instrument. One or more of the following criteria were used as a basis for inclusion of an item: the items discriminated by age, contributed to the test as a whole, discriminated between the clinical and primary school groups of children, or were clinically important. In addition, a wide range of items was incorporated into the format of the SYSTEMS from the original Mini-Mental State Examination (such as serial sevens subtraction). Important criteria for inclusion were that items be unbiased by socioeconomic indicators or by sex. The SYSTEMS took 7 to 12 minutes to administer, with items again scored as one for correct and zero for incorrect. The summing of correct items provided a total score. The SYSTEMS was also translated into Arabic and Vietnamese.

The third measure was an abbreviated form of the StanfordBinet Intelligence Test, 4th edition. The Abbreviated Form of this test took between 30 and 50 minutes and included the core tests of vocabulary, bead memory, quantitative scale, memory for sentences, pattern analysis, and comprehension.11

The fourth measure to be used was a neurologic judgment given by a neurologist on all clinical participants. Neurologists indicated whether their patients' cognitive functioning was normal, equivocal, or impaired. Normal was defined as average or aboveaverage mental functioning. The decision might be based on diagnoses known not to compromise mental functioning, for example peripheral disorders, a history of average to good school performance, or documentation of investigative results. Equivocal judgments were defined as situations in which impairment in mental functioning was suspected but not certain. This could include children with borderline IQ (70 to 84), attention-deficit hyperactivity disorder (ADHD), Tourette syndrome, or seizure disorder, but where it was unclear whether their neurologic condition was having an impact on their mental functioning. Impaired mental functioning was defined as mental retardation with known or estimated IQ less than 70 or other cognitive defects such as ADHD, specific learning disorder, or memory or language dysfunction such that the child had a clear alteration in school performance.

The final assessment measure was the Differential Ability Scales, which took between 45 and 90 minutes to administer. The Differential Ability Scales are an intellectual assessment test used to provide a measure of children's cognitive abilities. It is used by psychologists to identify, analyze, and diagnose children with learning or other disabilities. Subtests that determine a General Cognitive Ability score include recall of designs, word definitions, pattern construction, matrices, similarities, and sequential and quantitative reasoning. Diagnostic subtests include recall of digits, recall of objects (immediate and delayed), and speed of information processing. Achievement tests include basic number skills, spelling, and word reading.12

Participants

Study 1: first School Study

These children (n = 614) from 21 government primary schools in the Metropolitan East region of Sydney, New South Wales, Australia, were administered the Pediatric Mini-Mental State Examination. Schools were selected on the basis of Australian Bureau of Statistics (1990) Socio-Economic Indices for Areas-Index of Education and Occupation.13 Boys and girls were selected from seven age groups of 5 to 11 years. All participants spoke fluent English, did not have any known neurologic conditions, and were not impaired in their communication skills. In addition 42 of these children were administered an abbreviated version of the Stanford-Binet Intelligence Test, 4th edition.11

Study 2: Second School Study

These children (n = 399) from six government primary schools selected from within a 20-kilometer radius of the New Children's Hospital, Westmead, New South Wales, Aust were administered the SYSTEMS. Schools were from various socioeconomic groups (as classified in Study 1).13

Study 3: First Clinical Study

These children were all new patients (n = 78,69% boys and 31%girls) from the ages of 5 to 11 years who presented at the neurology clinics of the New Children's Hospital. The children were administered the Pediatric Mini-Mental State Examination and neurologic judgments were obtained.

Study 4: Second Clinical Study

Children were all new patients (n = 76, boys and girls) as per the first clinical study. These children were administered the SYSTEMS and the Differential Ability Scales, and neurologic judgments were obtained.

Study 5: Interrater Reliability

Of the children in the second school study, 69 were tested twice by two different researchers (researcher I and researcher 2) in a counterbalanced order. The second administration of the SYSTEMS took place within one day of initial testing and each child was administered the test by both researchers.

