THE UNIVERSITY OF ALBERTA

A VALIDATIONAL INVESTIGATION OF THE LURIA - NEBRASKA NEUROPSYCHOLOGICAL BATTERY FOR USE WITH ADOLESCENTS

by

NORMAN E. BRODIE

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

IN SCHOOL PSYCHOLOGY

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY

EDMONTON, ALBERTA

SPRING, 1988

ISBN 0-315-42764-7

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: NORMAN E. BRODIE

TITLE OF THESIS: A VALIDATIONAL INVESTIGATION OF THE LURIA - NEBRASKA NEUROPSYCHOLOGICAL BATTERY FOR USE WITH ADOLESCENTS

DEGREE FOR WHICH THESIS WAS PRESENTED: DOCTOR OF PHILOSOPHY

YEAR THIS DEGREE GRANTED: 1988

14532 - 26 Street
Edmonton, Alberta
T5Y 1W7

DATED: April 22, 1988

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled A VALIDATIONAL INVESTIGATION OF THE LURIA - NEBRASKA NEUROPSYCHOLOGICAL BATTERY FOR USE WITH ADOLESCENTS submitted by Norman E. Brodie in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

....
Supervisor

E. E. Fox

R. G. McIntire

R H Short

J G Paterson

M E J Orn
....
External Examiner

Date April 22, 1988

Dedicated to

Leslea,

Mom and Dad,

who have helped to

make it possible.

ABSTRACT

The study investigated the reliability, validity and utility of the Luria-Nebraska Neuropsychological Battery (LNNB) as an instrument for the adolescent age range. The LNNB was administered to a group of adolescents along with the WISC-R and WRAT. The LNNB literature review indicated that reliability and validity studies have been conducted with adult groups, but have not been performed on the adolescent age range.

The initial investigations explored the reliability and construct validity of the LNNB clinical scales. An abbreviated scale with equal reliability and improved item to scale consistency was identified. Factor analyses were conducted on both batteries to investigate the dimensional nature of the test. Studies of the concurrent validity were conducted for both the LNNB and the shortened format LNNB-A, with investigations of correlations with the WISC-R and WRAT, application of multiple regression analyses to determine the ability to predict WISC-R I.Q. scores and WRAT grade scores, and through use of linear discriminant analyses to determine the ability to discriminate between a priori defined learning ability groups. Cluster analyses using the clinical scale scores were conducted to explore the utility of both batteries to identify educationally meaningful subgroups.

The results of the study provided partial support for the use of the LNNB in the adolescent population, with an adequate level of reliability and robust correlations with intellectual and

educational tests. The content and construct validities of the battery were criticized, with the battery showing an excessive overlap between scales and an inadequate coverage of nonverbal cognitive abilities. The battery was found capable of discriminating between a priori defined subgroups of normal and delayed learners, but cluster analysis procedures failed to generate subgroups that differed in type of deficit, with separation only by level of impairment. The LNNB is judged to have potential for the assessment of educational disorders, but some expansion of its dimensional coverage is required to allow for practically meaningful subtyping of learning disabilities.

Discussion of the results addressed theoretical and applied implications, as well as the limitations of the current study and directions for future research.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

TABLE | PAGE

LIST OF FIGURES

CHAPTER I

INTRODUCTION

One of the most pressing and demanding problems facing the practicing school psychologist is the accurate and meaningful assessment of children and adolescents with severe learning disorders or disabilities. It has been estimated that at least 10 to 15 percent of children show seriously deficient academic attainment (Gaddes, 1976; Rutter, 1978) and in about half of these children, there are objective behavioural indicators of some neurological damage or dysfunction (Myklebust & Boshes, 1969). Longitudinal studies have demonstrated that children with such specific learning disabilities tend to show an enduring pattern of deficient scholastic performance, even with intensive remedial training (Trites & Fiedorowicz, 1976). It is also widely recognized that the ongoing experience of such frustration and failure leads to adverse emotional reactions and deficiencies in the student's self-concept (Bloom, 1976) and the learning disabled population is statistically more "at risk" to develop significant emotional and/or behavioural disturbances, particularly in adolescence (Rourke & Fisk, 1981).

While the educational, emotional and behavioural impact of learning disabilities is easily seen, the exact causation and essential nature of these failures to achieve remain more obscure, despite the extensive research that has been done to date. Theories regarding the etiology of learning disabilities abound, and represent radically diverse opinions on the subject (c.f., Learner, 1974). Similarly, theories of remedial instruction also present very diverse views, but with generally disappointing results on outcome studies.

What is known is that the learning disabled student experiences severe handicaps in his/her attempts to master the standard academic curriculum, despite at least average overall intellectual capabilities, no significant peripheral sensory or motor impairments, and adequate exposure to scholastic experiences. The most commonly accepted explanation for the stable learning deficiencies is that of some form of disordered neuropsychological functioning that results in selective central processing deficits, while sparing many of the other intellectual capabilities (c.f., Hynd & Obrzut, 1981). The adoption of such a definition or conceptualization of learning disabilities would imply the need to be able to identify and specify the presence and nature of the underlying disturbance in neuropsychological functioning, and would suggest that such a neuropsychological description should be valuable in specifying the nature of the learning disorder and in designing a remedial educational programme.

A major challenge confronting the school psychologist is therefore the need to make decisions about such neuropsychological dysfunctions, especially in view of limitations in the availability of specialized neuropsychological consultation services.

Neuropsychology is a relatively new discipline in the applied field of psychology, and the assessment techniques and methods of interpretation used in making a neuropsychological diagnosis are skills that are not widely represented among school psychologists at this time. A major limiting factor in this connexion has been the costly nature of neuropsychological assessment in the past, with the most frequently used assessment battery (the Halstead - Reitan Neuropsychological Battery) requiring the purchase of very expensive and bulky equipment that is not easily portable, and the time required for an individual evaluation (up to six to eight hours of test administration time) has also weighed against its routine application to the school population.

In view of the implicit need for attention to neuropsychological functioning contained in the modern definitions of learning disability and the difficulties in applying the Halstead - Reitan approach to neuropsychological assessment to the school population on a routine basis, a need is apparent for an alternative approach to neuropsychological evaluation that is more time and cost efficient. A recent addition to the field of clinical neuropsychology has been the publication of the Luria - Nebraska Neuropsychological Battery (Golden, Hammeke, & Purisch, 1980) which is a comprehensive assessment instrument that is relatively inexpensive, portable, and can be administered in most cases in two to two and a half hours. This battery is applicable to both adults and adolescents (starting from age 13) and has reportedly high reliability and validity when used for assessment of the effects of neurological damage and disorders (Golden, Purisch, and Hammeke, 1985). Its application to the adolescent population and specifically to the identification and description of learning problems in the adolescent age range has, however, not yet been dealt with extensively in the literature, and its utility for identification and specification of learning disability subtypes is not known.

This study was approached as a multi-stage investigation of the Luria - Nebraska Neuropsychological Battery in the context of its applicability to identifying educational problems in an adolescent population. The initial stage of this study was to investigate the reliability and construct validity of the Luria - Nebraska Neuropsychological Battery (LNNB) on a population of adolescents aged 13 to 16. This was accomplished through an evaluation of the internal consistency (Alpha coefficients) of the clinical scales of the battery, and a study of the correlations between the individual items of the LNNB and the clinical scale scores, to identify the extent to which the items were correlating specifically with their assigned scale.

Based on the observations of the initial reliability evaluation and the investigation of the degree of item to scale consistency for the LNNB clinical scales, it was decided to pursue the development of an abbreviated battery that would improve the construct validity and consistency of the clinical scales, and a shortened form of the

LNNB was created by elimination of selected items of poor discriminatory power or which showed poor item to scale correlation consistency.

A second stage was to investigate the factorial or dimensional nature of the LNNB, to determine more specifically what is the range of functions or abilities tapped by this battery, and to aid in interpreting the profile patterns or configurations that emerge in the clinical profiles. A factor analysis of the LNNB clinical scale scores was conducted and analyzed.

A third stage of the study was to investigate the concurrent validity of the LNNB clinical scales of Intellectual Processes and academic functioning by correlating these scales with the Wechsler Intelligence Scale for Children - Revised (WISC-R) and the Wide Range Achievement Test (WRAT), and by use of multiple regression analysis to identify the clinical scales of the LNNB that are maximally effective in predicting academic achievement on the WRAT and intellectual functioning on the WISC-R.

A fourth stage of the investigation was to explore the ability of the LNNB to discriminate between subgroups of students that display distinct patterns of academic performance on the WRAT (specific arithmetic deficit, global academic delay, normal academic performance). A linear discriminant analysis was applied to the LNNB data to determine its power or utility in discriminating between such subgroups defined a priori on the basis of the WRAT academic achievement scores.

A final stage of the investigation was to utilize the LNNB data to derive sub-groups by application of cluster analytic methodology, to determine the presence and nature of subtypes of learning delayed students. A major objective of this stage was to investigate whether the LNNB would be able to generate homogeneous sub-groups of learning disabled students that had apparent educational significance. Analysis of the differences between the derived sub-groups on measures of intellectual functioning (WISC-R) and academic attainment (WRAT) were conducted using multivariate analysis of variance procedures.

Discussion of the results addressed both theoretical and applied implications, as well as limitations of the present study and directions for future research.

CHAPTER II

SELECTIVE REVIEW OF THE LITERATURE

Review of the Literature on Learning Disabilities

Nature of the Problem

Learning disabilities have been described and defined for many years, with a variety of causative or etiological factors being implied by the definitions (Learner, 1974). One early definition of learning disability was offered by Kirk and Bateman (1962, p. 73):

> "A learning disability refers to a retardation, disorder, or delayed development in one or more of the processes of speech, language, reading, writing, arithmetic, or other school subjects resulting from a psychological handicap caused by a possible cerebral dysfunction and/or emotional retardation, sensory deprivation, or cultural or instructional factors."

One primary feature of this definition of learning disabilities is its rather broadly inclusive nature. This definition would include within its classification system almost all children with significant delays in academic attainment that could be attributed to a psychological handicap, regardless of the etiology of the handicap. Other definitions of specific learning disabilities have since been advanced that offered more explicit inclusion and exclusion criteria, often by reference to such factors as standards of academic deficiency, needs for special educational programming, or uneven patterns of psychological development:

> "Children with developmental imbalances are those who reveal a developmental disparity in psychological processes related to education of such a degree (often four years or more) as to require the instructional programming of developmental tasks appropriate to the nature and level of the deviant

developmental processes." (Gallagher, 1966, p. 28)

The central feature of many definitions is the attention given to the discrepancy between the child's estimated intellectual abilities or learning potential and the actual level of academic achievement. Kirk and Bateman (1965) addressed this factor in their description of learning disabled children, who:

> "manifest an educationally significant discrepancy between their estimated intellectual potential and actual level of performance related to basic disorders in the learning processes, which may or may not be accompanied by demonstrable central nervous system dysfunction, and which are not secondary to generalized mental retardation, educational or cultural deprivation, severe emotional disturbance, or sensory loss." (p. 220)

In the Kirk and Bateman definition the role of neurological dysfunctions in the genesis of learning disorders is acknowledged as present in some cases, but the ability to demonstrate evidence of the dysfunction is not held to be critical to inclusion in the learning disabled category. The actual cause(s) of the learning disorder are not addressed in this definition. A number of other authors have been more explicit in identifying the central role of neuropsychological dysfunctions in the genesis of specific learning disabilities:

> "we refer to children as having a psychoneurological learning disability, meaning that behavior has been disturbed as a result of a dysfunction of the brain and that the problem is one of altered processes, not of a generalized incapacity to learn." (Johnson & Myklebust, 1967 p. 8)

These same authors also emphasized the factor of uneven patterns of skills or abilities as a central feature of learning disabilities, such that some neurologically related capabilities are affected, while others are spared:

> "In those having a psychoneurological learning disability, it is the fact of adequate motor ability, average to high intelligence, adequate hearing and vision, and emotional adjustment together with a deficiency in learning that constitutes the basis for homogeneity." (Johnson & Myklebust, 1967 p. 9)

On the basis of a review of the various definitions of learning disabilities noted above, it is noted that they describe the presence of a more or less restricted area of learning handicap that is not caused by a generalized intellectual retardation or peripheral or environmental factors, and that most emphasize the importance of uneven patterns of development and/or the presence (either demonstrated or inferred) of neurological dysfunction(s). It is also apparent that the various definitions allow for the identification of a rather wide array of specific learning disorders, and that such disorders may manifest themselves in a variety of academic areas. It is evident that while neuropsychological dysfunctions are considered by many to be a central cause of learning disabilities, the exact nature of such dysfunctions as well as the cause of these deficits is subject to considerable inter-individual variability. This is consistent with recent research reviews in the area of learning disabilities, which have also emphasized the role of multiple factors or causes in the genesis of such disorders, and interactional nature of reading, language and neuropsychological deficits (c.f., Doehring, Trites, Patel & Fiedorowizc, 1981).

The multifaceted nature of learning disabilities and the manner in which they may be expressed in widely differing manners in individuals has recently been summarized by the Alberta Education Advisory Committee on Learning Disabilities, who have offered the following definition:

> "Learning Disabilities is a generic term that refers to a heterogeneous group of disorders that may have as its basis an identifiable or inferred central nervous system dysfunction. This group of disorders may be manifested by difficulties in one or more processes such as attention and concentration, perception, coordination, both social competence and emotional development, memory, reasoning, organization and planning resulting in deficits in one or more areas such as communication, reading, spelling, writing and mathematics.
>
> Learning disabilities affect individuals, with average or above average learning potential, in such a manner that the individual's unique learning characteristics require accommodation in and modification to the instructional process and the learning environment.
>
> Learning disabilities are not due primarily to visual, hearing or motor impairments, to mental retardation, emotional disturbance or environmental disadvantage." (AEAC, 1984)

Subgroups in Learning Disabilities

Considering the broad range of possible areas of handicap subsumed under the rubric of learning disabilities, and the tremendous range of individual differences or variance in the manner of expression of learning deficits implied by the AEAC definition, it is apparent that the meaningful assessment of the learning disabled child should include some means of accurately determining his/her current status on more than a cursory measure of intelligence and global estimate of academic attainment. To provide useful information that would assist educators in devising appropriate remedial strategies it is necessary to identify with some specificity the nature of the underlying psychological deficiency that is producing the educational handicap. Only then can the educational diagnosis of a learning disability be logically related to a remedial prescription based on the specific problems and needs of the individual student. What is implicit in this argument is that the broad range of learning disorders can be meaningfully subdivided into more homogeneous subgroups composed of individuals that share common features of specific learning problems, and which would likely respond similarly to specific remedial approaches.

Learning Disability Subtyping Research

The problem of identifying more homogeneous subtypes or subgroups of the learning disability population has been addressed by many investigators, using differing forms of methodology. These attempts have ranged from approaches based upon clinical diagnostic classifications (e.g., Boder 1973) to statistical approaches that rely upon computer generated subgroupings using either factor analytic (Petrauskas & Rourke, 1979) or cluster analytic methodologies (Doehring, Hoshko, & Byrons, 1979). These various techniques and approaches have given rise to a number of varying typologies, which will be examined with respect to the nature of subgroups identified.

Clinical Approaches to Subtyping Research

Boder (1973) used a clinical analysis approach to subtype reading disabled children into diagnostic classification groups, based upon the qualitative nature of their reading and spelling errors. She proposed (based upon clinical observations) three classification groups: 1) Dysphonetic dyslexics (about 67 percent of her sample), who have auditory-verbal deficits that are manifested in the process of analyzing and translating words into phonemes and which result in poor knowledge of letter-sound correspondences, 2) Dyseidetic dyslexics (about 10 percent of the sample), who display deficiencies in visual-spatial analysis and who fail to perceive words in their whole gestalt form, and 3) Mixed dysphonetic/dyseidetic dyslexics (about 23 percent), who display evidence of deficiency in both the verbal/phonetic and nonverbal/eidetic aspects of reading and spelling. It is generally assumed that the dysphonetic dyslexia deficits would be related to impairments of the left cerebral hemisphere, the dyseidetic dyslexic deficits to impaired right hemispheric functions, and the mixed deficit would be caused by combined left and right hemispheric impairments. These hypotheses have received some empirical support from studies of electroencephalographic asymmetries in normal and dyslexic readers (Fried, 1979; Fried, Tanguay, Boder, Doubleday, & Greensite, 1981) as well as from studies of perceptual asymmetries in the processing of verbal and nonverbal information (Pirozzolo & Rayner, 1979; Dalby & Gibson, 1981).

Another clinically based diagnostic classification of reading disability subtypes based on lateralized brain deficiencies was proposed by Bakker (1973, 1979, 1982, 1983), who hypothesized that each cerebral hemisphere plays a prominent role in the processing of reading material, but that developmental shifts occur as the learning process progresses from the early stages of perceptually analyzing unfamiliar script (described as primarily a right hemispheric process) to syntactic-semantic analysis of written material (a primarily left hemispheric process) at more advanced stages. Bakker hypothesized that deficiencies in the processing capability of either the right or left hemisphere could give rise to different patterns of reading deficit, with a deficit in right hemispheric processing resulting in an over-reliance upon left hemispheric linguistic strategies (L-type dyslexics), with reading errors characterized by a relatively fast reading rate with numerous substantive errors (e.g., omissions and substitutions/additions) as the person over-looks the spatial-perceptual features of the text.

Other children, with intact right hemispheric functioning but deficient verbal-linguistic abilities (P-type dyslexics), may tend to ignore the symbolic representational nature of the letters and words, and show a tendency to read slowly and make time consuming errors such as fragmentations and repetitions. These clinically defined reading disability subtypes have also been shown to have differential patterns of lateralization of language functioning on dichotic stimulation studies, with P-type dyslexics displaying a higher than normal incidence of right hemispheric dominance for language/speech functions as reflected in a left ear advantage on verbal dichotic stimulation testing (Bakker, Licht, Kok, & Bouma, 1980).

Mattis, French and Rapin (1975) applied the clinical approach to the analysis of neuropsychological test battery data on a group children who fell within three broad categories: brain damaged with no reading problems, brain damaged with dyslexia, and dyslexia without evidence of structural brain damage (i.e., developmental dyslexia). No significant whole group differences emerged between the two dyslexic groups, but three identifiable subgroups emerged that were seen in both dyslexic groups. The first group (about 38 percent of the sample) was characterized by the presence of a general language disability, the second group (37 percent of the sample) displayed pronounced motor-speech difficulties (articulation-graphomotor dyscoordination), and a third group (about 16 percent of the sample) was described as having a visuospatial perceptual impairment. The language disorder group was typified by deficits in naming, listening comprehension, oral sentence repetition, and speech sound discrimination tests. The articulation-graphomotor group displayed deficiencies in sound blending and graphomotor ability, and the visual perceptual disorder group displayed deficiencies in a variety of visual nonverbal tasks, including WISC Performance I.Q., Raven's Progressive Matrices, and the Benton Visual Retention Test.

In a cross-validation study (Mattis, 1978) a larger group of reading disabled children was similarly classified on the basis of nonreading test deficits, with 63 percent of this second sample classified as having language disorders, 10 percent as having articulation-graphomotor dyscoordination deficits, 5 percent with visual perceptual disorder, and 9 percent as showing mixed disorders incorporating features of two or more of the classification types. An additional 10 percent of this cross-validational group was also classified as having a temporal sequencing disorder, with deficits in such tests as sentence repetition, digit span, picture arrangement, and comprehension. These two classification studies, based upon neuropsychologically related tests of intellectual and other nonreading abilities, were consistent with the propositions of Boder that there is a visual perception disorder subgroup that comprises a relatively small percentage of the reading disabled population. The auditory/linguistic form of reading disability was, however, subdivided into as many as three distinct subtypes.

Denckla (1977, 1979) also described a variety of reading disability subtypes, based upon clinical observation and description, and six distinct groups were reported, including: 1) a global-mixed language disorder with general deficiencies in most language skills, 2) an articulation-graphomotor disorder with poor articulation as well as problems with fine motor coordination and pencil work, 3) an anomic-repetition disorder typified by poor naming (making frequent semantic errors) as well as deficient digit span and sentence repetition, 4) a dysphonemic sequencing disorder with poor digit span, sentence repetition and naming but with a preponderance of phonemic substitutions or reversal errors, 5) a verbal learning and memorization disorder, and 6) a correlational disorder in which reading was relatively normal but well below expectations for intelligence level.

Myklebust (1978) and his associates (Johnson & Myklebust, 1967) also described subtypes of reading disabilities based upon clinical observations, and identified basic auditory and visual types of disability (similar to the Boder dysphonetic and dyseidetic classifications) as well as several additional groups: an inner-language dyslexia or "word-calling" deficit that was described as a deficiency of integrative-neurosensory learning characterized by normal phoneme-grapheme encoding but poor comprehension; auditory dyslexia described as a deficit in intraneurosensory learning based upon difficulty in "cognitively auditorizing," symbolizing, and coding the written form of language but without impairment of spoken language comprehension; visual dyslexics who suffer from visual-verbal agnosia, a deficit in visualization or symbolization of written language forms; and intermodal or cross-modal dyslexia which was described as an interneurosensory learning deficit with intact intraneurosensory auditory and visual learning, but a deficit in translating the visual processes into auditory processes.