Study 6. Test-Retest Reliability

A number of children (n = 135) from the second school study were tested on two occasions over various time intervals. The first testing session (Time A) established baseline competence for the SYSTEMS. The second testing session (Time B) took place at intervals varying for different groups of children as follows: B1 2 weeks later, B2 = 4 weeks later, and B3 = 12 weeks later.

Study 7: Cultural Language Study

The Arabic and Vietnamese versions of the SYSTEMS were administered to children (n = 20) from Arabic- and Vietnamese-speaking backgrounds by native speakers of each language. All children were also administered the English version of the SYSTEMS by a different researcher.

Procedures

For the first and second school group studies socioeconomic indicators (education and occupation) for the area of Sydney were grouped into three categories: low, medium, and high. Participating children were randomly selected from these three groups. With the permission of school principals and with parental consent, one-on-one testing took place during class time in rooms at the schools. The following instructions were given by the researcher: "Hi, my name is... I am here today to ask a few questions and do some other activities. I'd like you to answer as well as you can. If you can't answer any just let me know."

Standard testing procedures for the Pediatric Mini-Mental State Examination and the SYSTEMS (English, Arabic, and Vietnamese versions) demanded that the items be evenly paced, and that there be no feedback to the child about correct or incorrect responses. The intelligence test was administered to 42 randomly selected (within age groups) children following the administration of the Pediatric Mini-Mental State Examination. The interrater reliability, test-retest reliability, and cultural studies involved the administration of the SYSTEMS on at least two separate occasions.

The first and second clinical studies involved the pediatric neurologists and their patients at the New Children's Hospital. Permission was obtained in advance from the parents of patients presenting for consultation at the outpatient clinics or private consulting rooms. Consenting parents were asked that their child either arrive early, stay afterwards, or return on another day for testing conducted in an office close to the clinic according to standard protocol. The researcher saw the patient without knowledge of the diagnosis or clinical problem. The neurologist performed a standard consultation without knowledge of the information collected by the researcher and provided a judgment of cognitive functioning. For the second clinical study the SYSTEMS was administered first, followed by the Differential Ability Scales. Scoring was completed only after both tests had been fully administered.

RESULTS

Establishment of the Normative Group

Comparisons of results from the two school studies were completed to establish a normative group. The Pediatric Mini-Mental State Examination with 98 items used in the first school study (n = 614) contained the 46 items of the SYSTEMS. The participants were rescored for the SYSTEMS from the pool of 98 items. A comparison of the SYSTEMS results between the two studies did not indicate any significant differences between the SYSTEMS test scores by age group, as indicated in FIgure 1. Analysis of variance showed no significant difference between the means of the first and second school studies (f = 4.72, df = 1/1010). These results indicate that the results from the two school studies could be combined to form the normative data of the SYSTEMS, including 1013 children.

Significance of Sex and Socioeconomic Indicators

Results indicated that the test is unbiased by sex (boys versus girls; f^sub obs^ = 0.751 P > 0.05) or by socioeconomic indicators (f^sub obs^ = 0.16, P > 0.05).

Distribution of Scores by Age

It was expected that for each age group the distribution of the SYSTEMS score would vary somewhat, particularly with younger children. Figure 2 shows the boxplots for each age group. The lower boundary of the large box shows the lower 25th percentile and the upper boundary shows the upper 75th percentile. Fifty percent of the cases have values within the box. The center bar represents the median score for each age group. The outer lines indicate the range of scores (high to low). Any dots outside these areas are outliers. From the boxplots it can be seen that the 5 and 6 year olds, are quite similar; the 10 and 11 year olds are also simlar. As age increases the scores have a smaller distribution, with a ceiling at around 10 years.

Concurrent Validity

The IQ scores of 42 children (every 15th child in the first school study) were converted to mental age (mental age = chronologic age x IQ/100). This score was then compared with the score obtained by the same child on the SYSTEMS. FIgure 3 shows that the two scores were highly correlated (r = 0.88). The analysis indicates that the SYSTEMS reflects level of cognitive functioning.

Clinical Studies

Sensitivity and Specificity

The sensitivity and specificity of a test are in a trade-off relationship, in that either can be improved but only at the expense of the other. Sensitivity is the proportion of true positive cases identified by the test, whereas specificity is the proportion of true negative cases identified by the test. Positive predictive value is the probability of detecting disease if the test is capable of detecting disease (a higher score would indicate that the test does detect disease)." The likelihood ratio compares the probability of knowing from the test result that the patient was a true positive case (does have a cognitive problem) with the corresponding probability that the patient was without cognitive deficit.14 For the purposes of the current research the neurologists' own decision with regard to patients' cognitive functioning was used as the initial measure for diagnosis. However, certain theoretical constraints must be acknowledged. Altman suggests that to calculate the sensitivity and specificity of a test it is necessary to know the patient's true condition as well as the diagnosis and the results of the test in question. When evaluating a diagnostic test's accuracy in predicting the given diagnosis, Altman argues that "...we do not necessarily know that the diagnosis is always correct...".16 That is to say, the diagnosis might not reflect the patient's true condition. Thus, the determination of sensitivity and specificity gives an evaluation of the screening test's ability to "...predict the diagnosis rather than the patient's true disease status."" To determine sensitivity, specificity, positive predictive value, and the likelihood ratio, the cases of true positives and true negatives need to be defined.

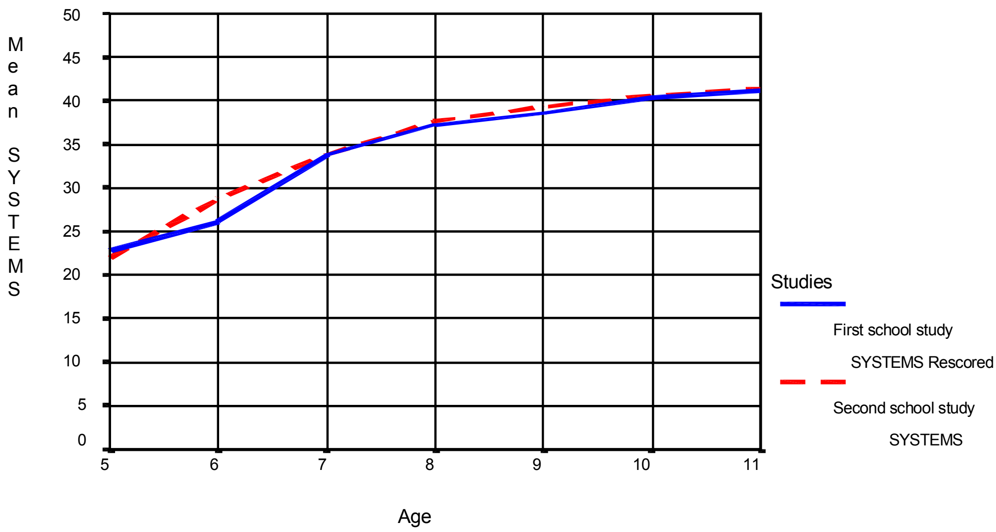

Figure 1.

Comparison of SYSTEMS results in the two school studies.

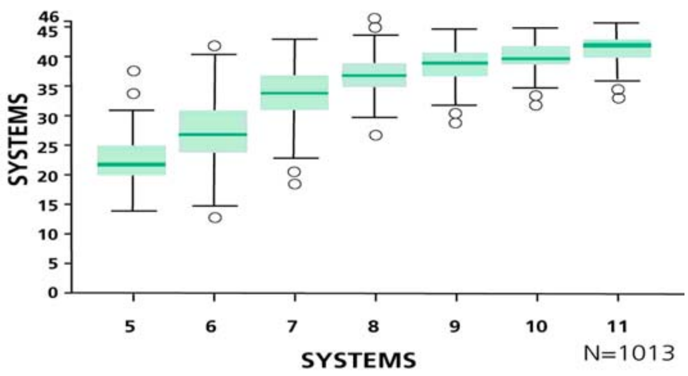

Figure 2.