These early clinical diagnostic approaches were stimulating and of great heuristic value, but like all clinical diagnostic ventures, they can be criticized due to the difficulty in establishing the reliability and validity of the emerging typologies or classifications. The use of unstandardized clinical observations and/or visual inspection of psychometric profiles for classification research purposes has also been criticized as being easily influenced by a priori biases and assumptions regarding the essential nature of learning disabilities (Fletcher & Satz, 1985). It was noteworthy, however, that these various studies identified rather consistently at least two types of learning or reading disabilities with primarily verbal or primarily visual-perceptual deficits, based upon presumed dysfunctions of the right and left cerebral hemispheres respectively (and frequently a combination or mixed disability group related to bilateral cerebral dysfunctions). The various clinical observation approaches also identified a number of potentially significant patterns or variations in subskills underlying the academic learning process, although the makeup of the more detailed or specific subtypes was less consistent and often reflected probable differences in the composition of the samples studied.

<u>A Priori Grouping Research</u>

Another approach explored has been to investigate the neuropsychological functioning of learning disabled children subtyped a priori on the basis of their pattern of performance on academic achievement tests, in an attempt to determine if the differential academic deficits are related to different patterns or loci of neuropsychological impairment. Rourke and Finlayson (1978) classified children on the basis of their pattern of relative strengths and weaknesses on the Wide Range Achievement Test measures of reading, spelling and arithmetic performance: Group 1 composed of children with uniformly deficient reading, spelling and arithmetic; Group 2 composed of children who were relatively adept at arithmetic as compared to their performance in reading and spelling; and Group 3 composed of children with average or better competence in reading and spelling but relatively deficient arithmetic performance.

The performances of these three groups were contrasted on a series of 16 measures of intellectual and neuropsychological abilities. The major finding was that the performances of Groups 1 and 2 were superior to Group 3 on measures of visual-perceptual and visual-spatial abilities, while Group 3 was superior to Groups 1 and 2 on measures of verbal and auditory-perceptual abilities. The authors suggested that these subgroups based upon patterns of academic test performance performed on the neuropsychological tests in a manner consistent with the presence of a right cerebral hemisphere dysfunction in Group 3, while Groups 1 and 2 performed in a fashion similar to patients with left hemispheric dysfunctions. Such a hypothesis of differing foci of cerebral dysfunction in subgroups of learning disabled children was also supported by studies of learning disabled youths examining the presence of differential patterns of lateralized motor deficits using the Grooved Pegboard Test (Rourke, Yanni, MacDonald, & Young, 1973) and on a test of speeded sequential analysis and visual scanning, the Trail Making Test (Rourke & Finlayson, 1975).

In addition to the above research into the lateralization of neuropsychological dysfunctions in children with specific patterns of learning disabilities, Rourke and his associates have also conducted some longitudinal studies of learning disabled students, and in one study (Rourke & Orr, 1977) a group of reading disabled children were reassessed after a four year period to evaluate the accuracy of a series of tests of reading, spelling, intelligence, and a neuropsychological test of visual discrimination in predicting long-range academic retardation. The results of this longitudinal investigation suggested that even within the restricted category of reading retardation there were at least two distinct subgroups, one group (comprising about 74 percent of the sample) that continued to display serious deficiencies in reading attainment throughout the four year span (with a tendency to fall progressively further behind their age peers in achievement level) and another group (about 26 percent of the sample) that made substantial improvements in their academic performance to the point that they were able to read at normal or near normal levels.

The evidence of two distinct groups with clearly significant differences in pattern of developmental progress was interpreted as suggesting that for some reading disabled children (those in the smaller group that essentially "caught up" over the four year span) a developmental lag hypothesis was consistent with their progress, whereas in the larger subgroup, the consistent disability (and progressively worsening relative position in their peer group achievement distribution) pointed to the presence of a persistent or permanent neuropsychological deficit. These authors also reported that a neuropsychological screening measure (the Underlining Test) was a more accurate predictor of eventual subgroup membership than any of the specific reading, spelling, or intellectual measures employed in the study, suggesting that neuropsychological factors may be especially useful in providing not only descriptive data, but also in forming the basis for accurate predictions of developmental outcome for subgroups of learning disabled children.

While studies of learning disabled children based upon a priori classifications into different "groups" or "types" of learning disability have been successful in identifying some general factors or differences with regard to lateralization of brain dysfunction, such studies are intrinsically limited with respect to identifying the wide array of possible subtypes, due to the implicit biases or assumptions of the investigator who defines the initial group criteria. A major goal of classification research is to identify naturally recurring subtypes or clusters of individuals with similar features (Cormack, 1971), and if such subtypes of learning disability really exist, it should be possible to identify them and separate them from subgroups of other children with different learning problems and from children of normal learning ability. An approach that has come to be applied with increasing frequency in this area is the application of multivariate statistical analysis of objective data, to search for the naturally occurring subtypes or clusters on the objective basis of correlations or similarities between members on objective tests.

Multivariate Statistical Approaches

An early application of the multivariate statistical approach was reported by Doehring and Hoschko (1977), who applied the Q factor analysis procedure to data from a series of reading skills tests to identify subgroups among children with reading problems. The tests used in this investigation were selected to systematically evaluate basic component skills of reading, such as reading letters, syllables, and words, and presenting the materials by four different procedures: 1) visual matching to sample - pointing to which of three printed choices exactly matched a printed sample, 2) auditory-visual matching to sample - choosing which of three written choices corresponded to a spoken sample, 3) oral reading - reading aloud a series of printed items, and 4) visual scanning - scanning rows of printed items and underlining a specified target stimulus among other items of a similar type. Additional tests of sentence comprehension and spelling were also administered, but not included in the main classification process, due to the high error rate on these more complex measures by children with reading problems. A total of 31 reading test scores for each child were used in the Q technique factor analysis, and the product-moment correlation coefficient was utilized as an index of similarity between children.

On the basis of this Q factor analysis, three identifiable factors emerged, with each factor representing a different reading test battery profile pattern. The first factor, comprising 35 percent of the sample, was characterized by oral reading problems, with poor performances on tests of oral reading of words, phrases, and sentences, but near normal silent reading skills on the matching to sample tasks. The second factor (about 32 percent of the sample) was described as having an intermodal association problem, and the children in this group had particular difficulties with rapidly matching spoken and printed letters, words, and syllables. The third factor group (24 percent of the sample) had the greatest problems with sequential relations, and was notably poor on the visual and auditory-visual matching of words and syllables as compared to letters. No factor related to visual perceptual disabilities was identified in the first studied group (restricted to children with primary learning disability diagnoses), but such a factor was obtained in a second sample utilizing children with mixed types of learning problems (including some cases with diagnoses of childhood aphasia and mild mental retardation). A validational check of the reliability of the factor classifications was conducted using a cluster analysis procedure on the same data, and despite the use of a radically different methodology, the same types of reading disability subgroups emerged, with 85 percent of the restricted learning disability sample grouping into the same type or subgroup by both methods, and 90 percent of the mixed learning problem group being classified into identical groups by both factor analytic and clustering procedures.

Additional multivariate classification studies have been conducted by other researchers, using different batteries of test variables. Petrauskas and Rourke (1979) utilized a Q technique factor analysis on neuropsychological test data for 160 children, composed of 133 children that met standard diagnostic definitions for reading disability, and 27 normal readers without academic delays. The tests employed in this study were derived from a standardized neuropsychological assessment battery that covered six ability areas: 1) tactile perception, 2) sequencing ability, 3) motor skills, 4) visual-spatial ability, 5) auditory-verbal skills, and 6) abstract-conceptual tasks, with a total of 20 measures analyzed for each individual subject. Pearson product-moment correlation coefficients between individuals were utilized as a similarity index, and the resulting correlation matrix was factor analyzed to identify factors that represented reading disability subtypes. A split-half procedure was employed with each half of the total sample independently factor analyzed to check for the reliability of the resulting factor structure, plus a third factor analysis with the combined or total group sample.

In each case six factors or subtypes emerged from the data and three of these proved to be reliable and consistent across both half sample groups. Type 1 was characterized by well developed visual-spatial and eye-hand coordination skills, average or near-average tactile-kinesthetic abilities, abstract reasoning~~, and~~ nonverbal concept formation, near-average performance on word definitions, mildly impaired word blending, immediate memory and store of general information, and moderately to severely impaired on tests of verbal fluency and sentence memory. Type 2 was characterized by average or near-average kinesthetic, psychomotor, visual-spatial constructional and word-definition abilities, as well as average scores on tests of nonverbal problem-solving and abstract reasoning skills. Borderline to mildly impaired performances were seen on tests of immediate memory for digits and other sequencing tasks, on the store of general factual information, sound blending, verbal fluency, and verbal concept formation, and moderate to severely impaired performances were displayed on finger recognition, immediate visual-spatial memory, and memory for sentences. Type 3 was characterized by average finger recognition (for the left hand only), as well as average ability on kinesthetic, visuo-spatial constructional ability, vocabulary and sound blending abilities, and nonverbal concept formation. Borderline to mildly impaired performance was seen on finger recognition (right hand), immediate memory for digits, eye-hand coordination under speeded conditions, store of general information and nonverbal abstraction. Mildly to moderately impaired scores were displayed on verbal fluency, sentence memory, and immediate visual-spatial memory, while moderate to severely impaired scores emerged on verbal concept formation tests. Type 4 was characterized by average to above-average scores on all neuropsychological measures (composed almost exclusively of the normal reading control group). Type 5 (somewhat lower in reliability than the preceding groups) was characterized by average to high-average finger recognition and kinesthetic skills, visuo-spatial construction, vocabulary, and sound blending abilities, with near-average immediate memory for digits, eye-hand coordination, store of general information and nonverbal abstract reasoning and concept formation, and mildly to moderately impaired verbal fluency, sentence memory, and immediate visual-spatial memory.

On the basis of the patterning of the neuropsychological test results for these groups, Petrauskas and Rourke suggested that the Type 1 pattern was essentially similar to the pattern in adults associated with left temporal lobe lesions, that the Type 2 and the Type 5 pattern were suggestive of dysfunction of the posterior regions of the left hemisphere, and that the Type 3 deficit pattern was similar to that seen in both adults and children with lesions of the left frontal lobe.

A subsequent statistical study of the neuropsychological profile patterns of older children with combined deficiencies in both reading and spelling (Fisk & Rourke, 1979) also employed Q-type factor analysis of the neuropsychological test data also identified subtypes that were reliable over the age span from 9 to 14 years of age. The first group (19 percent of the sample) was characterized by extremely poor finger localization, the second group (also about

19 percent) was distinguished with respect to very poor speech sounds perception, and a third group (about 15 percent of the sample) had very poor finger-tip number writing test scores. The first group of the Fisk and Rourke study was similar to Type 2 of the Petrauskas and Rourke study with respect to the deficient finger localization skill, while the second group in Fisk and Rourke is similar to Type 1 of Petrauskas and Rourke with respect to poor verbal and speech sounds perception skills. Despite significant differences in selection criteria for subjects and in the neuropsychological variables selected for inclusion into the Q-type factor analysis in the two studies, two consistent patterns of neuropsychological deficits associated with reading disabilities emerged, one associated with pronounced deficits in verbal skills related to the functional integrity of the left temporal lobe, and one that included a deficiency in finger localization that was suggestive of impairment of the posterior left parietal lobe.

A different approach to classification research in the area of reading disabilities was pursued by Satz and his associates, who applied a cluster analysis procedure to subgroup poor readers on the basis of their reading, spelling and arithmetic achievement scores on the Wide Range Achievement Test (WRAT). One aspect of this research group's approach that was distinct from that of Rourke and his associates was to use clustering procedures to first identify and select a sample of poor readers from an unselected sample of students. Satz rejected the use of a predefined criteria for reading disability based upon observations of a relatively high rate of reading problems in a longitudinal study of academic achievement, which was considered surprising in view of the adoption of a definition by exclusionary criteria for dyslexia. The initial three year evaluation (Satz, Friel, & Rudegair, 1976) indicated approximately a 12 percent incidence rate for serious reading problems, as well as significant numbers of children with additional academic problems (e.g., arithmetic).

These observations prompted an evaluation of the concept of specific developmental dyslexia (Taylor, Satz, & Friel, 1979) in which a group of children with reading delays that met the World Federation of Neurology definition of dyslexia (i.e., excluding by definition any retarded reader with less than adequate exposure to instruction, intelligence, or socio-cultural opportunities) were compared to a group of same aged "nondyslexic" disabled learners that failed to met one or more of the definition criteria (e.g., low IQ, SES, etc.), as well as to groups of nondisabled learners. The results of this study indicated that the "dyslexic" and "nondyslexic" reading disabled groups did not differ significantly from each other on any variable, including neuropsychological tests, other academic tests, severity of reading problems, reversal and letter confusion errors, parental reading proficiency, neurological examinations, or personality questionnaires, although both reading disabled groups were clearly different from the control groups on these same variables. On the basis of these null findings, Satz et al. challenged the use of traditional exclusionary definitions of learning disabilities, and they initiated a program of classification research.

In their first classification study, Satz and Morris (1981) noted that prior studies had classified children on the basis of either achievement dimensions (e.g., Boder, 1973; Doehring & Hoschko, 1977; Rourke & Finlayson, 1978) or on the basis of processing deficiencies on cognitive and neuropsychological tests (Fisk & Rourke, 1979; Mattis et al., 1975; Petrauskas & Rourke, 1979). Satz and Morris (1981) therefore attempted to classify children on the basis of both academic achievement data as well as independently on cognitive or processing-deficiency variables, in an attempt to explore the degree of consistency in group makeup that would emerge from the two different types of testing. The first step in their study was to identify academic achievement defined clusters of children by applying an average-linkage hierarchical agglomerative clustering procedure (Wishart, 1975) to the data from the WRAT for all 236 students in their longitudinal follow-up population, using squared Euclidean distance as the similarity coefficient.

This initial cluster analysis resulted in the generation of nine clusters or subgroups, which included 98 percent of the sample. These clusters displayed significantly different patterns of achievement on the WRAT, with subgroups ranging from consistently above average scores on all three academic tests to consistently below average scores on all three variables. Consistent with the findings of differing patterns of neuropsychological functioning in the research of Rourke and his associates, patterns of deficiency on the WRAT were observed in several of the cluster analytic derived subgroups that matched closely Rourke's clinically based groupings on the basis of WRAT pattern. As in Rourke's research, one group emerged that was significantly impaired on all three subtests of the WRAT, another displayed relative weakness of both the reading and spelling subtests, but average performance in arithmetic, and another displayed the reverse pattern, with average level performance on both both reading and spelling but significantly weaker achievement on arithmetic.

This finding of distinctive subgroups, with learning disability subgroups similar to Rourke's clinical classifications, has also been obtained in two independent replications, one in The Netherlands (van der Vlugt & Satz, 1985) and another in the continental United States (Johnston, 1984). Both of these replication studies added females to their subject pools, as well as a measure of reading comprehension. No significant changes in group composition or ability pattern emerged with the addition of female subjects or the introduction of an additional reading measure, suggesting that the three subgroups are reliable and common across cultures and in both sexes.

In the second stage of the Satz and Morris study, two of the subgroups that performed about one standard deviation below the sample mean on the WRAT were selected as representing most of the disabled learners in the sample. Additional subtyping distinctions were then explored on these subgroups, by cluster analyzing their performances on the Beery Test of Visual-Motor Integration, a Recognition-Discrimination Test, the Verbal Fluency Test, and the WISC-R Similarities subtest. These tests were selected to represent a brief neuropsychological screening of both nonverbal and verbal processing abilities. This second stage clustering yielded five profiles representing distinct disability subtypes. Subtype 1 was observed to be impaired on both verbal measures (Similarities and Verbal Fluency) but had average scores on both visual-spatial tests. This group was termed a "general verbal" disability subtype. Subtype 2 was impaired on only the Verbal Fluency Test, and was termed a "specific verbal" (naming) group. Subtype 3 was impaired on all of the cognitive measures and was classified as a "mixed-global" subtype. Subtype 4 was impaired primarily on nonverbal perceptual tests and was termed a "visual-spatial" group. Subtype 5 demonstrated no significant impairment on any of the neuropsychological variables, and was labeled as an "unexpected" subtype.

External validational studies (comparing the derived subgroups on independent variables not included in the clustering process) demonstrated significant differences across the subgroups with respect to neurological status (including birth history) and parental reading levels. Children included in the general verbal, mixed-global and visual-spatial subgroups had a higher proportion of neurological "soft signs" while the specific verbal and unexpected subtypes had fewer neurological findings. The parents of the latter two subgroups were also seen to have significantly higher reading scores than the parents of children from the other three subtype groups.

What is observed from a consideration of the brief historical review of learning disability subtyping research that there has been a trend for researchers to identify various subgroups of learning disabled students, and to described significantly different types or patterns of academic and/or cognitive deficits for each subgroup. It is also noted that some of the early proposed clinical groupings (e.g., Boder's dysphonetic versus dyseidetic typology) have been at least partially supported and expanded upon by the more recent applications of multivariate statistical classification approaches. Specifically, in all of the research studies there has emerged a subgroup (frequently further subdivided into two or more subtypes) that would correspond to Boder's dysphonetic group, such as in Rourke & Finlayson (1978), who identified the presence of selective deficiencies in verbal and auditory-perceptual functions and intact visual-spatial skills in children with patterns of lower language arts (reading and spelling) skills than arithmetic on the WRAT. Also similar to Boder's classification of a dyseidetic reading disability group has been the observations of patterns of selective deficit in nonacademic neuropsychological tests that are generally sensitive to right hemispheric functions in some other groups of learning disabled children in the Rourke and Finlayson study.

In the Petrauskas and Rourke (1979) factor analytic study it was also reported that for a group of reading disabled children there were also a number of separate subgroups which displayed differing patterns of neuropsychological deficiencies in verbal and nonverbal skills, with Type 1 children displaying good visual-spatial and eye-hand coordination skills but significant impairments in such verbal functions as word blending, verbal memory for digits, long term verbal information, and especially significant impairments in verbal fluency and sentence memory tests. A second subtype identified in this study showed evidence of mixed deficits in both verbal and nonverbal abilities, with impaired averaged scores on memory for digits, long-term verbal memory, sound blending, verbal concept formation, as well as on tactile finger recognition, visual-spatial memory and memory for sentences. A third reading disability subtype was identified as having good visuo-spatial constructional abilities and nonverbal concept formation but impairments in verbal fluency, sentence memory, immediate visual-spatial memory and verbal concept formation. These authors interpreted their data as suggesting that the three identified reading disability subtypes in their sample were related to underlying neuropsychological dysfunctions that suggested the presence of impairments in different regions of the language dominant left hemisphere, with the Type 1 pattern similar to left temporal lobe lesion effects, Type 2 deficits being similar to posterior (parietal-occipital) lesions, and the Type 3 deficit pattern being similar to that seen from left frontal lobe lesions.

It is also noted that in independent studies, using a different methodology for identifying learning disabled statistically rather than by arbitrary definition, similar subgroup patterns of selective deficit emerged on standardized achievement test data, with the Satz and Morris (1981) study discovering naturally occurring subgroups with patterns of relative deficits on the WRAT that closely matched the a priori groupings of Rourke. It is also noted that on the cluster analytic subtyping of the identified learning disabled children, there emerged distinct subgroups, one with an apparent "general verbal" disability, another with a "specific verbal" deficit, a third with "mixed-global" deficits in both verbal and nonverbal functions, and a fourth with just "visual-spatial" dysfunction. A fifth group without any identifiable neuropsychological impairments was also observed in this study as well.

From this brief review, it is concluded that within any reasonably inclusive sample of learning disabled children, objective subtyping or classification procedures should give rise to multiple subgroups that will differ significantly on neuropsychological measures. Specifically, it is noted from the review that there has been a rather consistent observation of at least one subgroup with selective impairments in verbal skills (generally assumed to be reflective of left hemispheric dysfunction) but intact visual-spatial abilities (assumed to be reflective of adequate integrity of the right cerebral hemisphere). Also frequently observed has been the presence of a subgroup with mixed verbal and visual-spatial deficits, often described as a "global" or "mixed" deficit group, with dysfunctions attributed to both cerebral hemispheres. The presence of a subgroup with selective impairments of the visual-spatial functions but sparing of verbal skills has also been reported by several studies, but such a group has been reported to represent a considerably smaller proportion of the learning disability population, and in some studies (e.g., Denkla 1977, 1979; Doehring & Hoschko 1977) such a visual-spatial deficit group was not observed, most probably due to the sampling of students and/or the lack of systematic evaluation of nonverbal visual-spatial functions. The observation of an "unexpected" subtype group in the Satz and Morris (1981) study also suggests that in large samples of children with identified learning problems, especially if selected primarily on the basis of objective academic test criteria, there may emerge a subgroup that does not have identifiable neuropsychological deficits.

Literature Review on the Luria-Nebraska Neuropsychological Battery

The LNNB, as a new test battery in the area of neuropsychology, has been the focus of much research attention and some controversy in regards to its validity as a clinical instrument. The test battery was initially conceived by its authors as an attempt to operationalize and standardize the neuropsychological assessment procedures that were previously described by A. R. Luria (1966, 1973) but which had limited utilization in practice due to their lack of standardization, reliance on qualitative interpretation, and lack of objective criteria for rating the impairments displayed. Luria's approach and procedures to neuropsychology was criticized particularly by Reitan (1976) who noted that the only validation of Luria's procedures was Luria's own impression that they were effective, and as a result of the nonstandardized administration procedures and qualitative analysis involved, it was impossible to separate Luria's skills as an insightful clinician from the actual value of the tests themselves. Golden et al. (1980) indicated the intent of the LNNB was to:

> "establish a standardized version of the Luria Neuropsychological Battery... to overcome these limitations. It was our strong feeling that the standardization of Luria's items would not only vindicate Luria's theories and demonstrate the effectiveness of the items and the test battery as a whole, but would also make these procedures more widely available to psychologists in the United States and around the world. A standardized version would offer the advantages of testing techniques that would be the same from experimenter to experimenter as well as standardized scoring techniques. These characteristics would enable the researcher to attempt to replicate both results found by other researchers and those reported by Luria in his publications. The presence of standardized scoring would also allow for the experimental

investigation of the effectiveness of Luria's results. The claims for validity of the battery would not have to rest solely on the clinical impression."