Boxplot of SYSTEMS scores

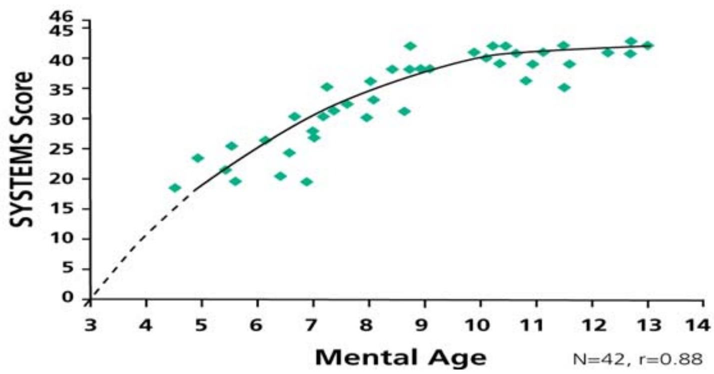

Figure 3.

Correlation between SYTEMS results and mental age

In the first and second clinical groups true positive describes children below the age-related cut-off value for the test who were also judged by the neurologist to be showing signs of impaired mental functioning. The term true negative describes children above the age-related cutoff value for the test who were judged by the neurologist not to be showing signs of impaired mental functioning. Equivocal neurological judgment cases (n = 45) were not included in the analysis.

SYSTEMS Versus MMSE-VAR and Pediatric MiniMental State Examination. Based on results from the first clinical study the MMSE-VAR revealed a sensitivity of 83%, specificity of 67%, positive predictive value of 52%, and likelihood ratio of 2.51. Pediatric Mini-Mental State Examination results were better with a sensitivity of 83%, specificity of 81%, positive predictive value of 65%, and likelihood ratio of 4.37. The SYSTEMS results revealed a sensitivity of 831/o, specificity of 76%, positive predictive value of 60%, and likelihood ratio of 3.46.

Table 4. SYSTEMS and Neurologic Judgment Results Correlation

	Neurologic Judgment, Aberrant	*Neurologic Judgment, Normal*	
SYSTEMS, aberrant	25	19	-
SYSTEMS, normal	4	61	-
Total no. of subjects tested	-	-	109

Table 5. SYSTEMS and Differential Ability Scales Results Correlation

	DAS, Aberrant Based on GCA Score	*DAS, Normal Based on GCA Score*	
SYSTEMS, aberrant	27	10	-
SYSTEMS, normal	3	36	-
Total no. of subjects tested	-	-	76

DAS = Differential Ability Scale; GCA = General Cognitive Ability

SYSTEMS Compared With Neurologic Judgment and the Differential Ability Scale. For the SYSTEMS cognitive screening test, two samples of children attending neurology clinics were studied to determine sensitivity and specificity, with respect to (1) the neurologic judgment of clinical impairment in the 109 nonequivocal cases from the first and second clinical studies and (2) the scores on the Differential Ability Scales, General Cognitive Ability, administered to all 76 children in the second clinical study, as the two "gold standards" (Tables 2 and 3).

SYSTEMS Versus Neurologic Judgment. Results showed that sensitivity = 86%, specificity = 76%, positive predictive value = 57%, and likelihood ratio = 3.63 (Table 4).

SYSTEMS Versus Psychologic Testing. In this situation true positive describes children below the age-related cutoff value for the SYSTEMS who were also classified as below the 25th percentile on Differential Ability Scales General Cognitive Ability scores. True negative describes children above the age-related cut-off value for the SYSTEMS with Differential Ability Scales General Cognitive Ability scores above age norms. Results indicate that sensitivity = 90%, specificity = 78%, positive predictive value = 730/6, and likelihood ratio = 4.14.

A closer examination of the Differential Ability Scales results for the above false positive cases (n = 10) enabled eight of these to be reclassified from Normal Differential Ability Scales/Aberrant SYSTEMS to Aberrant Differential Ability Scales/Aberrant SYSTEMS (that is, true positives). As well as analyzing the Differential Ability Scales General Cognitive Ability scores, the Differential Ability Scales diagnostic and academic tests were reviewed. These included tests of memory, processing speed, spatial skills, basic number skills, reading, and spelling. The eight cases were determined as being aberrant based on one or more of these tests indicating a clinical problem. The remaining two false positive cases not reclassified were either one point below the age-related cut-off value on the SYSTEMS or close to the 25th percentile on the Differential Ability Scales General Cognitive Ability scores (Table 5).

Table 6. SYSTEMS and Differential Ability Scales Results (Following Review)

	DAS, Aberrant GCA and Other Diagnostic and Achievement Tests	*DAS, Normal, GCA and Other Diagnostic and Achievement Tests*	
SYSTEMS, aberrant	35	2	-
SYSTEMS, normal	7	32	-
Total no. of subjects tested	-	-	76

DAS = Differential Ability Scale; GCA = General Cognitive Ability

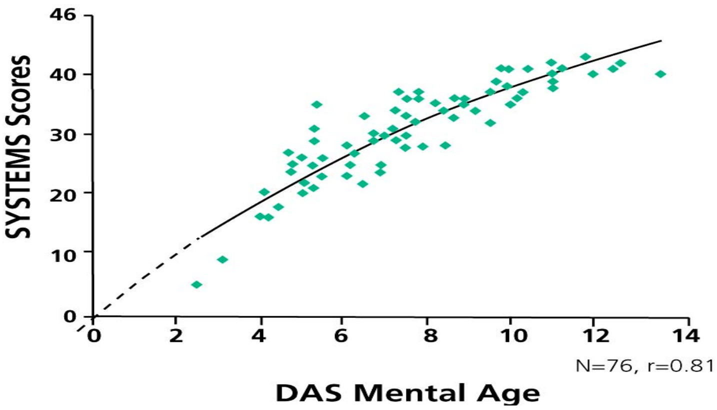

Figure 4.

Correlation between mental age (based on Differential Ability Scales, General Cognitive Ability) and SYSTEMS results.

A closer examination of the results for the above false negative cases (n = 3) showed that none of them was to be reclassified. The first case was a definite false negative. The second was a child who was impaired but also hyperlexic, which raised that child's SYSTEMS score into the normal range. The third was a child whose Differential Ability Scales General Cognitive Ability score was at the 23rd percentile, which is just below the normal range. Reclassified results indicate that sensitivity = 92%, specificity = 95%, positive predictive value = 95%, and likelihood ratio = 17.5 (Table 6).

Concurrent validity was established with the Differential Ability Scales for children in the second clinical study who presented at neurology clinics at the New Children's Hospital. The General Cognitive Ability scores for this group of children (n = 76) were converted to mental age (mental age = chronologic age x General Cognitive Ability/100). Figure 4 shows a strong correlation with SYSTEMS scores (r = 0.81) and Figure 5 shows a strong correlation between standardized SYSTEMS scores and General Cognitive Ability scores (r = 0.75).

Table 7. SYSTEMS Means and Standard Deviations for the Interrater Study

	Subjects' Age, years						
	5	*6*	*7*	*8*	*9*	*10*	*11*
Researcher 1							
Mean	22	32	37	36	39	42	43
Standard deviations	4	7	3	4	4	3	2
Researcher 2							
Mean	23	32	38	38	39	42	43
Standard deviations	3	6	3	4	4	2	2

Table 8. SYSTEMS Means and Standard Deviation Results for the Test-Retest Study

	Subjects' Age, years			
	5	*6*	*7*	*8*
Time A				
Mean	21	27	33	38
Standard deviations	5	5	6	3
Time B1 (2 weeks)				
Mean	20	30	33	39
Standard deviations	4	5	6	3
Time B2 (4 weeks)				
Mean	22	27	32	39
Standard deviations	3	4	6	2
Time B3 (12 weeks)				
Mean	25	30	35	40
Standard deviations	3	7	4	3

Interrater Reliability

Results reveal that the SYSTEMS administered by researcher 1 and researcher 2 were highly correlated ($r = 0.94$). These results were as expected, there being no bias with regard to two trained researchers administering the test. Table 7 shows the SYSTEMS means and standard deviations for each researcher by age group.