Reliability Studies

Since the initial release of the LNNB, there has been a rapidly growing number of studies that have examined the reliability and clinical validity of the battery, using a wide range of patient/subject populations. The initial reliability study (Golden, Hammeke, and Purisch, 1978) evaluated the scoring system by administering the battery in the presence of a second examiner, with each tester scoring the patient's performance independently. This double scoring procedure, applied on five patients, yielded an overall interrater agreement of 95% for individual item comparisons, and the interrater agreement for each of the patients ranged narrowly between 92 to 98%.

An internal consistency approach to reliability evaluation was undertaken by Golden, Fross, and Grabber (1981), who examined the split-half reliability of each subsection of the LNNB battery. Using an odd-even split approach on the data from 338 patients of diverse diagnostic categories (74 normal controls, 83 psychiatric, 181 mixed neurological patients) the split-half reliabilities for each summary scale were determined, with a range from .89 to .95 (see Table 1).

Table 1

Luria-Nebraska Item-Scale Consistency and Spilt-Half Reliabilities

Scale	Total no. of items on scale	No. of items with highest r on scale	Split-half r
Motor	51	46	.92
Rhythm	12	12	.90
Tactile	22	22	.90
Visual	14	12	.91
Receptive Speech	33	32	.91
Expressive Speech	42	34	.93
Writing	13	12	.94
Reading	13	13	.95
Arithmetic	22	22	.90
Memory	13	13	.89
Intellectual	34	32	.93
Total	269	250	

Source: Golden, Fross, and Graber (1981), p. 305.

A study of test-retest reliability was conducted by Golden, Berg, and Graber (1982) in which 27 patients with chronic, static injuries with initial test performances falling in the middle range of the normative distribution were retested with the LNNB. An average test-retest interval of 167 days (SD = 134 days) with a range of 10 to 469 days was reported, and the correlation coefficients for the test-retest scores ranged from .78 to .96 on the individual summary scales (see Table 2). The authors also reported that covariance analysis, parcelling out the effect of the test-retest interval, yielded little change in the resulting correlations, suggesting that the length of the interval had little effect on the results.

Table 2

Luria-Nebraska Clinical Scale Means and Standard Deviations for Initial Testing and Retest and Test-Retest Reliability

Variable	Initial Test M	Initial Test SD	Retest M	Retest SD	Reliability
Motor	62.2	20.9	60.8	21.9	.94
Rhythm	68.5	20.6	66.3	19.4	.90
Tactile	60.6	16.1	57.0	12.4	.78
Visual	65.0	12.8	61.8	11.3	.86
Receptive Speech	68.5	17.5	67.8	19.7	.87
Expressive Speech	76.5	24.7	66.7	24.3	.94
Writing	66.7	15.9	67.5	14.7	.92
Reading	63.3	16.0	63.2	16.0	.92
Arithmetic	84.7	33.6	83.5	32.0	.96
Memory	70.4	13.4	70.2	16.0	.84
Intelligence	78.8	17.4	75.5	17.9	.86
Right Hemisphere	53.7	19.6	50.6	14.0	.77
Left Hemisphere	54.9	19.0	53.9	16.2	.86

Note. df=25; all reliability coefficients are significant at $p<.001$.

Source: Golden, Berg, and Graber (1982), p.453.

Subsequent investigations have been undertaken to replicate the results of these initial reliability studies. Bach, Harowski, Kirby, Peterson, and Schulein (1981) conducted an interrater reliability study utilizing two test subjects and five raters. They pointed out that the initial interrater reliability study of Golden, Hammeke, and Purisch required empirical replication, and that the original study did not incorporate any statistical correction for the inflation of the reliability estimates due to the dichotomous scoring of most of the LNNB items. These investigators utilized a procedure of having five raters score the protocols of two test scoring criteria serving as the subjects, with the express intention of offering "marginal" or difficult to score responses to maximize the potential for scoring inconsistency. They also calculated estimates of interrater agreement that are uninflated by chance agreement (Cohen's Kappa). These authors indicated the inclusion of a larger number of raters, the correction of the interrater agreement for the chance agreement factor and the use of trained confederates to provide test responses of maximal scoring difficulty all served to yield a very conservative estimate of the LNNB's interrater reliability.

Under these more rigorous conditions, there was a significant decline in agreement rates for the LNNB clinical scale summary scores, with the greatest declines being observed on the scales on which the confederates offered "marginal" responses. However, even with the more severe conditions designed to maximize the interrater variability, an overall proportion of interrater agreement of .90 (K = .90) was obtained for the scales responded to in a "normal" fashion (i.e., to the best of the subjects ability), and an overall proportion of agreement of .75 (K = .74) for the same scales under the "marginal" response conditions (See Table 3). The authors concluded that: "Under conditions designed to maximize variability due to error between raters, the LNNB remains capable of reliable use. Proportions of interrater agreement and Kappa coefficients are, in most cases, within the acceptable range.... Further, the LNNB seems to be a rather robust instrument. Under conditions designed to optimize interrater variance ("marginal" instructions and five raters) the proportion of interrater agreement was .75 - surprisingly high given the conditions under which it was obtained." (Bach et al. 1981, p. 20).

Table 3

Proportion of Interrater Agreement on Individual LNNB Scales Under Normal and Marginally Scorable Response Conditions

Scale	n Items	Normal Proportion Agreement	Marginal Proportion Agreement
Motor	51	.82	.55
Rhythm	12	1.00	.92
Tactile	22	.95	.86
Visual	14	.79	.64
Receptive Speech	33	.94	.88
Expressive Speech	42	.93	.83
Writing	13	.84	.54
Reading	13	.84	.84
Arithmetic	22	.86	.91
Memory	13	1.00	.77
Intelligence	34	.94	.68
Total	269	.90	.75

Source: Bach, Harowski, Kirby, Peterson, and Schulein (1981), p.20.

Moses and Schefft (1985) also replicated the interrater reliability analyses of Golden, Hammeke and Purisch (1980) using a total of 36 patients from a Veterans Administration Hospital population, with a wide variety of diagnostic categories (cerebral contusion, intrinsic cerebral neoplasm, post-meningitic state, cerebrovascular accident, multiple sclerosis, primary degenerative dementia, and many others). As in the initial interrater reliability study the test was administered by one examiner in the presence of the other, and each item was independently scored by both examiners. Analysis of the results for the individual items on the 36 patients' test protocols (a total of 9,684 comparisons) demonstrated exact agreement in 95.71% of the comparisons. Of the 415 scoring disagreements between raters, 339 involved a one point difference which reflected "minor interpretive or scoring disagreements based on partial crediting of a response" (Moses & Schefft, 1985, p. 37). Calculation of interrater agreement frequency, mean difference scores and interrater correlations for the LNNB summary scales indicated that the there was an extremely close agreement between raters on these summary measurements from the LNNB, with no scale showing a mean difference score greater than one point, and interrater reliability correlation coefficients ranging from .97 to .99 (see Table 4).

Table 4

Interrater Mean Difference Score Statistics and Interrater Correlations for LNNB Summary and Localization Scales

LNNB Summary Scales	Difference Scores* Mean	SD	Pearson Correlation r	P
Motor	0.47	3.29	.97	.001
Rhythm	0.36	4.30	.98	.001
Tactile	-0.33	1.53	.99	.001
Visual	-0.14	2.03	.99	.001
Receptive Speech	0.00	2.39	.99	.001
Expressive Speech	-0.36	2.81	.98	.001
Writing	1.11	3.48	.97	.001
Reading	1.19	2.84	.97	.001
Arithmetic	0.81	2.98	.99	.001
Memory	0.08	1.42	.99	.001
Intellectual	1.47	2.91	.98	.001

Table 4 (cont.)

LNNB Localization Scales	Difference Scores*		Pearson Correlation	
	Mean	SD	r	P
Left Frontal	0.31	2.33	.99	.001
Left Sensorimotor	0.31	2.14	.98	.001
Left Parietal-Occipital	1.72	2.81	.98	.001
Left Temporal	0.61	2.22	.98	.001
Right Frontal	0.56	3.87	.96	.001
Right Sensorimotor	2.11	3.54	.96	.001
Right Parietal-Occipital	0.36	2.50	.99	.001
Right Temporal	0.19	2.03	.99	.001

* Difference score = (Rater one - Rater two)

Source: Moses and Schefft (1985), p.32.

The use of the split-half reliability coefficient approach applied by Golden, Fross, and Graber (1981) was based on the assumption of equivalence of variance of the resulting split halves, and this assumption was questioned by Maruish, Sawicki, Franzen, and Golden (1985) who noted that the equivalence of variances between parallel halves so created has not been empirically demonstrated. They also noted that the very heterogeneous sample of brain impaired, psychiatric, and normal subjects included in the Golden et al. (1981) study may have artifactually inflated the reliability coefficients, as reliability coefficients are generally inversely related to the homogeneity of the sample (Anastasi, 1982).

Maruish et al. (1985) analyzed the reliability of the LNNB summary scales using the alpha coefficient reliability approach, and separated subgroups of brain impaired, schizophrenic, mixed psychiatric (excluding schizophrenic diagnoses), and normal controls for separate analyses. The reliability estimates derived by this more conservative approach ranged from .84 to .93 for the brain impaired group, [illegible] to .93 for the schizophrenic group, .81 to .92 for the mixed psychiatric subjects, and .40 to .78 for the normal controls (see Table 5). The results were interpreted as demonstrating a continued high degree of internal consistency for the LNNB measures for all the subsamples except the normal controls. The markedly lower reliability coefficients on the normal group was attributed primarily to very limited variance in test scores among this group, as most normals find the vast majority of the LNNB items quite easy, and the standard deviation of the normal group's averaged scores was very small, resulting in marked truncation of the reliability estimates.

Table 5

Comparative Coefficient Alpha Reliabilities of Several Subsamples for the LNNB Clinical Summary and Localization Scales

Clinical Summary Scales	Sample* A	B	C	D
Motor	.93	.93	.92	.77
Rhythm	.84	.86	.83	.67
Tactile	.89	.89	.85	.54
Visual	.84	.83	.84	.59
Receptive Speech	.90	.89	.84	.69
Expressive Speech	.93	.93	.89	.74
Writing	.84	.86	.85	.59
Reading	.88	.88	.87	.68
Arithmetic	.93	.93	.90	.74
Memory	.84	.86	.82	.72
Intellectual Processes	.91	.92	.89	.78
Pathognomonic	.88	.87	.84	.68
Left Hemisphere	.85	.85	.81	.40
Right Hemisphere	.92	.87	.88	.53

Table 5 (cont.)

Localization Scales	Sample* A	B	C	D
Left Frontal	.92	.92	.90	.69
Left Sensorimotor	.88	.89	.84	.66
Left Parietal-Occipital	.90	.90	.88	.71
Left Temporal	.88	.87	.84	.66
Right Frontal	.77	.80	.77	.51
Right Sensorimotor	.81	.81	.78	.51
Right Parietal-Occipital	.89	.89	.87	.63
Right Temporal	.87	.90	.88	.69

*Samples: A= Brain Impaired (n=451)
B= Schizophrenic (n=414)
C= Mixed Psychiatric (n=128)
D= Normal (n=108)

Source: Maruish, Sawicki, Franzen, and Golden (1985)

A further replication of the test-retest reliability of the LNNB was conducted by Plaisted and Golden (1982) who examined the stability of scores on retest for the LNNB clinical summary scales, localization scales and the factor scales. In this study, 30 psychiatric patients were readministered the LNNB with a mean test-retest interval of 8.1 months (S.D.=6.0 months). The test-retest correlations for the 14 summary (clinical) scales ranged from .83 to .96, averaging .89 and for the 8 localization scales the reliability coefficients ranged from .78 to .95, averaging .89 (See Table 6). On the factor scales, two factors were observed to lack statistically significant reliability, and 6 of these thirty scales obtained reliability coefficients of less than .7, suggesting some limitation in the utility of the factor scales for test-retest comparisons. However, for the summary clinical and localization scales, the results of this study were highly consistent with those of previous reports, and supported the view of the LNNB as a reliable instrument capable of yielding stable data over time.

Table 6

Means, Standard Deviations, and Test-Retest Reliability Coefficients for the LNNB Clinical and Localization Scales

Variable Reliability* (Scale)	Initial Test		Retest		
	Mean	S.D.	Mean	S.D.	
~~Motor~~	~~33.3~~	~~19.0~~	~~33.7~~	~~21.7~~	~~0.87~~
Rhythm	12.3	6.6	12.7	6.9	0.90
Tactile	13.9	11.2	11.3	11.3	0.85
Visual	13.4	5.0	13.2	5.4	0.86
Receptive Speech	17.8	12.3	18.1	13.[illegible]	0.87
Expressive Speech	27.9	19.6	26.9	19.2	0.95
Writing	13.5	6.9	13.1	6.6	0.90
Reading	10.3	7.1	10.4	7.0	0.92
Arithmetic	17.2	14.0	16.3	14.0	0.96
Memory	15.6	6.2	15.9	6.4	0.83
Intellectual Processes	39.1	15.8	35.1	15.3	0.90
Pathognomonic	25.5	11.6	23.9	13.4	0.87
Right Hemisphere	11.9	10.1	11.7	10.7	0.91
Left Hemisphere	11.9	10.5	11.3	10.0	0.88
Left Frontal	30.2	17.0	30.0	17.2	0.95
Left Sensorimotor	27.2	13.4	27.2	13.7	0.91
Left Parietal-Occipital	23.5	13.3	23.2	13.1	0.93
Left Temporal	19.4	10.0	19.7	11.3	0.90
Right Frontal	14.7	5.3	15.4	7.4	0.78
Right Sensorimotor	14.9	7.4	15.4	7.1	0.86
Right Parietal-Occipital	22.9	11.7	21.6	11.9	0.88
Right Temporal	27.9	11.0	26.8	11.3	0.92

*df = 29, all significant at $p < 0.001$

Source: Plaisted and Golden (1982)

Validity Studies

The above referenced reliability studies have indicated that the LNNB has a sufficiently robust level of interrater reliability, internal consistency and test-retest reliability to warrant its use in experimental situations. The utility of a test battery in practical applied settings is, however, also dependent on its ability to yield valid results that are consistent with the purposes for which it was designed. The validational investigations of the LNNB have focused on several areas: 1) The concurrent validity of the LNNB to discriminate between brain damaged and non-brain damaged patients, 2) The concurrent validity or consistency of results between the LNNB and the Halstead - Reitan Neuropsychological Test Battery (HRNTB), and 3) Construct validity based on the correlation of results of the individual LNNB clinical scales with other standardized tests of similarly named functions (e.g., comparison of LNNB Intellectual Processes scale with WAIS or WISC-R results, comparison of the LNNB Memory scale scores to the Wechsler Memory Scale).

Discrimination of Brain Damaged Patients

In the initial standardization studies of the LNNB Golden, Hammeke, and Purisch (1978) examined the effectiveness of the individual test items for discriminating between samples of brain-damaged and normal patients. Utilizing a sample of 50 neurological patients (27 males and 23 females) with conclusive medical diagnoses of brain-damage and no evidence of current or previous psychiatric disorder and a control sample of 50 patients (24 males and 26 females) who were in hospital for non-neurological diseases or injuries (except for some patients with back injuries without head trauma) all items of the LNNB were compared for diagnostic discrimination power using t-test comparisons. A total of 252 items were found to discriminate significant at the .05 level or better, and a discriminant analysis using 30 of the most robustly discriminating items allowed for the classification of the 50 brain-damaged and 50 normal medical control patients with 100% accuracy.

Hammeke, Golden, and Purisch (1978) reported the results of a study utilizing the 14 summary or clinical scales to differentiate between brain-damaged and normal controls. Using the data from the same neurological and control groups described in the Golden et al. (1978) study, the neurological and control groups were found to show significant differences on all fourteen summary measures (See Table 7), and a discriminant analysis using the fourteen summary scale scores correctly classified all 50 of the control patients and 43 of the brain-damaged patients, an overall hit rate of 93%.

Table 7

Means, Standard Deviations and t-tests for LNNB Summary Variables

Variable	Brain Injured* Mean	Brain Injured* S.D.	Control* Mean	Control* S.D.	t**
Motor	44.36	19.70	19.24	8.73	8.24
Rhythm	12.21	5.60	5.89	2.64	6.70
Tactile	19.21	12.00	8.02	5.18	6.11
Visual	16.14	5.53	8.80	4.01	7.60
Impressive Speech	23.67	12.82	9.48	5.09	7.27
Expressive Speech	33.34	16.59	15.18	6.93	7.14
Reading and Writing	24.67	11.91	10.10	6.05	7.71
Mathematics	16.15	12.43	4.02	3.84	6.59
Memory	16.25	6.05	7.86	4.13	8.10
Intelligence	38.32	15.53	18.80	8.08	7.88
Pathognomonic	34.46	12.09	16.30	6.19	9.45
Left Hemisphere	18.82	10.90	7.38	4.41	6.88
Right Hemisphere	18.00	11.11	7.50	4.28	6.23

*N = 50
df = 98
** all t-test values $p < 0.05$

Source: Hammeke, Golden, and Purisch (1978)

In addition to the demonstration of the discriminative power of the LNNB to differentiate between neurological and cognitively normal medical patients, the test authors also conducted further validational studies to examine the battery's ability to discriminate between neurological and schizophrenic patients, as this differentiation is one of considerable clinical importance, and is recognized as a particularly difficult discrimination using the more established Halstead - Reitan neuropsychological test battery (Golden, 1977, 1978). Purisch, Golden, and Hammeke (1978) investigated the efficacy of the LNNB to make this discrimination, by administering the LNNB to a sample of 50 chronic schizophrenic patients (who had been screened for the possibility of neurological disorders and all of which had normal physical neurological examinations and normal EEGs) and comparing their results to those of the neurological patient group described in Golden et al. (1978).

Two-tailed t-test comparisons on the individual test items found the schizophrenics to perform significantly better (at the .05 level) than the neurological patients on 72 items, and the brain-damaged patients scored significantly better on only two items. A step-wise discriminant analysis using 40 items was able to achieve 100% accuracy in classification of the patients in both groups. The results using the 14 summary scale measures indicated that the schizophrenic group out-performed the brain-damaged group on 10 of the summary measures, and a discriminant analysis of the summary scale scores allowed for the correct classification of 92% of the schizophrenic group and 84% of the neurological group, an overall hit rate of 88% (See Table 8 for details on the means, standard deviations and t-test statistics on these samples).

Table 8

Means, Standard Deviations, and t-tests for LNNB Summary Variables

Variable	Brain Injured Mean	S.D.	Schizophrenic Mean	S.D.	t
Motor	44.36	19.70	34.20	17.53	2.73*
Rhythm	12.21	5.60	12.26	5.70	-.05
Tactile	19.32	12.00	13.62	8.58	2.73*
Visual	16.14	5.53	12.82	5.32	3.06*
Impressive Speech	23.67	12.82	19.48	10.49	1.79
Expressive Speech	33.34	16.59	24.86	14.34	2.73*
Reading and Writing	24.67	11.91	17.51	10.51	3.19*
Arithmetic	16.15	12.43	10.07	9.21	2.78*
Memory	16.25	6.05	15.84	6.61	0.33
Intelligence	38.32	15.53	33.24	13.80	1.73
Pathognomonic	34.46	12.10	22.20	9.46	5.65**
Left Hemisphere	18.82	10.91	13.24	8.37	2.87*
Right Hemisphere	18.00	11.12	12.44	8.35	2.83*

N = 50
df = 98
* $p < 0.01$
** $p < 0.001$

Source: Purisch, Golden, and Hammeke (1979)

A cross-validation of the initial validational study was conducted by Moses and Golden (1979), using new samples of 50 neurological and 50 control patients. The results of this study were very consistent with those of the original, and a discriminant analysis using the 14 summary scales yielded an overall hit rate of 96%, compared to the 93% hit rate reported by Hammeke et al (1978). The original study comparing schizophrenic and brain-damaged patients was also replicated by Moses and Golden (1980), who used the same neurological sample as in their 1979 study, but a new sample of 50 schizophrenic patients. In this cross validational study, a classification hit rate of 87% was achieved, which compares favorably, with the 88% rate reported by Purisch, Hammeke and Golden (1978). Shelly and Goldstein (1983) also replicated the Purisch et al. (1978) study with completely independent samples. In this study, a total of 30 neurological and 30 schizophrenic patients were assessed with the LNNB, and a discriminant analysis of the data from the 14 summary scales yielded a classification accuracy of 90%.

Identification of Lesion Lateralization or Location

In addition to the above studies that demonstrated the discriminative power of the LNNB to differentiate between neurological and non-neuropsychological patients, the battery has also been evaluated with regard to its ability to differentiate between neurological patients with lateralized and diffuse brain injuries. Osmon, Golden, Purisch, Hammeke, and Blume (1979) used the LNNB to evaluate 20 patients with left hemisphere damage, 20 patients with right hemisphere damage and 20 patients with diffuse brain damage, with all groups matched for age, education and chronicity. A discriminant analysis using the 14 summary measures resulted in correct classification of 59 of the 60 patients.

Lewis, Golden, Moses, Osmon, Purisch, and Hammeke (1979) reported on the ability of the LNNB to localize brain damage to the frontal regions, sensorimotor areas, temporal lobes, and occipital-parietal lobes of each hemisphere. The sample for this investigation consisted of 24 patients with focalized right brain damage and 36 patients with focal left brain damage, as confirmed by surgery, angiogram, CAT scan, or a combination of these procedures.

Discriminant analyses were conducted using the summary scales, and resulted in correct localization classifications of 22 of the 24 right hemispheric cases and 29 of the 36 left hemisphere cases.

McKay and Golden (1979) also reported on the derivation of a series of localization scales for the LNNB based on individual items that discriminated between patients with focalized lesions in the eight cerebral areas described in the Lewis et al. (1979) study. In the McKay and Golden (1979) study, 77 normal controls and 53 neurological patients with localized brain damage (subdivided into eight groups dependent on the localization of their brain damage) were assessed with the LNNB, and multiple t-tests were conducted to compare the results of each brain-damaged group with the normals on each LNNB item. Items that discriminated between normals and selected localization groups were included in that group's scale, with no items being included in more than two localization group scales and no two scales having more than two items in common. Each resulting scale was standardized on the results of the normal group, and the results demonstrated that each locally impaired group scored highest (most impaired) on its corresponding localization scale. In individual comparisons on the localization scales, 47 of the 53 patients scored highest on the scale corresponding to their actual area of diagnosed brain damage, with the majority of misses in classification resulting on patients with T-scores above 70 on all localization scales, reflective of very generalized neuropsychological impairment, regardless of the localization of the structural neuropathology. Application of the localization scales to patients without diffuse neuropsychological impairment resulted in the correct classification of 42 out of 43 cases.

Luria-Nebraska and Halstead-Reitan Comparisons

A standard procedure in the development of a new psychological test is to evaluate the degree to which it produces results comparable to already established tests that purport to measure the same variables. In the field of neuropsychology, the most widely used test battery has traditionally been the Halstead - Reitan Neuropsychological Battery (HRNB), which was standardized and validated by R.M. Reitan (Reitan, 1964) and represented an extension and adaptation of Ward Halstead's original battery of neuropsychological tests (Halstead, 1947). The HRNB has received wide usage in many settings, and has a large literature related to its efficacy for identifying brain damage (c.f., Reitan & Wolfson, 1985).

An initial investigation of the comparability of the LNNB to the HRNTB was conducted by Kane, Sweet, Golden, Parsons, and Moses (1981), who compared the accuracy of the two neuropsychological batteries to correctly classify 45 "hard-to-diagnose" patients. These authors reported that the use of each battery by an "expert" examiner (i.e., one with extensive experience with the test being utilized) yielded agreement on 37 of the 45 cases. In the 8 remaining cases the LNNB correctly classified 5 cases, while the

HRNTB correctly classified 3 cases. The authors concluded that in reference to making a discrimination between brain damage and no brain damage, the two batteries were essentially equivalent.

Another approach to the evaluation of the comparability of the two test batteries is to examine the correlation between the scores of the two batteries. Golden, Kane, Sweet, Moses, Cardellino, Templeton, Vicente, and Graber (1981) demonstrated that the 14 summary scales of the LNNB were significantly correlated with the major variables or tests of the HRNB. These authors reported that there was approximately 73% shared variance between the LNNB and HRNTB, indicating that the two batteries overlap to a great extent with regards to the basic skill areas they assess. Utilization of a discriminant analysis with the scores from the LNNB summary scales achieved a hit-rate of 87% (42/48) in the neurological group of this study, while 88% (53/60) of the control group were correctly classified. On the same subjects, the discriminant analysis of the HRNB data achieved a 90% (43/48) hit-rate for the neurological group and an 84% (50/60) hit-rate in the control group. These hit-rates, derived from an objective, statistical analysis and classification procedure are consistent with the results of the Kane et al. (1981) study in displaying an equivalence of diagnostic accuracy between the two batteries in adult populations. Shelly and Goldstein (1982) also compared the LNNB and HRNB on a sample of adult neuropsychiatric patients, and reported very high correlations between measures of general impairment on each battery, as well as significant commonality between them on a factor analysis.

Berg, Bolter, Ch'ien, Williams, Lancaster, and Cummins (1984) reported a study devised to evaluate the comparability of the LNNB and HRNB approaches for children and adolescents. In this study, either the LNNB or the Luria - Nebraska Neuropsychological Battery for Children (LNNB-C) (composed primarily of items from the adult LNNB that have been demonstrated to be applicable to children under 13) was administered, depending on the age of the youth, as well as either the HRNB or Halstead - Reitan Neuropsychological Battery for Children (HRNB-C). Thirty juvenile patients aged 9 years or older with chronic idiopathic epilepsy were assessed with both the age appropriate Luria and Halstead - Reitan batteries. The two batteries were then compared as to their consistency in identification of neuropsychological impairment, based on objective criteria previously established for each test battery.

Combination of both the children's and adult's battery data resulted in an overall agreement of 87% (26/30) between the respective Halstead - Reitan and Luria - Nebraska batteries. The percentage of agreement was more robust for the younger age groups, with a 91% agreement between the LNNB-C and HRNB-C, and for those patients aged 13 to 14 years who were administered the adult version LNNB and the HRNB-C (because of the different age ranges covered by each battery) there was a 91% (11/12) agreement. The authors concluded: "The results of the current study indicate strongly that the Luria - Nebraska and the Halstead - Reitan batteries are essentially equivalent in terms of diagnostic power.... These findings do not suggest that one battery is superior to the other when using objective criteria to determine presence or absence of dysfunction." (Berg, et al., 1984, pp. 201-202).

A similar approach to comparison was also applied in evaluating the comparability of the LNNB-C and HRNB-C approaches for older children. Tramontana, Sherrets and Wolf (1983) reported that the Children's version of the LNNB yielded results very consistent with those of the HRNB for older children, with an 86% agreement in the classification of impairment between the two batteries when impairment was defined on the LNNB as three or more scales above a T-Score of 60 points.

Comparisons with the WAIS

In addition to the comparisons with the HRNTB, the LNNB has also been studied as to its correlations with other psychological instruments, most notably the Wechsler Adult Intelligence Scale and the Wechsler Memory Scale. These comparisons have largely been directed towards establishing that the LNNB summary scales relate to their named construct in a meaningful manner. In the Golden, Kane, Sweet, Moses, Cardellino, Templeton, Vincente, and Graber (1981) study, correlations between the LNNB scales and all HRNB variables, including the WAIS, were calculated. In this study, it was reported that the LNNB Intellectual Processes scale correlated highly with the WAIS I.Q. variables, with a correlation of -.84 with the Verbal I.Q., a correlation of -.74 with Performance I.Q., and a correlation of -.84 with the WAIS Full Scale I.Q. score, all highly significant.

Prifitera and Ryan (1981) investigated the validity of the LNNB Intellectual Processes scale (IP) as a measure of adult intelligence, and on a sample of 33 psychiatric patients (31 males, 2 females) they reported correlations between the IP and WAIS I.Q. scores as follows: V.I.Q.= -.86, P.I.Q.= -.76, and F.S.I.Q.= -.86, all correlations being highly significant ($p < .001$).

McKay, Golden, Moses, Fishburne, and Wisniewski (1981) reported a cross validation of the Prifitera and Ryan (1981) study, using a substantially larger and more diverse population, and also expanding the analysis to investigate the relationships between the WAIS I.Q. scores and all the LNNB summary scales. In this study a sample of 280 subjects (110 males, 170 females) of diverse clinical status (97 neurological, 110 psychiatric, 28 normal, and 45 without definitive diagnosis) were tested with both the LNNB and WAIS. Correlations between the LNNB Intellectual Processes Scale and the WAIS I.Q.'s were extremely consistent with those reported by Prifitera and Ryan (1981), with r's of: V.I.Q.= -.84, P.I.Q.= -.74, and F.S.I.Q.= -.84. It was also reported that the correlations between all the other LNNB summary scales and the WAIS I.Q. scores were also significant at the .01 level, with correlations ranging from a low of -.47 between the Verbal I.Q. and LNNB Tactile scale to -.81, between V.I.Q. and the LNNB Arithmetic Scale (see Table 9).

Table 9

Pearson Correlations Between WAIS and Luria-Nebraska Scales

LNNB Scale	V.I.Q.	P.I.Q.	F.S.I.Q.
Motor	-.62	-.67	-.67
Rhythm	-.72	-.68	-.74
Tactile	-.47	-.56	-.53
Visual	-.61	[illegible]	-.69
Receptive Speech	-.77	-.66	-.76
Expressive Speech	-.79	-.69	-.78
Writing	-.77	-.66	-.76
Reading	-.78	-.62	-.75
Arithmetic	-.81	-.70	-.80
Memory	-.73	-.71	-.76
Intellectual Processes	-.84	-.74	-.84
Pathognomonic	-.65	-.65	-.68
Right Hemisphere	-.55	-.61	-.60
Left Hemisphere	-.49	-.58	-.56

all correlations significant at $p < .01$

Source: McKay, Golden, Moses, Fishburne, and Wisniewski (1981)

These correlation results demonstrated that while the LNNB Intellectual Processes scale is highly correlated with the intellectual functions tapped by the WAIS, there are substantial intellectual factors involved in all the summary scale scores of this battery. This is a result that conforms to general neuropsychological principles, as brain injuries typically result in decrements in many aspects of adaptive functioning, particularly in intellectual abilities (c.f., McFie, 1975; Wheeler & Reitan, 1963). The significant correlations between the various LNNB summary scales and the three WAIS I.Q. scores lead to McKay et al. (1981) developing a series of regression formulae to predict the I.Q. scores, utilizing the T-Scores from the three best predicting LNNB scales for each I.Q. variable, with multiple correlation coefficients of 0.88 for the V.I.Q., 0.80 for the P.I.Q., and 0.87 for the F.S.I.Q. variable. In each of these multiple regression formulae, the LNNB Intellectual Processes scale was the predictor with the highest Beta weight, and the additional scales included in each formula reflected the general distinction between verbal and nonverbal intellectual skills tapped by the battery as a whole. Thus, the best predictor of Verbal I.Q. was the weighted combination of the Intellectual Processes, Reading, and Arithmetic scales, the best predictor of Performance I.Q. was the Intellectual Processes, Visual, and Memory scales combination, and the Full Scale I.Q. was best predicted by the combination of Intellectual Processes, Arithmetic, and Rhythm scales. It was concluded by the authors that: "selected Luria - Nebraska scores may provide reasonable estimates of summary IQ scores.... The Intelligence scale alone appears to provide a reasonable estimate of WAIS Verbal and Full Scale IQ scores, and the Intelligence, Memory, and Visual scales combined may give useful estimates of Performance IQ. By the same token the high correlations indicate that summary WAIS IQ scores account for about 50% of the variance on the Luria - Nebraska scales. Psychometric intelligence is clearly one of the major component skills that the Luria - Nebraska measures." (McKay et al., 1981, p. 945).

Picker and Schlottmann (1982) reported an investigation of the validity of the LNNB Intellectual Processes scale as a predictor of WAIS I.Q.'s in nonneurological students. In this study, 30 high school and 30 university undergraduate students were administered both the WAIS and the LNNB Intellectual Processes scale, and significant correlations were observed between the Intellectual Processes scale and the three WAIS I.Q. scores (correlation with V.I.Q.= -.67, with P.I.Q.= -.63, and with F.S.I.Q.= -.74, all correlations p < .0001). This relationship was observed, however, only for subjects with lower I.Q.'s on the WAIS, as separate analyses of the correlations between the Intellectual Processes scale and I.Q.'s for subjects with WAIS Full Scale I.Q.'s above and below 110 indicated that the significant correlations were observed only on the low I.Q. group (See Table 10).

Table 10

Correlations of WAIS I.Q.'s and Intellectual Processes Scale for Low and High I.Q. Groups

	Low I.Q.	High I.Q.
V.I.Q.	-.589***	-.172 NS
P.I.Q.	-.626***	-.020 NS
F.S.I.Q.	-.764***	-.137 NS

*** = p < .001
NS = Not Significant

Source: Picker and Schlottmann (1982)

This finding was interpreted as reflecting the relative simplicity of the majority of items on the LNNB IP scale, and the observation that most subjects of better than average intelligence make few errors on the scale, resulting in an inadequate "ceiling" effect. Golden et al. (1980) acknowledge the limitation of the IP scale in evaluating intellectual abilities above the average level, but assert that for individuals of I.Q.'s up to 100, the results of the LNNB IP scale and the WAIS are directly comparable. Picker and Schlotmann (1982) concluded that:

> "the Luria - Nebraska Intellectual Processes scale is a fairly useful screening device for the assessment of intelligence with a non-brain damaged population. Its correlation with the WAIS is comparable to many widely used IQ screening devices (e.g., Peabody Picture Vocabulary Test - Revised, Dunn & Dunn, 1981, p. 63). As predicted by the Luria - Nebraska's authors, the IP scale is a much more sensitive device for subjects with average to below average intelligence than for those subjects with above average intelligence. The simplicity of many of its items fails to test the limits of cognitive capacities for above average subjects" (p. 123).

The above referenced studies indicate clearly that intellectual functions are tapped by the LNNB, and particularly so by the Intellectual Processes scale. There is evidence that the level of correlation between the LNNB IP scale and the WAIS I.Q.'s is significantly related to the functional level of the individual, with valid reflections of actual ability being obtained only for persons of average to below average ability. The other summary scales of the LNNB also correlate significantly with intellectual ability, however, and the specificity of the other summary scales as measures of non-intellectual abilities has been called into question by this finding (Chelune, 1982). In addition, other authors have questioned the construct validity of the individual summary scales, and have pointed out that the nature of the tasks, particularly the rather heavy verbal emphasis on instructions, task stimuli, and responses can result in inflated error scores on scales due to deficits totally unrelated to the stated function of the scale (Crosson & Warren, 1982; Delis & Kaplan, 1982; Spiers, 1981). For these reasons, the ability of the LNNB to reflect functional abilities other than, or in addition to, those related specifically to intelligence has become an area of considerable interest.

Comparisons with the Wechsler Memory Scale

An initial study of the relationship between the LNNB Memory scale and the Wechsler Memory Scale (Wechsler, 1945) Memory Quotient was conducted by Ryan and Prifitera (1982). These authors reported that in a sample of 32 psychiatric patients there was a significant correlation between these variables of .65, and that they shared a common variance of 42 percent.

McKay and Ramsey (1983) also reported a study of the relationship between the LNNB Memory scale and the Wechsler Memory Scale (WMS). These authors noted that the WMS Memory Quotient score actually represents both a raw ability score and an age correction factor, and that this would likely reduce the correlation with the LNNB Memory scale score, which does not incorporate any such age correction. In this study, 38 male alcoholic inpatients were administered both the LNNB and WMS, and raw scores from each scale correlated 0.82, and accounted for 67% of the variance in the two tests. They also investigated the relationship between the WMS subtests and the factor scores (McKay & Golden, 1981) for the LNNB Memory scale. Significant correlations between the LNNB total Memory scale and factor scores with the individual subtests of the WMS were obtained for all WMS subtests except for Information and Orientation (see Table 11).

Table 11

Correlation of WMS and LNNB Memory Measures

WMS Scale	LNNB Memory Scale Total	LNNB Verbal Memory Factor	LNNB Visual & Complex Memory Factor
Information	-.18	-.22	-.19
Orientation	-.15	-.09	-.16
Mental Control	-.31	-.36*	-.21
Logical Memory	-.42**	-.45**	-.43**
Digit Span	-.67****	-.56***	-.48**
Visual Retention	-.56***	-.45**	-.64****
Associate Learning	-.57***	-.51***	-.47**
WMS raw score	-.82****	-.76****	-.75****
WMS MQ	-.60****	-.58****	-.47**
Digits Forward	-.48**	-.40*	-.29
Digits Backward	-.73****	-.65****	-.61****

* $p < .05$ ** $p < .01$ *** $p < .001$ **** $p < .0001$

Source: McKay and Ramsey (1983)

A factor analysis of the correlational data revealed two significant factors, with the first representing a "general retention" factor, and the second reflecting "attention and concentration" abilities. The LNNB Memory factor scales (the Verbal Memory and the Visual and Complex Memory scales) both loaded primarily on the general retention factor, with moderate secondary loadings on the attention and concentration factor. The WMS subtests displayed a significant pattern of divergent factor loadings, with the Mental Control and Digit Span factors loading heavily on the attention and concentration factor, while the Logical Memory and Associate Learning subtests loaded exclusively on the general retention factor, and Visual Retention displaying intermediate loadings on both factors (See Table 12).

Table 12

Factor Loadings of LNNB and WMS Subtests with Orthogonal (Varimax) Rotation

Scale	Factor 1	Factor 2
Mental Control	-.02	.80
Logical Memory	.72	-.08
Digit Span	.23	.85
Visual Reproduction	.47	.45
Associate Learning	.71	.09
LNNB Memory: Verbal Memory Factor	-.71	-.49
LNNB Memory: Visual & Complex Memory Factor	-.78	-.38

Source: McKay and Ramsey (1983)

The authors concluded that the LNNB Memory scale provides a reasonably good estimate of the WMS raw score, and that the scale is sensitive to both the general retention and attention/concentration factors that have previously been described in factoral analyses of memory tests (Kear-Colwell & Heller, 1978). The data also suggested the LNNB scale is more reflective of the general retention factor than the attention/concentration factor, and that both of the LNNB factor scales loaded on the two factors in a similar manner. An implication of this result is that in comparison to the WMS, the LNNB Memory scale may be less susceptible to attentional deficits that often interfere with immediate recall tasks, without affecting actual retention of material once perceived.

Comparisons with Academic Achievement Tests

The relationship of the LNNB with measures of academic achievement has also been studied. Hale and Foltz (1982) reported that a slightly modified form of the LNNB Pathognomonic scale proved to be the best single predictor of academic attainment (as measured by the Wide Range Achievement Test) in a population of educationally handicapped adolescents aged 12 to 16, and that the WISC-R Full Scale I.Q. did not account for significant variance in achievement over and above that already accounted for by the Pathognomonic score. These authors concluded that these results suggested that the neuropsychological information available from the entire LNNB battery would likely prove useful for educational planning purposes, as even an abbreviated selection of the LNNB items was able to demonstrate a higher degree of predictive relationship with academic achievement than the WISC-R Full Scale I.Q. score.

Shelly and Goldstein (1982) reported a study of 150 male, neuropsychiatric inpatients who were assessed with the LNNB, WAIS, and the WRAT, which allowed for the investigation of the relationships of the LNNB with both intellectual and academic abilities. Significant correlations were obtained between the WRAT Grade level scores and the LNNB summary scales, but in each case the

WRAT academic test (Reading, Spelling, or Arithmetic) correlated most strongly with the corresponding LNNB academic scale (Writing, Reading, or Arithmetic). A factor analysis of the correlational data (See Table 13) between these three test batteries revealed that the LNNB, intellectual, and academic variables tended to "load on some of the same factors in a readily interpretable manner" (p. 166) such that measures of academic achievement derived from the WRAT loaded on the same factor as the academic scales of the LNNB and measures of visual-spatial and construction abilities such as the WAIS Object Assembly and LNNB Motor and Vision scales load together.

Table 13

Rotated Factor Matrix for LNNB, WAIS, and WRAT Variables

Variable	Factor Loadings			
	Factor 1	Factor 2	Factor 3	Factor 4
LNNB Scales:				
Motor	.70	.20	.16	.53
Rhythm	.48	.31	.22	.52
Tactile	.62	.24	.16	.46
Visual	.66	.15	.16	.50
Receptive Speech	.40	.28	.39	.64
Expressive Speech	.35	.33	.37	.63
Writing	.32	.62	.19	.55
Reading	.10	.67	.20	.53
Arithmetic	.39	.55	.29	.42
Memory	.52	.26	.34	.41
Intellectual Processes	.49	.28	.51	.49

Table 13 (cont.)

Variable	Factor Loadings Factor 1	 Factor 2	 Factor 3	 Factor 4
WAIS Scaled Scores:				
Information	.33	.46	.61	.22
Comprehension	.28	.25	.74	.22
Arithmetic	.41	.54	.40	.16
Similarities	.40	.30	.54	.28
Digit Span	.28	.52	.25	.28
Vocabulary	.18	.42	.81	.19
Digit Symbol	.67	.18	.25	.27
Picture Completion	.63	.26	.42	.30
Block Design	.76	.29	.17	.09
Picture Arrangement	.63	.22	.30	.27
Object Assembly	.78	.09	[illegible]	.11
WRAT Grade Level:				
Reading	.10	.80	[illegible]	.17
Spelling	.20	.81	[illegible]	.13
Arithmetic	.54	.59	[illegible]	.05
Percent Explained Variance	81.58	9.75	4.64	4.03

Note: Luria-Nebraska Scores have been reflected so that lower scores indicate poorer performance.