Test-Retest Reliability

Results indicated that the SYSTEM scores at Time A and Time B were highly correlated ($r = 0.94$). Table 8 shows the similarities between SYSTEMS mean and standard deviation scores at the various times. These results were as expected, there being no significant difference between scores with regard to the time interval between two administrations of the SYSTEMS, whether it be 2, 4, or 12 weeks following Time A.

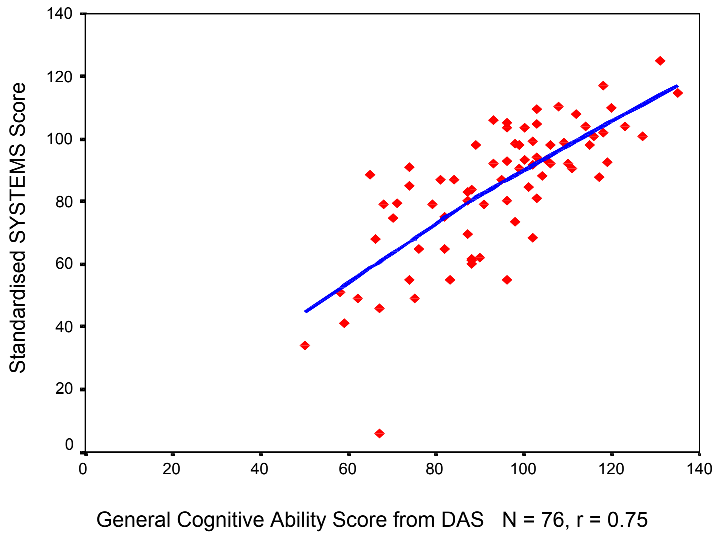

Figure 5.

Correlation between standardized SYSTEMS and General Cognitive Ability scores.

Cultural Language Study

A small study of community language groups suggests that the SYSTEMS can be administered in English rather than the children's home language. A sample of children who speak Vietnamese or Arabic as their first language at home, and have been at school in Australia for at least 2 years, were administered both the English-language SYSTEMS and a translated version of the SYSTEMS. Results showed a correlation (r = 0.81) between the scores on the SYSTEMS in English and in the other language (FIgure 6). For English SYSTEMS and Arabic SYSTEMS the correlation was r = 0.89 and for English SYSTEMS and Vietnamese SYSTEMS the correlation was r = 0.84.

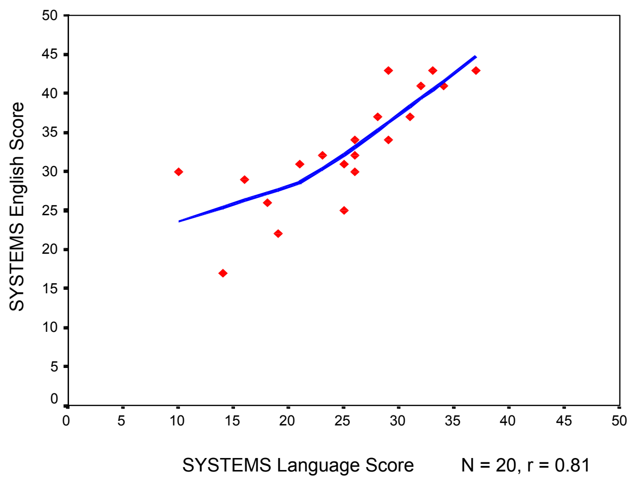

Figure 6.

English-language version of the SYSTEMS compared to the SYSTEMS in other languages.