Source: Shelly and Goldstein (1982)

These authors concluded from the factoral analysis of their data that:

> "the Luria - Nebraska battery on the one hand and the WAIS and the WRAT on the other appear to be tapping into common domains utilizing different methods. The implication of this finding would appear to be that both sets of procedures are essentially equally equipped to test the same neuropsychological hypotheses.... from the points of view of loading on the same factors, as described above, and the high correlations between Luria - Nebraska and WAIS variables, now found in three independent studies, the Luria - Nebraska battery may be viewed as assessing several of the same domains of intellectual functioning as the WAIS. The present study also supports the view that the Luria - Nebraska shares a common domain with an established achievement test, the WRAT."

Critiques of the Luria-Nebraska Battery

As with any new major psychometric instrument, the LNNB has received a considerable degree of attention in the literature, and several authors have published critical reviews of it, in some cases accusing the battery's authors of releasing it for publication and clinical usage prematurely (Adams, 1980a, 1980b, 1984; Spiers, 1981). The earliest such published criticism was that of Adams (1980a), who expressed a number of concerns and allegations about the LNNB. Following a somewhat lengthy discourse in which he soundly criticized the lack of objectivity and replicability of Luria's behavioural neurological approach to evaluation, Adams then outlined a series of "problems" or charges against the LNNB that he felt compromised its validity as a clinical instrument or test.

The first specific objection raised was that: "There are many items in this particular Luria protocol that are quite subjective (e.g., "show me how" or "score for quality") and certainly cannot be seen as tests" (p. 512). He went further to assert that the utilization of such "subjective" items on the LNNB was a problem as: "The many variations in the qualitative aspects of patient performances make actual evaluation more difficult than manuals or instructions typically anticipate. That is, the exceptions to the clinical scoring systems quickly exceed the expected range of responses in clinical use" (p. 512). The net result of this potential source of error would be a reduction in the reliability of the scoring process for individual items. This particular criticism was responded to by Golden (1980) who noted that while some items on the LNNB are scored for the subjective dimension of quality (particularly drawings) these items are scored according to a detailed set of objective standards or criteria which are fully explained in the test manual. The fact that the interrater reliability of the LNNB on these items was better than 95% (Golden et al. 1978) was also pointed out as well as the fact that other widely used and accepted tests such as the Wechsler Adult Intelligence Scale also employ items scored according to qualitative (albeit structured) criteria.

Adams (1980a) also asserted that there was a contamination of the procedures used to determine the reliability of the scoring system by the: ""adjusting" to the independent variable of group membership" a charge that was denied by Golden (1980), and which is refuted by the test manual, in that all subjects are scored on items according to a preset standard or criteria, and no allowance for altering the scoring of item responses according to diagnostic category is made.

A further assertion by Adams (1980a) is that on those items that are scored according to objective criteria (e.g., time or errors) have also been compromised by the: "resorting to relatively ordinal summary indices." Each item on the LNNB is scored and rated according to a three point system, with a score of "0" representing normal performance, a "2" representing clearly abnormal performance, and a score of "1" reflecting an intermediate or borderline performance level. Adams then asserts that the use of such a scoring procedure: "takes the neuropsychological testing process out of the realm of quantified performance and provides no public way of inspecting or comparing individual or cohort differences on specific tasks." This criticism was also addressed by Golden (1980) who noted that this allegation was "meaningless" in that the transformation from raw score to scaled score is objectively determined and does not affect the objectivity or quantification of the score in any manner. Golden also noted that in the initial development research on the LNNB there was an evaluation of the effect of using the three point scoring system versus raw score data, and that no loss of diagnostic power took place in adopting the simpler three point scores. The fact that there is some constriction on the range of scores that each item can display and a resultant potential loss of descriptive power would seem inconsistent with the above reviewed validational studies on the discriminative power of the LNNB to identify neurological patients. As Golden (1980) noted: "It might further be noted that if the current high hit rates reflect a "loss of power," then the effectiveness of the test must actually be higher than we have reported" (p. 518).

Adams (1980a) again criticized the use of "adjusting" cutoff scores to the independent variable of group membership in the Purisch et al. (1978) study and described this as a "mistake" in procedure. Golden (1980) noted that this procedure of publishing optimal cutoff scores for differing diagnostic groups is a common practice and there is no clear reason for criticizing this or classifying it as a mistake.

A further discourse made by Adams (1980a) was to attempt to demonstrate: "the problems encountered in the display of ordinal summary indices" by comparing means and standard deviations of the schizophrenic group of the Purisch et al. (1978) standardization sample with the scores reported for what he claimed was: "the same information for a comparable sample and technique recently reported by Diamant and Hijman." [illegible]nted table of scores displayed radically different means and standard deviations for the two studies, which Adams claimed made: "it difficult to directly determine whether the method can be considered reliable across clinical settings" (p. 512). Golden (1980) addressed this criticism by noting that the Diamant and Hijman study did not use the LNNB test at all, but rather some other procedure (actually the Luria's Neuropsychological Investigation procedure described by Christensen, 1975) and that the scores reported for this other study are actually incompatible with the LNNB in that the reported mean score on some of the scales actually exceed the maximum score on the related LNNB scale. It would appear obvious that the fact that the means and standard deviations for similarly named scales on two separate test batteries are not similar does not necessarily imply any lack of validity or reliability for either test.

Subsequent complaints raised by Adams (1980a) included a criticism that the subjects selected for the validation studies were quite diverse in terms of age and education and that the neurological diagnoses of the brain-injured group and the criterion tests were not reported in detail. Golden's response to these concerns was to note that while Adams criticized the existence of educational differences between the subject groups he had failed to comment on the fact that an analysis of covariance controlling for these educational differences failed to reveal any significant relation between education and individual item performance (Golden et al., 1978, p. 1264) and that despite Adams (1980a) implication that there were significant differences between the groups with respect to age, this was incorrect and no such age differences between groups actually existed. Thus, the subject selection criticisms with respect to both age and educational level were observed to be of little substance or significance. Golden (1980) also indicated that detailed descriptions of the patient diagnoses, neurological test results, etc. were also not reported in the initial validational studies simply due to space restrictions, and he confirmed that no patient was included in the brain damaged group for the validational studies: "that was not confirmed neurologically by tests- mostly computerized axial tomography (CAT scan), electroencephalogram (EEG), or angiogram- or history (e.g., individuals with head traumas causing unconsciousness for more than 24 hours)".

A criticism by Adams (1980a) that the psychotropic medication levels were not reported in either validation study was also addressed by Golden (1980) who noted that a paragraph regarding this had been edited out while the manuscript was in review, and that all patients were tested while on whatever medications their physicians had prescribed, [illegible] schizophrenic patients generally taking one of the antipsychotic medications in a "wide variety of doses." This testing condition was felt to most closely conform to the typical testing situation facing a clinical neuropsychologist, and Golden also noted that subsequent evaluations of schizophrenic performance on the LNNB "failed to find a relation between past or present medication and Luria performance" (p. 518).

Adams continued with noting that the chronicity, number of previous hospitalizations and age at onset differ significantly between the brain-injured and schizophrenic groups and he also suggested that the brain-injured group in the these studies might be a mixed psychiatric and brain-injured population from a previous study (Golden, 1977). Golden (1980) denied that the brain-injured group in the two validational studies was in any way related to the group of neuropsychiatric patients reported in the 1977 study, and he reiterated that "no patient with a history of any serious psychiatric condition was included in the normal or brain-damaged group" and that: "We did not include in the study any schizophrenic for whom there was a suspicion from history, EEG, physical neurological examination, or doctor's opinion that the patient was brain damaged." The significant differences between the schizophrenics and neurological patients on chronicity, number of hospitalizations, and age of onset was described as intentional, and it was noted that the effect of these differences would have been to make it more difficult to discriminate between the schizophrenic and neurological groups, thus: "offering a maximally difficult test for the battery" (Golden, 1980, p. 519). In a separate investigation (Lewis, Golden, Purisch, & Hammeke, 1979) it was reported that length of hospitalization and chronicity of schizophrenic disorder failed to show significant effects (both main effects and interaction effects) on a multivariate analysis of variance on the 14 summary indices from the LNNB, which further weakened the significance of this particular criticism of the initial validational studies.

Adams (1980a) next criticized the methodology used in analyzing the significance of the initial pool of 285 items in discriminating between neurological and medical control groups. He also asserts that this methodological error was compounded by the task of making: "the essential discrimination being made is probably between very seriously neurologically impaired (and, perhaps, psychiatrically disturbed) patients and medical controls." He added that the reported finding that nearly 90% of the ratings of the tasks as useful and sensitive indicators of brain dysfunction: "may only reflect the more gross distinction between patients with suspected cerebral damage and medical comparison subjects" (p. 514). Golden (1980) acknowledged that the use of 285 sequential t tests to assess the individual item significance was less than an ideal manner to analyze data, but he justified this as a procedure for screening the battery for ineffective items, which were dropped from later versions of the test. He also rebutted again Adams assertion that the brain-damaged group included any subjects with psychiatric disturbance, and also noted that no patient was included with seriously acute disorders that interfered with ability to follow test instructions, give informed consent, cooperate with testing procedures, or whose disorder interfered with consciousness. It was further reported by Golden (1980) that over 80% of the neurological patient group was tested at least six months after their initial injury, and that 30% were tested more than 5 years after the onset of their neurological condition, allowing ample time for most of the neurological patients to show recovery from acute symptoms. The allegation that the neurological patient group was unusually impaired or more severely dysfunctional than is typical for a neurological patient population was adamantly refuted by Golden (1980).

A final criticism raised by Adams (1980a) was the assertion that both validational studies: "report discriminant functions using a base in which the number of variables is nearly three times the number of subjects. This is a gross violation of discriminant technique (Adams, 1979), which renders the claims of efficient classification in both articles invalid and casts the conclusions about the utility of the battery into doubt" (p. 514). This final specific criticism was countered by Golden (1980) who pointed out that while it is clearly true that a discriminant function run on 285 items for a sample of 100 would be meaningless, the actual procedure used in the two validational studies was to run the discriminant analyses using only the best discriminating 30 items (Golden et al., 1978) or best 40 items (Purisch et al., 1978) which placed the subjects to variable ratio into an acceptable range. Golden also noted that Adams (1980a) ignored the report of the discriminant analysis using just the fourteen summary variables, which yielded an 88% diagnostic accuracy in the Purisch et al. (1978) study, and which could not be ruled as invalidated by subject to variable ratio problems. Golden also noted that the test manual also reported the results of two additional cross-validational studies using entirely new patient populations in which the 14 summary indices displayed hit rate accuracies of 87% and 93% (Moses & Golden, 1979; Moses & Golden, 1980). As noted in the preceding section of this study, subsequent validational research by independent investigators has also confirmed the validity of the LNNB to discriminate between neurological and nonneurological groups using the summary scale variables.

Overall, the variety of criticisms and complaints raised in the Adams (1980a) article were rather consistently responded to by Golden (1980) in a convincing manner, and in many instances the particular assertions and statements regarding the "problems" inherent in the validational studies are described by Golden as inaccurate or not based on the facts of the actual studies. While the critique by Adams was particularly vitriolic in its tone and stated censure of the LNNB, it did little to provide any specific basis for objectively evaluating the validity or utility of this new battery.

Spiers (1981) noted that the previous debate between Adams and Golden in 1980 had raised some doubts as to the adequacy of the validational procedures and statistical methodology used in standardizing the battery, and he also expressed a concern that a major need in the validational process is to establish the validity and utility of the content of the test, with a need to: "examine the actual composition of the Luria - Nebraska and evaluate whether the items chosen are likely to provide information about the functions that they purport to test". Spiers initiates his review with a general observation that the LNNB is composed of only 248 actual tasks, and that 26 out of the total of 269 item scores are separate scorings for a preceding item, typically the time taken to perform the task. He adds that this would mean that subjects with: "initiation, anxiety, or motivation problem may fail almost 10% of the items on this test, strictly because they require more time to produce a response" (p.p. 332-333). He also criticized the scoring system in that no allowance is made for crediting an item answered correctly after the expiration of time limits, such that it is possible for a patient to receive the highest or most impaired score while still giving the correct answer. He also notes that Luria's clinical investigation approach was primarily concerned with determining qualitatively <u>why</u> a function was impaired, rather than attempting to provide a summary or quantitative index of how much it was impaired. Spiers also criticized the use of the three point item scoring system on the basis of it being "meaningless" to attempt to equate: "independent, clearly localizable activities and functions by the application of scaled scores".

The balance of the Spiers (1981) article is devoted to what he describes as a "content analysis" of a selection of 6 of the 11 LNNB clinical scales. In these subsequent discussions Spiers describes the item content of the various reviewed scales and classifies each of the scale items according to what he feels the item reflects, with a relatively high number of items reflecting (in his assessment) functions other than that specified by the scale name or title. For example, in discussing the Motor scale, he asserts that: "only 14 of these 51 items are unequivocally related to motor functioning" and that of the remaining items 4 are more closely related to "position sense", 15 are measures of representational, nonrepresentational and bucco-facial praxis, 6 are more closely related to right-left orientation and 12 are visuo-spatial constructional tasks. Having summarily dismissed the majority of the Motor Scale items as not assessing motor functioning he then asserts that remaining items: "simply do not provide a sufficient sampling of motor behavior and do not reliably or adequately measure motor speed, dexterity, alternation, oral agility or written motor alternation, given the small number of trials administered on each task" and adding that: "the majority of items making up the Motor Functions scale may be failed for reasons having nothing to do with motor systems compromise" (p. 334). Similar criticisms of item content and composition was tendered for the Tactile, Receptive and Expressive Speech, Memory and Intellectual Processes scales.

Spiers (1981) discussion of the LNNB and his "content analysis" of the individual items was marked, however, by a lack of specification of what method was used in analyzing the actual content of the items, and no objective basis or technique for assigning the individual items into the categories he described was specified. His determination of what the individual items were measuring was apparently based entirely upon his subjective evaluation of the item's content or functional demand. His criticism of the item testing procedures was also criticized by Hutchinson (1984), who noted a variety of errors in Spiers' analysis, including a substantial number of "errors of fact," "errors of interpretation or understanding," and "errors in reasoning." While Spiers was especially critical of the LNNB scale content and the fact that each clinical summary scale contained items that could be impaired by a variety of factors in addition to deficits in the specific functional system identified for the scale, Hutchinson noted: "Finally, content validity criticism is not necessarily a reliable or valid method for refutation of empirically derived item correlations. Such refutation is best accomplished by presentation of empirical counterevidence" (p. 540). In conclusion, Hutchinson noted that Spiers did make a limited number of valid criticisms about the LNNB, including the fact that the Memory scale was composed only of immediate memory items and lacked coverage of delayed or remote memory, that some items are potentially influenced by peripheral sensory impairments such as colour blindness and that the initial marketing of the LNNB may have overstated the case in describing the battery as assessing "all major neuropsychological functions." It was suggested, however, that the actual utility and content validity of the LNNB scales and items would be better evaluated by careful empirical study rather than by subjective speculation or casual perusal of the test items.

Spiers (1981) concerns with the content validity of the LNNB have also been shared by other authors. Crosson and Warren (1982) criticized the battery for its heavy reliance on verbal instructions and response formats, and noted that aphasic deficits could conceptually give rise to inflated estimates of impairment on nonverbal scales (e.g., Visual Functions). The coverage and content of the speech scales were also criticized by Delis and Kaplan (1982) on a conceptual basis as they were not oriented to commonly accepted theories of aphasia and provided a case example of how the LNNB summary scores could contribute to misdiagnosis of an aphasic syndrome. The major criticisms of both of these critiques were that the content validity of the scales was confounded by the fact that assessment of cognitive and motor functioning is dependent upon the integrity of other functions (most notably language) and that the assessment of primary cognitive functions (such as confrontation naming) is inadequate because of the use of too few items. Delis and Kaplan (1983) also noted that while the LNNB items were initially selected on the basis of their ability to discriminate between brain-damaged and nonbrain-damaged patients, and that no attempt was made to evaluate either the items or the summary scores against other measures of the functions that they purport to assess. They noted further that: "For the patient who suffers from known neurological involvement... knowledge of whether the LNNB can predi[illegible]ge, is insignificant. Rather, what is crucial is an [illegible]ssment of the patient's spared and impaired [illegible] and delineation of the mechanisms underlying the [illegible]

[illegible]on the criteria of discriminating between n[illegible] and nonneurological hospitalized patients for inclusion of items in the LNNB was also criticized by Stambrook (1983), who noted that: "While such a procedure may make the battery relatively efficient in diagnosing brain damage, it is not consistent with the goal of developing an instrument that is capable of providing a comprehensive assessment of neuropsychological functions that allows for the isolation of defects in functional systems or for the development of rehabilitation programs" (p. 253). Stambrook also shared with other evaluators concerns about the adequacy of coverage of the LNNB, such as in the notable lack of assessment of reading comprehension on the LNNB Reading scale and the failure to include measures of recent or remote memory functions on the Memory scale. He did, however, note that a variety of studies related to the battery's criterion-related (diagnostic) validity had been supportive of the LNNB's utility in diagnosing the presence of brain damage and in discriminating between neurological and psychiatric patients. In contrast to the essentially conceptual criticisms offered by Spiers (1981) and Delis and Kaplan (1982, 1983), Stambrook also noted that the construct validity of the battery can be evaluated by reference to other indicators, such as the internal consistency of the summary scales, factor analysis of the battery and correlations of the battery with other instruments.

With respect to the issue of internal consistency of the summary scales, Golden, Ariel, Moses, Wilkening, McKay and MacInnes (1982) asserted that while the summary scales are not homogeneous and are in fact purposefully heterogeneous in item content, the items of each individual summary scale tap the same general construct. This assertion was examined empirically by Golden, Fross, and Graber (1981) who reported that of the 269 items of the battery, 250 correlated most highly with the scale on which they were placed than with any other summary scale. These authors also reported that split-half reliabilities for the eleven summary scales ranged from 0.89 for the Memory scale to 0.95 for the Reading scale, indicating a rather uniformly high level of internal consistency for all summary scales. Stambrook (1983) noted, however, that these results were insufficient to allow for an evaluation of the homogeneity of the scales in terms of their underlying constructs as the magnitude and patterns of the individual item correlations are not reported.

A different approach to evaluating the construct validity of the LNNB has been the factor analytic study of the items of the summary scales, which Golden and his colleagues reported in a series of investigations (e.g., Golden, Purisch, Sweet, Graber, Osmon, & Hammeke, 1980; McKay & Golden, 1981)~~. While~~ these factor analytic studies indicated that, for the most part, the factor structures of the summary scales were generally consistent with what would be predicted by Luria's (1980) theory of brain organization, there were a number of methodological problems which have been cited that raise questions about these results. One concern has been the use of factor analytic methodology on ordinal scales with few categories (c.f., Comrey, 1978) but a more potent criticism was that these factor analytic studies were conducted on mixed samples of 90 neurologically intact, 90 psychiatric, and 90 brain-damaged patients, with no consideration of the likelihood that the factor structures could differ for each of these patient population groups (Spiers, 1982). Supporting this concern is the observation that in a separate factor analytic study using a different proportion of patient types (105 neurological intact, 94 psychiatric, and 218 brain-damaged) in a re-examination of the Receptive Speech scale McKay and Golden (1981) reported a markedly different factor structure (seven factors identified) compared to the initial structure of only two factors in the Golden, Purisch, Sweet, Graber, Osmon, and Hammeke (1980) study. While there was an overlap in subject composition between the two studies (207 patients from the initial factor analysis study were included in the second investigation) there was a substantially higher proportion of brain-damaged subjects in the second study, and the disparity between the two factor structures is suggestive of a different factor structure for brain-damaged than for neurologically intact and psychiatric patients. Based upon a consideration of these methodological issues, Stambrook (1983) concluded:

> "these studies cannot be used to speak to the construct validity of the summary scales of the LNNB. A factor analysis of the items is possible if the items are scored with many categories or if the summary scales are used as the data points, although different issues would be addressed by each approach. In this light, it seems unfortunate that Golden (1980) chose to forgo a five-point scoring system in favour of one which was dichotomous and trichotomous." (p. 262)

An alternative approach to construct validity is to examine how well a new instrument correlates with well established instruments, with the intent of comparing the construct validity of the new measure with that of others in current use (Anastasi, 1976). Golden, Kane, Sweet, Moses, Cardellino, Templeton, Vicente, and Graber (1981) reported that using a multiple correlation procedure it was observed that each of the 14 major LNNB scales correlated significantly and highly with 14 major variables from the Halstead - Reitan Neuropsychological Battery, with a shared variance of approximately 73 percent between the two neuropsychological test batteries. Similar indications of a high degree of overlap or shared variance between the LNNB and HRNB has been observed in demonstrations of essentially identical hit-rates for the two batteries in a linear discriminant analysis between the groups of control and neurologically impaired subjects in this study. Comparable hit-rates for both batteries were also reported using a blind, expert, clinical interpretation of the profiles in a mixed psychiatric and brain-damaged sample (Kane, Sweet, Golden, Parsons, & Moses, 1981).

Stambrook (1983) noted in a review of this construct validational research that while the shared variance between the LNNB and HRNB would suggest that both are assessing the same general set of functions, it is also necessary to determine how well the individual LNNB summary scales assess the constructs that are thought to underlie each scale (e.g., intelligence, memory, reading, etc.). While the LNNB Intellectual Processes scale is observed to correlate highly with the Wechsler Adult Intelligence Scale (WAIS), the specificity of this finding is reduced by the observation that all the other LNNB summary scales also correlate significantly with WAIS I.Q.'s, with reported correlations ranging from -.53 for the Tactile scale to -.80 for Arithmetic, with a mean of -.70 (McKay et al., 1981). Thus, while the Intellectual Processes scale of the LNNB reflects psychometric intelligence (as defined by the WAIS), so to do the balance of the other summary scales to a significant extent, a finding that may call into question the validity of separating them into distinct scales. It has also been observed that if the WAIS I.Q. scores are excluded from the correlations between the HRNB and LNNB, the next strongest predictor of the LNNB summary scales is the HRNB - Aphasia Screening Test (Golden, Kane, Sweet, Moses, Cardellino, Templton, Vicente, & Graber, 1981). As the Aphasia Screening Test is a collection of items that primarily assesses various aspects of receptive and expressive speech skills, the high correlations with all of the LNNB summary scales has been interpreted as indicating that language skills are involved in the performance of all of the tasks tapped by the LNNB (Stambrook, 1983), a finding that supports the contention that scales that are not specifically intended to assess language functions (e.g., the Motor, Tactile, Visual scales, etc.) are subject to contamination by the reliance upon language skills (Crosson & Warren, 1982; Delis & Kaplan, 1982; Spiers, 1981) such that for aphasic patients these scales may reflect their language disorder, rather than the functions that they have been named for. Stambrook (1983) summarized these concerns by noting:

> "Additional research, using patient groups that differ in their presentation of neuropsychological symptom patterns, is needed before it would be possible to suggest the nature of the constructs that underlie the summary scales on the LNNB. Although the data suggest that the LNNB and the HRNTB overlap considerably in the skills that they assess, and that the LNNB measures much of what is considered to be psychometric intelligence, little is known regarding the construct validity of the LNNB summary scales (the exception being the Intellectual Processes scale)." (p.264)

CHAPTER III

RATIONALE AND HYPOTHESES

Rationale

From consideration of the literature reviewed in the previous chapter regarding both the area of learning disabilities and the neuropsychological assessment approach as represented by the Luria - Nebraska Neuropsychological Battery, it is possible to derive a number of general impressions and conclusions. It is noted that the general emphasis in the learning disabilities area has been directed toward establishing a logically consistent definition of these types of learning disorders, and to define specific criteria for diagnosing the presence of such a disability. This has often been based upon a process of diagnosis by exclusion, by indicating that the learning disabled child is one that has normal intellect, sensory/perceptual functions, cultural and educational exposure, etc., yet still is displaying a deficit in academic attainment. However, as demonstrated in the Taylor, Satz and Friel (1979) study, learning delayed students who fail to meet one of more of these exclusionary criteria are not observed to differ in any identifiable way with respect to patterns of academic performance deficits, family or personal health and risk factors, or in neuropsychological test performance.

A more recent approach has been to shift the focus in the area of learning disabilities from diagnosis by exclusion to the adoption of more inclusionary criteria, particularly through incorporating into standard definitions some statements regarding the presence of identifiable or inferred neurological dysfunctions as underlying the learning disability (c.f., AEAC, 1984). A further trend and emphasis in recent research in the area of learning disabilities has been to search for and identify meaningful subgroups of learning disabled children, with reference to identifying the underlying deficits in processing and neuropsychological functioning that distinguishes the disparate types of disability (c.f., Rourke, 1985).

What is inherent in this shift in emphasis from an exclusionary diagnostic process to the identification of unique patterns of underlying neuropsychological impairments to be used to subgroup learning disabled children is that there is a need for a reliable and valid instrument to evaluate these underlying deficits for the individual child. Thus, if the theoretical position is that learning disabilities are caused by or are primarily attributable to patterns of selective neuropsychological deficit that may be different for individuals, it is necessary to have valid neuropsychological assessment procedures for identifying such individual differences in the underlying processes critical to educational attainment.

While the reviewed literature on the Luria - Nebraska

Neuropsychological Battery indicates that this new battery has adequate basic reliablity with respect to the scoring of the individual items and internal consistency of the scales and there is a growing body of evidence that supports its validity to be able to diagnose and identify the presence and general location of brain damage in adults, there are several areas on which it has been criticized, particularly the construct validity of the individual clinical scales. With respect to the value of such a neuropsychological battery. for the identification of specific sub-groups of learning disturbance, it is especially important that the instrument be shown to have adequate construct validity with respect to its scale composition, and that there are reliable correlations between it and other measures that are commonly accepted as measures of constructs such as intelligence and school achievement.

Related to this, it is necessary to establish that the individual scales of the LNNB are both internally consistent, and that there is a high correlation between the scales and other measures of similarly labelled functions. As previously noted, the summary scales of the LNNB have been criticized for their heterogeneous nature, and for the over-reliance upon language functions on scales that are not specifically oriented to language abilities. It, has also been criticized by some for the high correlation of all scales with psychometric intelligence, with the suggestion that the separation of the items into distinct scales is artifactual and not based upon demonstrated differences in underlying functional constructs.

In addition to these criticisms of the LNNB construct validity (which would be especially important for its use in making individual assignments to separate learning disability subgroups), it is also noted that while this scale has been promoted and marketed as useful for adolescents down to the age of 13 years, there is relatively little reported in the literature regarding its use in this age range. Particularly as the factoral structure of individual summary scales has been shown to be dependent upon the sample characteristics, the issues of internal consistency and construct validity raised in the use of this scale with adult populations are even less assured in the adolescent age range.

Hypotheses

Based upon the preceding rationale, the utility of the Luria-Nebraska Neuropsychological Battery for an adolescent age population was investigated, as well as its validity for identification of learning problems or deficits. The basic applicability of this battery for use in the adolescent age range was first investigated by examination of the psychometric property of reliability. This was approached by examining the reliability of the individual summary scales through an internal consistency approach (Cronbach's Alpha) and through evaluating the individual items of the battery with respect to their correlation with their own scale as well as other summary scales to establish the level of overlap and degree of scale related specificity of the items. Hypotheses based upon these initial investigations were as follows:

Hypothesis 1

The clinical scales for the LNNB will display internal consistency reliability estimates for the adolescent aged sample that are equivalent to published estimates of mean reliabilities from adult investigations.

Hypothesis 2

The clinical scales for the LNNB will display levels of consistency between the item to scale correlations in this adolescent sample that are equivalent to those reported in adult investigations.

Based upon the findings of these initial investigations into the psychometric properties of the LNNB in the adolescent age range, it was decided to identify an abbreviated version of the battery, based upon a selection of the LNNB items that would allow for maximizing both the internal consistency of the clinical scales as well as the level of consistency of item to scale correlation. This was achieved by eliminating most items that either displayed inadequate variance or correlation with the scale they were located in or which displayed stronger correlations with other than their assigned scale. Hypotheses formulated regarding the psychometric properties of this abbreviated LNNB (the LNNB-A) were as follows:

Hypothesis 3

The internal consistency reliability estimates of the LNNB-A clinical scales will be equal to or higher than those of the full LNNB for this sample.

Hypothesis 4

The level of item to scale correlation consistency for the LNNB-A will be significantly improved from that of the LNNB, and will be equivalent to such consistency estimates for the LNNB in adult populations.

As a component of the investigation of the construct validity of the LNNB and LNNB-A batteries, a factor analysis was conducted to explore the dimensional structure of the tests. Hypotheses formulated regarding the factor analysis of the LNNB and LNNB-A were as follows:

Hypothesis 5

Principal components factor analysis of the LNNB will reveal the presence of a significant primary factor related to level of performance or impairment level, with strong loadings of all clinical scales on this primary principal component factor.

Hypothesis 6

In addition to the loadings on an overall general ability factor, group factors will be identifiable for the LNNB that will correspond to neuropsychologically meaningful dimensions in the data, particularly with respect to idenification of factors related to verbal and nonverbal cognitive processes.

Hypothesis 7

Principal components factor analysis of the LNNB-A will reveal the presence of a significant primary factor related to level of performance or impairment level, with strong loadings of all clinical scales on this primary principal component factor.

Hypothesis 8

In addition to the loadings on an overall general ability factor, group factors will be identifiable for the LNNB-A that will correspond to neuropsychologically meaningful dimensions in the data, particularly with respect to identification of factors related to verbal and nonverbal cognitive processes.

Hypothesis 9

The factor structures of the LNNB and LNNB-A will be equivalent and display no significant loss of coverage of neuropsychological functions from the abbreviation of the test battery.

The concurrent validity of the LNNB Intellectual Processes and academic scales (Writing, Reading, and Arithmetic) was also examined to determine their usefulness in describing these educationally important abilities in an adolescent population. This was investigated initially by examination of the intercorrealtions between the scores from the LNNB, WISC-R, and WRAT. Hypotheses related to this correlational investigation were as follows:

Hypothesis 10

The LNNB Intellectual Processes scale will correlate significantly with the I.Q. scores from the WISC-R, and it will correlate higher with WISC-R estimates of intelligence than any of the other LNNB clinical scales.

Hypothesis 11

The LNNB Writing scale will correlate significantly with the Spelling grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Spelling grade score and any other LNNB scale.

Hypothesis 12

The LNNB Reading scale will correlate significantly with the Reading grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Reading grade score and any other LNNB scale.

Hypothesis 13

The LNNB Arithmetic scale will correlate significantly with the Arithmetic grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Arithmetic grade score and any other LNNB scale.

The degree of construct validity of the identified selection of items comprising the educationally related scales of the LNNB-A was also investigated with respect to the correlations between the LNNB-A Intellectual Processes scale and the WISC-R I.Q. scores and between the LNNB-A Writing, Reading, and Arithmetic scales and the related scales from the WRAT. Hypotheses relating to these correlation comparisons were as follows:

Hypothesis 14

The LNNB-A Intellectual Processes scale will correlate significantly with the I.Q. scores from the WISC-R, and it will correlate higher with WISC-R estimates of intelligence than any of the other LNNB-A clinical scales.

Hypothesis 15

The LNNB-A Writing scale will correlate significantly with the Spelling grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Spelling grade score and any other LNNB-A scale.

Hypothesis 16

The LNNB-A Reading scale will correlate significantly with the

Reading grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Reading grade score and any other LNNB-A scale.

Hypothesis 17

The LNNB-A Arithmetic scale will correlate significantly with the Arithmetic grade score from the WRAT, and this correlation will be higher than the correlations between the WRAT Arithmetic grade score and any other LNNB-A scale.

A further approach to explore the educational utility of the LNNB and LNNB-A batteries was to determine the ability of these neuropsychological test batteries to reflect educational attainment by use of multiple regression analyses, with calculation of multiple correlation coefficients and identification of which clinical scales are able to contribute to the prediction of intellectual and academic test scores. Hypotheses arising for this investigation are as follows:

Hypothesis 18

Muliple regression formulae will be derived that allow for highly significant prediction of WISC-R I.Q. and WRAT grade score values from the LNNB clinical scale data. For the purposes of this hypothesis a correlation in excess of 0.70 will be considered to be evidence of a high level of predictive validity.

Hypothesis 19

Muliple regression formulae will be derived that allow for highly significant prediction of WISC-R I.Q. and WRAT grade score values from the LNNB-A clinical scale data. For the purposes of this hypothesis a correlation in excess of 0.70 will be considered to be evidence of a high level of predictive validity.

Use of the summary scale scores (from both the original and abbreviated neuropsychological batteries) for discriminating between a priori classified groups of adolescents with externally defined patterns of academic performance was investigated as a test of the utility of the measure for identifying the presence of different patterns of academic delay. Hypotheses formulated regarding this linear discriminant analysis investigation were as follows:

Hypothesis 20

Application of a linear discriminant function to the clinical scale data from the LNNB will allow for the discrimination between cases that have different patterns of academic attainment on the WRAT at a level higher than chance.

Hypothesis 21

Application of a linear discriminant function to the clinical scale data from the LNNB-A will allow for the discrimination between cases that have different patterns of academic attainment on the WRAT at a level higher than chance.

A final series of investigations was to use the summary scale scores from the LNNB and the LNNB-A in cluster analysis procedures to examine the utility of this instrument to identify educationally relevant homogeneous subgroups of learning delayed students. Hypotheses formulated for this cluster analytic portion of the study were as follows:

Hypothesis 22

Application of appropriate cluster analytic methodology to the clinical scale data from the LNNB will result in the identification of subgroups within the sample that display significant differences with respect to LNNB profile elevation and/or shape.

Hypothesis 23

Identified subgroups based upon the cluster analysis of the LNNB clinical scale data will demonstrate significant between group differences on the external criteria of WISC-R I.Q. scores.

Hypothesis 24

Identified subgroups based upon the cluster analysis of the LNNB clinical scale data will demonstrate significant between groups differences on the external criteria of WRAT grade scores.

Hypothesis 25

Application of appropriate cluster analytic methodology to the clinical scale data from the LNNB-A will result in the identification of subgroups within the sample that display significant differences with respect to LNNB-A profile elevation and/or shape.

Hypothesis 26

Identified subgroups based upon the cluster analysis of the

LNNB-A clinical scale data will demonstrate significant between group differences on the external criteria of WISC-R I.Q. scores.

Hypothesis 27

Identified subgroups based upon the cluster analysis of the LNNB-A clinical scale data will demonstrate significant between groups differences on the external criteria of WRAT grade scores.

These series of analyses were designed to establish the internal consistency of the clinical scales of the LNNB, the consistency of item/scale assignment for an adolescent population, the concurrent validity of several of the scales most relevant to educational achievement, and the factorial structure of this instrument for an adolescent population. The predictive validity of the battery for identifying externally defined subgroups of learning abilities was examined, as well as the ability of the scale to identify meaningful subgroups of learning delayed students. The intent of this study was to evaluate the applicability of the LNNB (either in its original form or in some modified or abbreviated form that is most appropriate for this age range) to an adolescent population, with a particular emphasis on establishing its reliability and validity as an instrument to measure academic learning ability and achievement, and to discriminate between students with normal and disordered academic learning abilities.

CHAPTER IV

METHOD

Subjects

Subjects for this investigation represented a total of 120 consecutive adolescent cases referred to the writer for clinical neuropsychological evaluation between the dates of October 1983 and October 1986. All cases represented an adolescent (between ages 13 and 16 inclusive) referral in which concerns were raised about combinations of academic, behavioural and emotional functioning, and for which complete data was available for the Luria - Nebraska Neuropsychological Battery (LNNB), Wechsler Intelligence Scale for Children - Revised (WISC-R) Verbal, Performance and Full Scale I.Q. scores, and the Wide Range Achievement Test (WRAT) scores for Reading, Spelling and Arithmetic subtests. The sample consisted of 24 females subjects and 96 male subjects, representing a male to female sex ratio of 4 to 1. This obtained sex ratio was observed to be consistent with that frequently reported among unselected clinical referrals of youths for both educational and behavioural problems, as most investigators and practitioners in the area of school psychology have noted a similar over-representation of males in groups of school aged children with identified problems in academic attainment (Brentzen, 1961) and which has been hypothesized to be reflective of sexually related superiority of left hemispheric language functioning in females (Townes, Trupin, Martin, & Goldstien, 1980) and/or a generalized vulnerability of the male central nervous system to any form of early (including prenatal) trauma or maldevelopment (Mathura, 1979). As a major focus of the current examination was to determine the utility of the LNNB in typical assessment situations facing the practicing school psychologist, this unequal representation of the sexes in the sample was felt to be consistent with the population generally served by this discipline.

The mean age of the total sample was 14.367 years with no significant difference observed between the mean ages for females and males (female mean age = 14.375, male mean age = 14.365, F = 0.00, $p > .05$). Overall intellectual functioning levels as reflected by the WISC-R Full Scale I.Q. were also equivalent for the male and female subsamples (female mean 87.625, male mean 92.573, F = 2.98, $p > .05$), and no significant differences emerged on standardized academic attainment (WRAT) grade scores in Reading, Spelling or Arithmetic between the female and male subsamples, as tabulated below:

Table 14

WRAT Grade Score Attainment by Sex

Variable	Males		Females			
	Mean	S.D.	Mean	S.D.	F Ratio	P
Reading	6.39	1.85	6.60	1.69	0.24	> .05
Spelling	5.45	2.06	6.13	2.17	2.06	> .05
Arithmetic	4.12	1.30	4.17	0.87	0.03	> .05

As the male and female subsamples were observed to be equivalent with respect to age, general intellectual potential, and academic attainment levels in all three basic areas tapped by the WRAT, the male and female subgroups were combined for subsequent data analysis procedures.

Procedure

Subjects for this study were assessed using a battery of tests administered under standard procedures, following the administration prodecure instructions published for the test manuals for the LNNB (Golden, Purisch, & Hammeke, 1985), WISC-R (Wechsler, 1974) and WRAT (Jastak & Jastak, 1978). All neuropsychological testing with the LNNB was completed by the study's author, and in approximately 97 percent of the cases (116 out of 120) the WISC-R and WRAT were also administered by the author as well. In four cases the WISC-R and WRAT test scores were provided from other certified psychologists who had administered these tests within the past six months, and repetition of these intellectual and academic tests was not attempted to avoid contamination by practice effects. Testing order was counterbalanced over the course of the data collection period, with half of the subjects administered the LNNB first, followed by the WISC-R and then the WRAT, with the other half of the subjects receiving the WISC-R and WRAT first, followed by the LNNB.

Subsequent to neuropsychological and psychometric testing, the data were compiled for statistical analysis. Initial analyses focused upon establishing the reliability of the LNNB clinical scales through calculation of estimates of internal consistency (Cronbach's Alpha). An individual item analysis was also conducted by means of examination of the correlation of each item with the scale it was assigned to and comparison with correlations with other scales of the LNNB. Identification of items with inadequate correlation with their assigned clinical scale and of items with zero variance lead to the identification of an abbreviated set of LNNB items that reduced the overlap between the seperate clinical scales and which maintained a high internal consistency and correlation with the original scales, while offering a substantially reduced number of total items and administration time. Validity of the LNNB for psycho-educational evaluation purposes was examined first by means of correlations between the LNNB and accepted instruments of intelligence (WISC-R) and academic achievement (WRAT), and secondly by means of a linear discriminant analysis in which the LNNB clinical scales were utilized to predict group membership that was defined a priori with reference to WRAT academic achievement patterns. The dimensional structure of the LNNB was evaluated by means of a factor analysis of the clinical scale scores for this sample. Evaluation of the clinical utility of the LNNB was conducted by means of utilizing the clinical scale scores in a cluster analysis, to derive subgroups of adolescent students that were evaluated with respect to their educational significance by comparisons of group mean scores on the WISC-R and WRAT variables that were not included in the clustering process. Similar tests of validity and clustering were also conducted on the abbreviated set of LNNB items to determine its consistency with the full LNNB battery.

CHAPTER V

RESULTS

Reliability Evaluation of the LNNB

An initial objective of this investigation was to evaluate the psychometric properties of the LNNB and its performance in the adolescent age range. This aspect was first investigated with respect to reliability of the individual clinical scales, utilizing Cronbach's Alpha as an index of internal consistency. From the total subject pool of 116 cases that had complete data (including all 269 item scores from the LNNB), descriptive statistics were prepared, using the DERS DEST02 programme. From this analysis, it was found that the eleven clinical scales that comprise the full complement of LNNB items displayed a mean internal consistency reliability estimate of 0.787, with a range from 0.707 on the Receptive Speech Scale to a high of 0.873 on Arithmetic (see Table 15). These reliability estimates, using an alpha coefficient approach, are comparable to the results reported by Maruish et al. (1985) on a series of designated samples of brain impaired, schizophrenic, mixed psychiatric and normal controls. The internal consistency estimates obtained in this sample of adolescents fall in a range intermediate between the various clinical groups and the normal control group reported by Maruish et al. (1985).

Table 15

Comparative Coefficient Alpha Reliabilities of Several Subsamples for the LNNB Clinical Summary Scales

Clinical Summary Scales	Adult Samples* A	B	C	D	Adolescent Sample
Motor	.93	.93	.92	.77	.81
Rhythm	.84	.86	.83	.67	.78
Tactile	.89	.89	.85	.54	.73
Visual	.84	.83	.84	.59	.73
Receptive Speech	.90	.89	.84	.69	.71
Expressive Speech	.93	.93	.89	.74	.84
Writing	.84	.86	.85	.59	.78
Reading	.88	.88	.87	.68	.85
Arithmetic	.93	.93	.90	.74	.87
Memory	.84	.86	.82	.72	.75
Intellectual Processes	.91	.92	.89	.78	.81
Average of Coefficients	.88	.89	.86	.68	.79

*Adult Samples: A= Brain Impaired (n=451)
B= Schizophrenic (n=414)
C= Mixed Psychiatric (n=128)
D= Normal (n=108)
Source: Maruish, Sawicki, Franzen, and Golden (1985)
Adolescents: (n=116)

From an examination of the reliability estimates on these basic LNNB scales from the various samples reported by Maruish et al. (1985) and from the current investigation, it is observed that the estimates derived from this adolescent sample fall consistently below the reliabilities for all three adult clinical samples (brain injured, schizophrenic and mixed psychiatric) but above the estimates for the adult normal control group in the Maruish et al. study. As the sample of adolescents in this study was based principally upon out-patient referrals for evaluation on the basis of learning and/or behavioural problems rather than for assessment of the effects of documented brain damage or injuries or for the evaluation of severe psychotic level psychiatric disturbances, the range of reliabilities obtained on this sample would be consistent with the expected level of general disturbance in higher cognitive functioning for this group. As noted by Maruish et al. (1985), the relatively lower range of reliabilities obtained from the normal control group was not an unexpected finding, in view of the relatively restricted variance of test scores in this group, due to the relative lack of difficulty level of most of the LNNB items for normal control subjects. The level of reliability observed for a particular test is dependent upon the nature of the group upon which the reliability is being measured, particularly the range of individual differences within the group (Anastasi, 1982).