DISCUSSION

Prior to the development of the SYSTEMS no rigorously validated, short, initial screening test of cognitive function in childhood was available. Previous studies have examined children's responses to a systematic short test of higher mental function, but neither study used a proper control population. Using a minor modification of the Mini-Mental State Examination, Ouvrier et al studied 117 neurology clinic patients and a group of 29 "quasi-controls".7 Besson and LaW studied 20 "clinical" and 79 "nonclinical" patients using another variation of the Mini-Mental State Examination, the Modified Mini-Mental State. The "nonclinical" patients were volunteered by their parents for participation in the latter study following recruitment through local schools, daycare centers, and well-baby clinics, as well as by community and media advertisements.17 Neither of these studies, therefore, provided psychometrically accurate normative data or extensive psychometric analyses for the administered test.

The present research, in contrast, systematically studied normative and clinical population samples using rigorous epidemiologic principles. The samples were carefully chosen with specific criteria. Children in the school studies were selected in each of seven age groups with relatively equal numbers of boys and girls. Children tested were drawn from schools in low, medium, and high socioeconomic status areas. The mean of the Socio-Economic Index for the whole school group (1068) was similar to the mean of the Australian population (1000).12 The only testing exclusions were for known neurologic or communication disorders (including nonfluent English, unless the children were involved in the Arabic and Vietnamese cultural language study). Children in the clinical studies were selected from neurology clinics of the New Children's Hospital. Special conditions ensured that the researcher and neurologist were blind to each other's results until all testing and judgments were completed.

The research demonstrated that SYSTEMS scores from the two school studies could be combined. The results were used as a normative sample of 1013 children aged 5 to 11 years for the SYSTEMS. An important aspect of this research was that along with the establishment of a normative sample the test was found to be unbiased by sex and socioeconomic indicators for geographic areas. There was a high correlation coefficient of 0.88 for the relationship between the SYSTEMS score and mental age in 42 randomly selected school children. In addition the SYSTEMS was subjected to further rigorous psychometrics and was found to be highly resilient. Acceptable levels of interrater and test-retest reliabilities, along with concurrent validity, were established. High levels of sensitivity, specificity, positive predictive value, and likelihood ratio for impairment, based on comparisons with the SYSTEMS and neurologic judgment plus Differential Ability Scales scores, were determined in clinical studies. In the first clinical study, where the "gold standard" was the neurologist's judgment, the sensitivity was shown to be identical (83%) for the MMSE-VAR, Pediatric Mini-Mental State Examination, and the SYSTEMS. Specificities were 67%, 81%, and 76%, respectively. The positive predictive values were 52% 65%, and 60%, respectively, and the likelihood ratios were 2.51, 4.37, and 3.46. These values clearly show the superiority of the SYSTEMS over the Mini-Mental State Examination. As expected, the Pediatric Mini-Mental State Examination, with 98 items (including all of the Mini-Mental State Examination and SYSTEMS items) was superior to the SYSTEMS but at the cost of extra administration time. Since the aim of the project was to create a brief and well-researched test of cognitive function, the SYSTEMS, taking 7 to 12 minutes to administer and having extensive psychometric results available, has an obvious advantage over any other form of the test (such as the MMSE-VAR and the Pediatric Mini-Mental State Examination).

Finally, SYSTEMS has been shown to have high levels of concurrent validity for children with less than average intellectual functioning determined by Differential Ability Scale resubs ($r = 0.83$, $P < .001$).18

CONCLUSION

The primary outcome of this research is the development of a cognitive screening test for children aged 4 to 11 years. It builds on previous research on the Mini-Mental State Examination in adults.1,3-5,19 and children.7 The SYSTEMS includes age-appropriate items and is based on a rigorous psychometric approach to test development with desirable levels of specificity, sensitivity, positive predictive value, and likelihood ratio for clinically diagnosed impairments. The SYSTEMS has high levels of reliability and reproducibility when administered serially by two trained researchers and over short to medium periods of time. The final version of the SYSTEMS is a valuable clinical tool to assist in the decision about the need for further cognitive assessment. It is also useful for the serial observation of cognitive functioning over short or long intervals and for the detection of specific neurologic deficits, such as aphasia and memory defects. It is, furthermore, a useful way for trainees in pediatrics and neurology to gain insights into the development of the cognitive processes of children.