In this connexion, it is logically consistent to observe a level of reliability for a mixed group of normal learning and mild to moderately learning delayed adolescents that falls in an intermediate range between normal adults and neurologically or psychiatrically disordered adults. Also consistent with this interpretation would be the observation that the LNNB scales that would theoretically be the most likely to display the greatest variance in a mixed group of learning delayed and nondelayed adolescents (the academic scales of Writing, Reading, and Arithmetic, plus the Intellectual Processes scale) generally display higher reliability coefficients than scales that would not be theoretically related strongly to academic attainment (such as the

LNNB Tactile, Visual, Memory, or Rhythm scales). It is most probable that the obtained reliability estimates, while falling within the lower end of the clinically acceptable range, are somewhat artifactually restricted due to the very limited age range of this sample (a range of only four years) and the relative lack of coverage of a broad range of ability levels. Relatively few adolescents were referred for neuropsychological evaluation who had above average general intellectual and/or academic functioning or who were functioning intellectually in the mentally deficient range. With these restrictions in mind, the observed Alpha coefficients are notable and suggest a relatively robust degree of reliability for this adolescent aged sample.

In respect to the observed mean Cronbach Alpha coefficient of 0.79, it is noted that this estimate of reliability is also considerably lower than that reported for investigations of split-half reliability such as Golden, Fross, and Graber (1981), who reported a mean split-half reliability of 0.92 for an adult population of mixed psychiatric, neurological and normal controls. The Alpha coefficient approach to estimating internal consistency, however, represents a metric of the average of all potential split-half reliability estimates from different splittings of a test, and as such it is similar to the Kuder-Richardson formula 20 in providing a conservative estimate of reliability, especially if the composition of the scales is not highly homogeneous.

This is a factor that is particularly notable on the LNNB, as all of the clinical scales, were constructed with a relatively heterogeneous item/task composition, and with many items assessing rather disparate features of a general ability domain (such as motor functioning) by requiring the subject to complete a variety of tasks, and with scores being based on quite different criteria for different items. As an example, on the Motor scale, the item tasks range from tests of relatively simple motor speed (such as opening and closing the hands or touching of finger to thumb sequences), tests of reproduction of motor movements either from visual presentation or based upon tactile/kinesthetic feedback, fine motor coordination and drawing tasks, and tests of the ability to alternate movements or to inhibit habitual responses in the presence of conflicting verbal commands. Test scoring is based for some items on the quality or accuracy of reproduction of the designated movements or motor products (e.g., reproducing finger or hand positions, scoring drawings for accuracy of reproduction), while on other items the dependent variable is the time or speed of response. With such a relatively heterogeneous collection of tasks, the use of Alpha coefficients would be prone to potential under-estimation of the reliability of the scales, even if the various tasks are themselves individually reliable.

It is concluded on the basis of this evaluation of internal consistency that the LNNB remains a reliable instrument even for an adolescent age population, with the obtained reliability estimates being adequate for clinical research and the slight restriction of the current estimates being attributable to the differences in severity of disorder as well as restrictions in individual differences in both age and general functioning level between the current sample and previous adult reliability studies. The level of observed reliability would be sufficient to accept Hypothesis 1 in predicting equivalent reliability for the adolescent as well as adult age ranges.

Construct Validity Evaluation

In addition to the determination of the internal consistencies of the clinical summary scales, an individual item analysis was also performed, to evaluate the degree to which the individual items on each scale correlated with the scale it was assigned to in comparison to the balance of the LNNB clinical scales. The construct validity of the LNNB has been called into question by authors such as Spiers (1981, 1982) and Delis and Kaplan (1982), who assert that the composition of the LNNB clinical scales is excessively heterogeneous and that failure on items from any scale can be caused by a variety of impairments unrelated to the function that each scale is purported to tap.

Golden, Fross, and Graber (1981) addressed this question by examining the proportion of items that correlated most strongly with its assigned scale as contrasted to correlating more strongly with another clinical scale, and they reported that on a mixed sample of normal controls, psychiatric patients and mixed neurological patients (total N = 338) that 250 out of the 269 items were maximally correlated with their own assigned clinical scale. A similar analysis was conducted with the data from this adolescent sample, and the results for each scale are presented in Table 16, with the equivalent results from Golden, Fross, and Graber (1981) included for comparison. Hypothesis 2 predicted equivalent levels of consistency for the current study as for previous adult investigations.

Table 16

Luria - Nebraska Neuropsychological Battery Item-Scale Consistency

Scale	Total no. of items on scale	Golden et al. (1981) No. of items with highest r on scale	Adolescents No. of items with highest r on scale
Motor	51	46	31
Rhythm	12	12	12
Tactile	22	22	17
Visual	14	12	11
Receptive Speech	33	32	18
Expressive Speech	42	34	20
Writing	13	12	11
Reading	13	13	11
Arithmetic	22	22	17
Memory	13	13	11
Intellectual	34	32	21
Total	269	250=92.9%	180=66.9%

On the basis of this table, it is apparent that the high degree of item to scale correlation that was reported by Golden et al. (1981) does not hold for this group of adolescents, with a reduction from a reported level of 92.9 percent to only a 66.9 percent level. Hypothesis 2 must therefore be rejected. This effect was especially marked for the clinical scales of Receptive and Expressive Speech as well as for the Intellectual Processes and Motor scales. Examination of the correlations that were higher for scales other than that to which the item was assigned failed to reveal any significant pattern, with all eleven major clinical scales having items from other scales correlating excessively with themselves, and the range of such inadvertent item correlations was relatively limited (from 5 to 9 items per scale), such that no one scale was considerably more likely to be the target of such "off assignment" maximal correlations.

Item analysis also revealed the presence of 12 items with zero variance across the adolescent sample, reflecting that the item was so simple that none of the 116 subjects failed to display perfect performance on the task. These nondiscriminating items were primarily grouped among the Motor scale (five items) and the Expressive Speech scale (three items) with the Visual, Receptive Speech, Writing and Arithmetic scales all displaying an additional single item with zero variance.

It is concluded that for the sample of adolescents with educational and/or behavioural problems used in this study, there are significant deficiencies on several of the LNNB clinical scales with respect to construct validity, such that a substantial number of the items on the battery either fail to display any variability or discrimination between lower and higher functioning subjects in the sample or they correlate more strongly with scales other than they are assigned to. This would suggest that the scales, as published, lack sufficient homogeneity to warrant a summation of individual item scores to calculate a meaningful summary clinical score. The finding of a substantial number of items that are more strongly correlated with one or more other scales than with their assigned scale would also suggest that the test may display an excessively high level of intercorrelation between the summary clinical scale scores, so that the validity of labelling the scales with distinct titles and treating them as measures of discrete functional abilities may be inappropriate.

Intercorrelation of LNNB Clinical Scales

Consistent with this interpretation, a variance-covariance analysis of the clinical summary scores yielded a Cronbach Alpha coefficient of 0.93, indicating that the scales were highly consistent in terms of the functions they reflected, and suggested that for this sample they are all likely tapping a single general factor. A table of correlations between the LNNB clinical scales is provided in Table 17, and corroborates the interpretation that there is a high degree of intercorrelation between these clinical scales, with all obtained intercorrelations being significant at beyond the 0.001 probability level.

Table 17

LNNB Clinical Scale Intercorrelation Table

Scale	1	2	3	4	5	6	7	8	9	10	11
1	1.00	.62	.57	.61	.42	.57	.43	.48	.50	.52	.43
2		1.00	.50	.58	.51	.64	.55	.54	.63	.54	.57
3			1.00	.57	.41	.52	.40	.39	.57	.45	.44
4				1.00	.54	.65	.53	.52	.64	.56	.61
5					1.00	.63	.54	.46	.67	.63	.65
6						1.00	.73	.76	.71	.61	.66
7							1.00	.82	.69	.47	.54
8								1.00	.62	.42	.49
9									1.00	.61	.71
10										1.00	.64
11											1.00

Note: Degrees of Freedom = 114; all correlations $P < .001$

Factor Analysis of the LNNB

A principal component factor analysis of this correlation matrix resulted in the identification of only a single factor with an eigenvalue greater than 1.0, with an obtained eigenvalue on this initial factor of 6.678. This single significant factor accounted for 60.7 percent of the total variance on the matrix, and this finding strongly supported the interpretation that the LNNB clinical scales were all highly saturated with this common factor, which appears to be reflective of a general level of adaptive functioning or intellectual ability, a finding that confirms the prediction of Hypothesis 5. Factor loadings of the eleven clinical scales on this single general factor are presented in Table 18.

Table 18

Factor Loadings of the LNNB

LNNB Scale	Factor 1 Loading
Motor	.710
Rhythm	.783
Tactile	.672
Visual	.795
Receptive Speech	.757
Expressive Speech	.879
Writing	.786
Reading	.762
Arithmetic	.864
Memory	.752
Intellectual Processes	.791

As can been seen on Table 18, the LNNB clinical scales are observed to display very high saturation with this single principal component factor, and no meaningful pattern of differential loading is apparent to suggest a specific qualitative interpretation of this factor. It is most likely that this factor is reflective of a generalized ability measure, most likely reflecting the overall level of performance on the test battery as a whole. Such a high saturation with a single factor would suggest that this battery, when used in a clinical setting with adolescent cases, appears to be principally sensitive to the general adaptive function, most probably reflecting a general intellectual ability level.

Evaluation of the obtained eigenvalue structure for the LNNB indicated that in addition to the highly significant general factor there were at least two additional factors that contributed some significant variance to the test, and which appeared to be reflective of group factors. The eigenvalue structure on the LNNB was observed to be very similar to that reported for the WAIS-R (Hill, Reddon, & Jackson, 1985), with a similar very high loading on the first or general factor, followed by two smaller loadings on group factors, and a significant discontinuity after the third factor. Similar to the multiple factor analytic studies conducted on the Wechsler scales, the patterning of these eigenvalues suggests that meaningful interpretations of rotated group factors may be possible, despite the strong loading of all scales on the first principal component.

Based upon this observation of meaningful contribution of the first three factors, a Varimax rotation was applied to maximize the simple structure of the rotated first three factors, and resulted in the identification of three interpretable factors that accounted for 76.7 percent of the total variance. Table 19 displays these rotated factoral loadings for the LNNB.

Table 19

Rotated Factor Loadings for the LNNB

LNNB Scale	H^2	Factor 1	Factor 2	Factor 3
Motor	.781	.186	.241	.830
Rhythm	.647	.379	.403	.584
Tactile	.706	.247	.147	.789
Visual	.683	.446	.316	.620
Receptive Speech	.766	.808	.284	.182
Expressive Speech	.813	.471	.661	.394
Writing	.873	.324	.855	.193
Reading	.905	.186	.896	.262
Arithmetic	.765	.623	.487	.374
Memory	.732	.760	.149	.364
Intellectual Processes	.774	.789	.294	.254
% Common Variance		36.0	32.4	31.6
% Total Variance		27.6	24.9	24.2

Inspection of the factoral loadings on the above table would indicate that Factor 1 is strongly represented by loadings on the LNNB scales of Receptive Speech, Memory, and Intellectual Processes, with a moderate secondary loading on the Arithmetic scale. Consideration of the specific item content of these scales would suggest that this factor is most probably reflective of a general verbal intellectual function, including as components functions such as attention to and discrimination of phonemic sound units, interpretation of the meaning of words and phrases, analysis of grammatical structures, recognition of numbers and performing both written and mental calculations, recall of word lists, sentences and story content, analysis and interpretation of parables, analogies and common idiomatic expressions, definition of words, identification of essential verbal similarities and differences and abstract verbal classification skills. This initial factor appears to cover a relatively broad selection of verbal skills, and in a manner similar to the subtest composition of the WISC-R, it reflects a combination of both fluid intellectual skills (abstract reasoning, memorization of new word lists, etc.) as well as skills of a more practiced of crystallized nature (defining vocabulary words, performing simple mathematical calculations, etc.). This first factor will be designated as a Verbal Intelligence factor.

Factor 2 emerged with pronounced loadings on the Writing and Reading scales, with secondary moderate loadings on the Expressive Speech and Arithmetic scales. This factor appears to be most strongly related to the level of attainment in the language arts area, with a particularly strong reflection of basic writing skills. The loading on the Expressive Speech scale is consistent with this interpretation as this particular LNNB scale, while labelled in a manner that would suggest a primary focus upon oral language functioning, has a large number of items (18 out of the total of 42 items) that require functional reading skills for completion. It is also noted from an analysis of the item to scale correlations that the Expressive Speech items that involved reading were generally among the scales with the strongest correlations to the total Expressive Speech scale score, with an average item to scale correlation of 0.385. The Expressive Speech scale items that did not involve or depend upon reading skills were, on the whole, somewhat weaker predictors of the scale total score, with an average item to scale correlation of 0.287. The Expressive Speech scale from the LNNB is therefore more strongly reflective of general facility with reading than of oral speech functioning, and the significant loading with the second factor would support an interpretation of this factor being principally reflective of language arts academic attainment. This second factor will be labelled as an Academic Attainment factor, and is viewed as reflecting principally crystallized abilities of phonemic decoding and oral pronunciation of letter and letter combination sounds, single word recognition and oral pronunciation of words in isolation, oral reading of short sentences and paragraphs, spelling, writing of short phrases and sentences to verbal cues, and performance of basic arithmetical calculations.

The third factor was characterized by strong loadings on the Motor and Tactile scales, coupled with moderate loadings on the Rhythm and Visual scales. This factor appears to reflect the level of performance on a variety of tasks that are less dependent upon verbal mediation and which may be equated to the visual-spatial and manual dexterity tasks that are included in the Performance subtests of the WISC-R. The loadings would suggest that this factor is maximally reflective of the functions of motor speed, coordination and dexterity, appreciation of spatial relationships and right/left discrimination, perceptual acuity of spatial relationships and configurations (via both tactile and visual modalities), and concentrational factors in the analysis of nonverbal auditory stimuli such as rhythm and pitch relationships. This factor will be labelled as a Perceptual/Motor factor.

Consideration of the observed factor structure on the LNNB would suggest that this battery of neuropsychological tasks has a strong saturation with a general ability factor or an overall adaptive functioning variable, but that interpretable group factors are also present. The loadings of these group factors would be quite consistent with factoral loadings observed for other broad ranged batteries of general functioning, such as the WAIS-R and WISC-R, and the three group factors obtained appear to reflect a strong verbal intellectual factor, a nonverbal intellectual or perceptual/motor factor and a factor that is reflective of general academic attainment, especially in language arts.

From a neuropsychological perspective, these factors would appear to be consistent with the theoretical expectations of a population of learning disabled and behaviourally disordered adolescent population. The strong loadings on the unrotated first factor would be consistent with the observation that in this sample of consecutive clinical referrals there was a substantial range of functioning levels displayed, with some cases displaying considerable levels of clinical impairment on most variables, including neuropsychological, intellectual and academic, while other cases (primarily behaviourally disturbed youths with no histories of specific learning problems) displayed essentially normal neuropsychological functioning and in a few cases significantly above average intellectual abilities. A strong loading on a global level of performance factor would also be consistent with the research on the effects of neuropsychological dysfunction on children, as the most pronounced effects of chronic/static neurological damage or maldevelopments are typically observed on measures of overall adaptive functioning such as intelligence (Reitan & Davison, 1974).

The identification of three interpretable group factors that reflect verbal and nonverbal intellectual skills as well as a more specific academic attainment factor would also be consistent with the previous research in neuropsychological functions, as verbal skills are generally observed to be primarily reflective of the functional integrity of the left cerebral hemisphere, while nonverbal or perceptual/motor tasks are generally interpreted as more reflective of the nondominant right cerebral hemisphere. The academic attainment factor would likely be reflective of a crystallized verbal skill, which while clearly dependent upon the functional integrity of the language dominant left hemisphere (particularly the left parietal-temporal-occipital region), would also be dependent upon the personal learning experience history, and would reflect the end-product of an interaction between basic learning competencies, exposure to appropriate learning environments and emotional/motivational factors that support or conversely discourage the application of effort to mastery of academic skills. The obtained factoral structure of the LNNB would therefore support the view that this battery is tapping a number of essential functions that are relevant to adaptation to the demands of living, and especially to the mastery of academic learning tasks. Both left (verbal) and right (nonverbal) cerebral functions appear to be evaluated by this battery, and both fluid as well as crystallized intellectual skills are assessed. The results of this factor analysis would confirm both Hypothesis 5 in predicting a strong general level of functioning factor and Hypothesis 6 in predicting the presence of group factors that are neuropsychologically meaningful.

Derivation of an Abbreviated LNNB

Based on the observation of a dramatically lower level of item to scale correlation consistency, the occurrence of a number of items that displayed no significant variability among a relatively heterogeneous sample of adolescents, and the observation of an extremely high level of intercorrelation between the LNNB clinical scales it was decided to explore the possibility of determining a subset of the LNNB items that would have maximal consistency with the scale they were assigned to and which would allow for retention of adequate reliability for each clinical scale.

An item analysis of correlations between each item and the eleven clinical scales was therefore conducted and an abbreviated scale scale derived, retaining only those items that correlated both significantly and highest with the scale it was assigned to. An initial attempt at eliminating all items that did not correlate highest with the scale on which they were placed plus all items that displayed zero variance resulted in the retention of a set of 175 items from the original 269 items of the LNNB, but resulted in significant losses in reliability (Alpha coefficient) for several scales. Modification of the inclusion rules to allow for retention of some items that correlated very highly (greater than 0.40) with the assigned scale, even if equally high or minimally higher correlations were observed on other scales resulted in the identification of a subset of the LNNB items (hereafter referred to as the Luria - Nebraska Neuropsychological Battery - Abbreviated or LNNB-A) that still allowed for a significant reduction in total number of items, elimination of all items with zero variance, a high level of consistency of item to scale correlation, and no loss of internal consistency for any of the basic eleven clinical summary scales (see Table 20).

Table 20

Internal Consistency and Correlations Between the LNNB and LNNB-A

Scale	LNNB N	LNNB Alpha	LNNB-A N	LNNB-A Alpha	r LNNB-LNNB-A	t	P
Motor	51	0.81	28	0.81	0.961	0.497	NS
Rhythm	12	0.78	12	0.78	1.000	0.000	NS
Tactile	22	0.73	14	0.74	0.957	0.424	NS
Visual	14	0.73	12	0.74	0.992	1.641	NS
Rec. Speech	33	0.71	18	0.71	0.961	0.052	NS
Exp. Speech	42	0.84	28	0.84	0.981	0.712	NS
Writing	13	0.78	11	0.79	0.997	0.691	NS
Reading	13	0.85	11	0.86	0.999	3.948	.001
Arithmetic	22	0.87	17	0.87	0.996	0.731	NS
Memory	13	0.75	11	0.71	0.975	1.455	NS
Int. Proc.	34	0.81	27	0.82	0.986	1.447	NS
Total N:	269		189				
Average Correlations:		0.79		0.79	0.982		

The use of this item analysis and selection procedure thereby resulted in the identification of an abbreviated number of items that retained (and in one case significantly improved) the reliability of the clinical scale, while resulting in a marked reduction in the overall length of the battery (with a savings of 29.7 percent in length) and which had an extremely high correlation with the full LNNB battery scales (average correlation between the full and abbreviated scales equal to 0.982). This abbreviated battery was considerably more consistent with respect to the correlations between the individual items and the assigned scales, with 175 out of the 189 retained items (92.6 percent) correlating highest with their assigned scales, a percentage that is essentially identical to that reported by Golden, Fross and Graber (1981) for their study with adult subjects. These results confirm the predictions of both Hypotheses 3 and 4.

Intercorrelations of the LNNB-A Clinical Scales

The scale intercorrelations of the LNNB-A were also examined to determine the extent to which this abbreviation of the LNNB affected the previously noted tendency for all scales to intercorrelate highly. The clinical scale intercorrelations for the LNNB-A are presented in Table 21.

Table 21

LNNB-A Clinical Scale Intercorrelations

Scale	1	2	3	4	5	6	7	8	9	10	11
1	1.00	.57	.47	.53	.32	.42	.40	.44	.47	.48	.40
2		1.00	.46	.59	.45	.61	.55	.55	.63	.49	.54
3			1.00	.51	.30	.46	.37	.36	.52	.43	.36
4				1.00	.45	.57	.51	.52	.63	.50	.55
5					1.00	.44	.48	.35	.55	.51	.55
6						1.00	.67	.75	.65	.44	.46
7							1.00	.81	.66	.39	.48
8								1.00	.61	.33	.43
9									1.00	.55	.65
10										1.00	.54
11											1.00

Note: Degrees of Freedom = 114; all correlations $P < .001$

On a comparison of this intercorrelation matrix with that of the original LNNB, it is noted that there is a small but highly significant reduction in the overall level of intercorrelation between the clinical scales, with the mean intercorrelation for the LNNB matrix being .621, compared with a mean intercorrelation of .554 on the LNNB-A matrix ($t = 5.17$, $P < .001$). However, there continues to be a significant degree of overlap existent, with all clinical scales intercorrelating beyond the .001 significance level, suggesting that a strong general factor is still present in this matrix.