Acknowledgment

This work was funded by the New Children's Hospital and the National Health and Medical Research Council. Much appreciation is extended to Heidi Lyneham for her enthusiasm and hard work while employed as a research assistant. The authors thank the primary school children from Sydney and their parents, and the teachers, staff, and principals of the schools involved. Gratitude is extended to neurologists at the New Children's Hospital (Dr Jayne Antony, Dr Paddy Grattan-Smith, Dr Elizabeth Fagan, A/Prof Peter Procopis, and Dr Kathryn North) and the children and parents who attended the clinics. Gratitude is extended to Julie Quinn, Dr Helga Rowe (ACER), Dr George Cooney (Macquarie University), Katrina Williams (NCH), Dr John Knight (NCH), Sarah Ouvrier, Lesley Russell, Sonia Hlavacek, and Katherine Coote, all of whom provided suggestions, assistance, and support.

References

1. Folstein M, Anthony JC, Parhad 1, et al: The meaning of cognitive impairment in the elderly. J Am Geriatr Soc 1985;33:228-235.
2. Katzman R, Brown T, Fuld P, et al: Validation of a short orientation-memory-concentration test of cognitive impairment. Am J Psychiatry 1983;140:734-739.
3. Folstein MF, Folstein SE, McHugh PR: Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975; 12:189-198.
4. Crum RM, Anthony JC, Bassett SS, Folstein MF: Population-based norms for the Mini-Mental State Examination by age and education level. JAMA 1993;269:2420-2421.
5. Feher EP, Mahurin RK, Doody RS, et al: Establishing the limits of the Mini-Mental State: Examination of 'subtests'. Arch Neurol 1992;49:87-92.
6. Dick JPR, Guiloff RJ, Steward A, et al: Mini-Mental State Examination in neurological patients. JNeurol Neurosurg Psychiatry 1984;47:496-499.
7. Ouvrier RA, Goldsmith RF, Ouvrier S, Williams IC: The value of the Mini-Mental State Examination in childhood. A preliminary study. J Child Neurol 1993;8:145-148.
8. Sattler JM:- Assessment of Children. San Diego, California, Sattler, 1988.
9. Rosenthal A: An evaluation of the Quick Neurological Screening Test. (QNST). Int J Neurosci 1983;21:85-95.
10. Nester MJ: Use of a brief assessment examination in a study of subacute sclerosing panencephalitis. J Child Neurol 1996;11: 173-180.
11. Thorndike RL, Hagen EP, Sattler JM: Stanford-Binet Intelligence Scale, 4th ed. Chicago, DLM Teaching Resources-Riverside, 1986.
12. Elliott CD: Differential Ability Scales: Administration and Scoring Manual. New York, Psychological Corp/Harcourt Brace Jovanovich, 1990.
13. Australian Bureau of Statistics: Information Paper. Socio-Economic Indices for Areas. Canberra, AGPS, 1990.
14. Ades AE: Evaluating screening tests and screening programmes. Arch Dis Child 1990;65:792-795.
15. Sackett DL, Richardson WS, Rosenberg W, Hayes RB: EvidenceBased Medicine: How to Practice & Teach EBM. New York, Churchill Livingstone, 1997.
16. Altman DG: Practical Statistics for Medical Research. Melbourne, Chapman & Hall, 1991.
17. Besson PS, LaW EE: Use of the Modified Mini-Mental State Examination with children. J Child Neurol 1997; 12:455-460.
18. Russell L: A Clinical Study of the Concurrent Validity of a New Cognitive Screening Test for Children, research diss. Sydney, Faculty of Science, University of Sydney, 1998.
19. Bleeker ML, Bolla-Wilson K, Kawas C, Agnew J: Age-specific norms of the Mini-Mental State Exam. Neurology 1988;38: 1565-1568.