Factor Analysis of the LNNB-A

A principal components factor analysis of the LNNB-A intercorrelation matrix revealed the presence of two factors with eigenvalues greater than 1.0 (with eigenvalues of 6.327 and 1.005), which is in contrast to the single factor with an eigenvalue greater than unity obtained on the same analysis of the LNNB correlations. This would suggest that the reduction in item overlap in the LNNB-A battery has resulted in some improvement in the differential factor loading of the individual scales, but with continued highly significant loadings for all scales on the first or general ability factor. As with the analysis of the eigenvalue values from the LNNB, the LNNB-A eigenvalues also displayed a marked discontinuity between the third and fourth value, and a three factor solution was also suggested. The factor loadings for these three principal component factors, after Varimax rotation to maximize simple structure are presented in Table 22.

Table 22

Rotated Factor Loadings for the LNNB-A

LNNB-A Scale	H^2	Factor 1	Factor 2	Factor 3
Motor	.689	.255	.180	.769
Rhythm	.643	.482	.383	.514
Tactile	.685	.164	.183	.790
Visual	.657	.374	.412	.589
Receptive Speech	.771	.234	.842	.079
Expressive Speech	.787	.698	.393	.380
Writing	.846	.864	.264	.171
Reading	.896	.907	.124	.241
Arithmetic	.746	.553	.554	.364
Memory	.696	.095	.712	.424
Intellectual Processes	.745	.353	.728	.300
% Common Variance		37.0	32.8	30.2
% Total Variance		27.4	24.3	22.4

From the data contained in Table 22, it is observed that the three factor solution for the LNNB-A accounts for 74.1 percent of the total variance for this abbreviated neuropsychological battery, which is equivalent to the 76.7 percent accounted for in the three factor solution for the LNNB. The factor structure obtained is also highly consistent with that of the full LNNB, with the first factor displaying primary loadings on the Reading, Writing and Expressive Speech, and Arithmetic scales (equivalent to factor two of the LNNB), the second factor having major loadings on the Receptive Speech, Memory and Intellectual Processes scales (equivalent to factor one of the LNNB) and the third factor displaying principal loadings on the Motor, Rhythm, Tactile and Visual scales, in a manner entirely consistent with that of the third factor from the LNNB factor analysis.

It is concluded that the factoral structure of the LNNB-A is essentially equivalent with that of the full LNNB, with a large eigenvalue (6.327) for the first unrotated factor indicating a high saturation with a general ability or level of performance factor, and three group factors reflecting Verbal Intelligence, Academic Attainment, and Perceptual/Motor abilities. As with the factoral structure for the complete LNNB battery, these results would suggest that the shortened LNNB-A is still adequately tapping both the verbal intellectual and academic achievement functions that are dependent upon the left cerebral hemisphere as well as the nonverbal sensory/perceptual, spatial organization and perceptual/motor skills that are typically ascribed to the right cerebral hemisphere. Despite the substantial shortening of the length of the battery, these factor loadings would suggest that no significant loss in the breadth of functional coverage has occurred, nor any significant alteration in the dimensional nature of the functions reflected by the individual scale scores. These results would confirm the predictions of Hypotheses 7, 8 and 9 regarding the factor structure of the LNNB-A.

Correlation of LNNB with other Measures

To assess the validity of the LNNB as a measure of functions that are relevant to educational attainment and identification of learning disorders, a correlational analysis was conducted between the LNNB clinical scales and the intellectual quotients from the

WISC-R and the grade scores of the WRAT. These correlations are summarized in Table 23.

Table 23

Correlations of LNNB with Intellectual and Academic Test Variables

LNNB Scale	WISC-R			WRAT Grade Score		
	V.I.Q.	P.I.Q.	F.S.I.Q.	Reading	Spelling	Arithmetic
Motor	-0.38	-0.61	-0.60	-0.42	-0.36	-0.37
Rhythm	-0.57	-0.39	-0.57	-0.57	-0.54	-0.45
Tactile	-0.36	-0.45	-0.57	-0.43	-0.39	-0.34
Visual	-0.49	-0.59	-0.65	-0.54	-0.45	-0.46
Rec. Speech	-0.58	-0.35	-0.55	-0.52	-0.46	-0.43
Exp. Speech	-0.67	-0.48	-0.68	-0.78	-0.70	-0.51
Writing	-0.55	-0.31	-0.51	-0.82	-0.85	-0.61
Reading	-0.55	-0.31	-0.53	-0.87	-0.84	-0.50
Arithmetic	-0.62	-0.39	-0.58	-0.65	-0.65	-0.70
Memory	-0.57	-0.49	-0.63	-0.49	-0.39	-0.45
Int. Proc.	-0.73	-0.45	-0.67	-0.57	-0.52	-0.55
Average:	-0.55	-0.44	-0.59	-0.61	-0.56	-0.49

N=120 All correlations P < .001

On this table of results from the LNNB, it is observed that there are highly significant correlations between all eleven clinical scales and the external variables of WISC-R I.Q.'s and WRAT grade scores. With respect to the dimension of verbal intelligence as reflected by the WISC-R V.I.Q. score, it is observed that the LNNB Intellectual Processes scale correlates strongly with this measure, with an inverse correlation of -0.73 (higher scores on LNNB scales reflect greater impairment or lower functioning). Performance I.Q. did not, however, correlate as strongly with the Intellectual Processes scale, and while the obtained value of -0.45 is highly significant from a statistical perspective, it is matched or exceeded by five other LNNB scale correlations. The highest single LNNB correlation with P.I.Q. was obtained for the Motor scale, with a correlation of -0.61, followed closely by a correlation of -0.59 with the Visual scale. This would be consistent with the fact that the WISC-R Performance I.Q. is based upon a series of perceptual-motor tasks that involve the analysis of visual-spatial relationships and/or the motoric manipulation of materials under time constraints. The overall or Full Scale I.Q. was observed to correlate strongly with both measures of nonverbal abilities on the Motor and Visual scales as well as with verbal tasks as tapped by the Expressive Speech and Intellectual Processes scales. The Memory scale, which is composed of a combination of both verbal as well as nonverbal immediate recall items, was also observed to correlate highly with the Full Scale I.Q. score. The correlation between the Full Scale I.Q. and the Intellectual Processes scale was relatively robust at -0.67, but it was marginally exceeded by the correlation of -0.68 between the Full Scale I.Q. and the Expressive Speech scale.

With respect to the prediction of Hypothesis 10 that the LNNB Intellectual Processes scale would display a higher correlation with the WISC-R I.Q. scores than any other LNNB clinical scale, it is observed that this only holds for the Verbal I.Q. score, while for both the Performance I.Q. and Full Scale I.Q. variables there are other LNNB clinical scales that correlate at the same or higher levels than the Intellectual Processes scale. Hypothesis 10 is therefore only partially supported, and the view that the Intellectual Processes scale is a measure of general (as contrasted with specific verbal) intelligence must be rejected.

It is apparent that this neuropsychological test battery has a substantial loading on intellectual functions, and that the various clinical scales of the LNNB display strong associations with the verbal and nonverbal intellectual abilities tapped by the WISC-R. These results are relatively consistent with those of McKay, Golden, Moses, Fishburne and Wisniewski (1981) who reported that in their sample of 280 adult subjects (including neurological, psychiatric, and normal control subjects) the correlations between the LNNB clinical scales and WAIS I.Q. scores were highly significant for all comparisons ($P < .01$). The average correlations between the eleven clinical scales and the WAIS I.Q. scores were considerably higher in the McKay et al. (1981) study, however, with average correlations with V.I.Q., P.I.Q., and F.S.I.Q. of -0.72, -0.67, and -0.73 being reported in that study. This would likely reflect the diverse sample utilized in that evaluation, with subjects included from populations ranging from significantly impaired neurological patients to normal controls of at least average general intelligence, in contrast to a considerably more restricted range of impairment for the current sample of adolescents.

While the LNNB appears to have a high level of saturation with the general intellectual factor (consistent with the results of the previously discussed factor analysis), it is also apparent that the

the Intellectual Processes scale is more clearly associated with or reflective of verbal intellectual functions than of the nonverbal intellectual skills such as are reflected by the WISC-R Performance I.Q. score. It is also apparent from a comparison of the average correlations between the I.Q. scores and the LNNB clinical scales that verbal functions are tapped more extensively than are nonverbal skills by the battery as a whole. This would be generally consistent with the theoretical orientation of A. R. Luria, upon whose work this battery was derived, as Luria had placed a considerable degree of emphasis in his theoretical writings to the primacy of the language dominant cerebral hemisphere in assuming a controlling or directing function, and a large percentage of his investigative efforts and developed techniques were addressed to exploring the intricacies of various neurologically based language disorders. As a result, the basing of the LNNB items on the theoretical position and the published evaluation procedures for the Luria investigation (Christensen, 1975) has resulted in a significant biasing toward more extensive coverage of verbal functions, with a somewhat less exhaustive evaluation of nonverbal intellectual abilities.

On the correlations between the LNNB and the WRAT grade scores, it is noted from Table 22 that these correlations are also all highly significant ($P < .001$ in all comparisons), and that the highest correlations are consistently observed between the WRAT grade scores and the related LNNB academic scale. Thus, the highest correlation for the WRAT Reading grade score is with the LNNB Reading scale, the highest correlation with the WRAT Spelling grade score is with the LNNB Writing scale, and the highest correlation with the WRAT Arithmetic grade score is with the LNNB Arithmetic scale. These results confirm the predictions of Hypotheses 11, 12 and 13, and provide support for the contention that the LNNB academic scales are valid measures of the scholastic skill areas they purport to assess. In all cases the correlations between the LNNB academic scales and the corresponding WRAT variables were significantly higher than the correlations between the WISC-R intellectual estimates and the WRAT scores, suggesting that the observed LNNB-WRAT correlations are attributable to more than just a general intellectual ability factor. This would be consistent with the observation of a third Academic Attainment factor in the LNNB factor analysis, and supports the interpretation of the LNNB as displaying a combination of both fluid and crystallized intellectual abilities.

Correlation of the LNNB-A with other Measures

The correlations between the shortened format LNNB-A and the WISC-R and WRAT were also investigated, and the obtained results are presented in Table 24.

Table 24

Correlations of LNNB-A with Intellectual and Academic Test Variables

LNNB-A Scale	WISC-R			WRAT Grade Score		
	V.I.Q.	P.I.Q.	F.S.I.Q.	Reading	Spelling	Arithmetic
Motor	-0.34	-0.56	-0.54	-0.38	-0.32	-0.36
Rhythm	-0.58	-0.40	-0.57	-0.59	-0.55	-0.47
Tactile	-0.32	-0.39	-0.42	-0.42	-0.37	-0.32
Visual	-0.49	-0.56	-0.62	-0.57	-0.46	-0.47
Rec. Speech	-0.53	-0.29	-0.48	-0.44	-0.3[illegible]	-0.36
Exp. Speech	-0.63	-0.47	-0.65	-0.78	-0.69	-0.50
Writing	-0.54	-0.32	-0.51	-0.83	-0.84	-0.61
Reading	-0.53	-0.33	-0.53	-0.87	-0.84	-0.52
Arithmetic	-0.60	-0.39	-0.57	-0.64	-0.63	-0.71
Memory	-0.48	-0.45	-0.56	-0.41	-0.32	-0.41
Int. Proc.	-0.73	-0.47	-0.68	-0.55	-0.48	-0.57
Average:	-0.52	-0.42	-0.56	-0.59	-0.53	-0.48

N=120 All correlations $P < .001$

On this table it is observed that the correlations between the LNNB-A scale scores and the intellectual variables from the WISC-R remain highly significant. As with the correlations with the full LNNB, the language functions reflected by the WISC-R Verbal I.Q. score were more strongly correlated with the LNNB-A than the nonverbal intellectual abilities tapped by the Performance I.Q. variable. It is also noted that the LNNB-A Intellectual Processes scale remains predominantly a measure of verbal intellectual ability, with a significantly stronger correlation with the WISC-R Verbal I.Q. than with the Performance I.Q. score. Nonverbal intellectual abilities are best reflected on the LNNB-A by the Motor and Visual scales, which is consistent with the results from the complete LNNB, and the general ability measure represented by the

WISC-R Full Scale I.Q. was observed to be most strongly associated with the Intellectual Processes, Expressive Speech, and Visual scales, reflecting both verbal and nonverbal tasks. The average correlations between the LNNB-A and the WISC-R I.Q. variables are only minimally lower (.02 to .03) than those displayed for the LNNB, suggesting that the reduction in the length of the battery has had a minimal effect upon the coverage of the intellectual functioning variables, with both verbal and nonverbal functions still significantly represented.

With respect to the prediction of Hypothesis 14 that the LNNB-A Intellectual Processes scale would correlate higher with the WISC-R I.Q. scores than any other LNNB-A clinical scale it is observed that this prediction is confirmed for both the Verbal as well as the Full Scale I.Q. variables, but is rejected for the Performance I.Q. score variable. The correlation between the Intellectual Processes scale and the WISC-R Full Scale I.Q. is also observed to be only marginally higher than the correlation between the Expressive Speech scale and this I.Q. variable. As with the full LNNB, it is concluded that the LNNB-A Intellectual Processes scale remains predominantly a measure of verbal intelligence, with considerably weaker coverage of the nonverbal intellectual abilities tapped by the WISC-R Performance I.Q. score.

Academic functioning was also observed to be significantly correlated with the LNNB-A, and as with the LNNB it is noted that the academic scales of the LNNB-A are the most robust correlates with the WRAT variables. Both the LNNB-A Reading and Arithmetic scales displayed the strongest correlations with the equivalent WRAT score, but the LNNB-A Writing scale failed to display a significantly higher correlation with the WRAT Spelling grade score, with the LNNB-A Reading score displaying an equal correlation to WRAT Spelling. This is essentially consistent with the findings on the LNNB, however, as even the full battery displayed only a minimally higher correlation for the Writing scale over the Reading scale (.85 compared to .84). The difficulty in establishing a clearly higher correlation with the Reading or Spelling WRAT scales is, however, also largely a reflection of the strong intercorrelation between these two WRAT variables, as in the sample of 116 subjects used for the LNNB investigation, the WRAT Reading and Spelling grade scores were observed to correlate at the 0.87 level, which is even higher than the intercorrelation of 0.81 between the LNNB-A Reading and Writing scales. The extremely high correlation between the two WRAT language arts scale variables would indicate that they are highly overlapping in functional content, such that any variable that correlates strongly with one must also correlate to an essentially similar level with the other.

There is, however, a much clearer differentiation possible between the academic attainment of the two LNNB-A language arts scales of Writing and Reading and the Arithmetic scale, such that the WRAT Arithmetic grade score is correlated with the LNNB-A Arithmetic scale at a significantly stronger level than with any other LNNB-A scale, including the language arts academic scales, and the two WRAT language arts scales (Reading and Spelling) correlate at a significantly stronger level with the corresponding two LNNB-A language arts scales than with any other clinical scale. These findings would confirm the predictions of Hypotheses 16 and 17, and give partial support to the prediction of Hypothesis 15, and would indicate that even with the substantial reduction in the length of the battery, the LNNB-A academic scales remain robust and valid reflections of the basic academic skill areas they purport to measure.

Multiple Regression Analyses of the LNNB and LNNB-A

An alternative approach to investigating the utility of the LNNB as a measure of educationally relevant functions is to explore the ability of the test battery as a whole to predict external criteria of intellectual abilities and academic attainment, using a multiple regression analysis. This allowed for the identification of which elements in the LNNB provided information that was related to the external criterion, and the estimate of a multiple correlation between the LNNB and each of the WISC-R and WRAT variables.

Prediction of WISC-R I.Q. Scores from the LNNB

Prediction of intellectual functioning was investigated and using a stepwise multiple regression analysis of the LNNB clinical scale T-Scores the following regression formula was derived to predict Verbal I.Q.:

Table 25

Multiple Regression Analysis of LNNB to Predict V.I.Q.

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Intellectual Processes	-0.664	0.115	33.261	0.000
Expressive Speech	-0.378	0.138	7.471	0.000
Rhythm	-0.383	0.230	2.759	0.046

Constant term in equation = 114.912

Multiple correlation R = 0.766; R^2 = 0.587

From this table, it is observed that the multiple correlation coefficient obtained between the three LNNB scales and the WISC-R V.I.Q. falls at a relatively robust level, and accounts for about 59 percent of the variance in this sample's V.I.Q. distribution. The combination of predictors included in this regression equation confirms that the LNNB Intellectual Processes scale is a strong predictor of general verbal intelligence, as well as the Expressive Speech scale. The substantial third contribution from the Rhythm scale was unexpected, however, as this scale is, on the surface, apparently nonverbal in its orientation, and no clear analogue for this type of task is present among the WISC-R Verbal subtests. The LNNB Rhythm scale is composed of a series of items that require sustained concentration, attention to auditory stimuli (both rhythmic and pitch perception) and on several items the application of counting of groups of beats. While the stimuli presented are nonverbal in nature, the task demands incorporated on this scale allow for and encourage the frequent application of verbally guided strategies, such as counting the number of beats, verbally labelling tones as being higher or lower in pitch, and assigning verbal labels to individual tones of lower and higher sound intensity to aid in recall of the patterning of the groups. In all cases, the presented stimuli of tones or beeps are presented with very limited ranges of variation, which also aids the utilization of simple verbal labelling. Thus, while the stimuli are of a nonverbal nature, the task demands are such that the application of an analytic strategy based upon verbal labelling is frequently used, and this scale is observed to load extensively on the verbal intellectual dimension.

Verbal intelligence, as reflected by the LNNB, would appear to be a combination of oral and written language skills (Expressive Speech), facility in higher level abstract reasoning using language (Intellectual Processes) and a facility with applying verbal analysis strategies even to problem solving tasks involving nonverbal stimuli (Rhythm). The obtained multiple correlation would indicate that the LNNB could be used as a valid measure of verbal intellectual functioning, at least within the ability range covered by this clinical sample.

The stepwise regression analysis for predicting Performance I.Q. was conducted in a similar manner and yielded the following equation:

Table 26

Multiple Regression Analysis of LNNB to Predict P.I.Q.

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Motor	-0.304	0.150	75.823	0.000

Constant term in equation = 117.345

Multiple correlation R = 0.632; R^2 = 0.399

On this table, it is observed that only a single predictor variable was included that met the minimal inclusion criterion at the 0.05 probability level. This single variable, the Motor scale, resulted in a multiple correlation coefficient estimate of 0.632, and accounted for approximately 40 percent of the sample variance in Performance I.Q. scores. As noted previously in the discussion of the intercorrelations between the LNNB and WISC-R, the LNNB appeared to display a generally higher level of correlation with verbal intellectual functions that with nonverbal abilities, and this has been substantiated by this multiple regression analysis. Outside of the one major loading on the Motor scale, the balance of the LNNB clinical scales do not present sufficient unique variance to significantly improve upon the prediction using the Motor scale alone. This would suggest that the LNNB Motor scale, which is composed of a variety of tasks that assess motor functioning through both gross motor movements as well as through displays of fine motor coordination and speed, would be viewed as tapping into the same general ability domain as the WISC-R Performance I.Q. score. As speed of response and reaction time are specifically included as components of the Motor scale, this would support interpretations of the WISC-R Performance I.Q. as being at least partially dependent upon such simple motor functioning factors (c.f., Sattler, 1974).

The WISC-R Full Scale I.Q. was next investigated, and the multiple regression equation for the LNNB clinical scales is presented in Table 27.

Table 27

Multiple Regression Analysis of LNNB to Predict F.S.I.Q.

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Intellectual Processes	-0.697	0.095	53.666	0.000
Motor	-0.621	0.115	29.403	0.000

Constant term in equation = 120.602

Multiple correlation R = 0.751; R^2 = 0.564

This regression equation table indicates that the WISC-R Full Scale I.Q. can be best predicted from the LNNB data through a weighted combination of the Intellectual Process scale (which has been observed to tap primarily verbal intellectual skills) and the Motor scale (as a measure of nonverbal intellectual skills). The obtained multiple correlation coefficient is relatively robust, and accounts for approximately 56 percent of the variance in WISC-R general intellectual functioning in this sample. The relative small number of LNNB variables included in this regression equation would suggest that there is a very strong degree of overlap between the different LNNB clinical scales with respect to intellectual functioning, such that a very small selection of the clinical scales is sufficient to account for most of the variance in the sample distribution, and the balance of the LNNB providing little further significant information regarding intellectual functioning.

Prediction of WRAT Grade Scores from the LNNB

In a similar manner, the LNNB clinical scale T-Scores were used in a stepwise multiple regression to predict the three academic attainment grade scores from the WRAT. The regression equation data for the prediction of the WRAT Reading grade score is summarized in Table 28.

Table 28

Multiple Regression Analysis of LNNB to Predict WRAT Reading

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Reading	-0.068	0.011	39.036	0.000
Writing	-0.045	0.012	14.652	0.000
Expressive Speech	-0.033	0.010	10.980	0.000

Constant term in equation = 14.969

Multiple correlation R = 0.901; R^2 = 0.[illegible]

From this table, it can be seen that the LNNB scale variables of Reading, Writing and Expressive Speech can be combined to yield a highly significant multiple correlation coefficient that accounts for approximately 81 percent of the variance in the WRAT Reading grade score for this sample. It is also noted that the LNNB Reading scale is the first variable to enter the stepwise regression equation, and this is consistent with the predicted results, and supports the validity of this scale as a measure of basic reading skills. As noted previously in the discussion of the intercorrelations between the LNNB Reading a Writing scales, these two variables are highly correlated, but the incorporation of both in the stepwise regression equation would indicate that the Writing scale also contains some unique variance that adds to the prediction of the WRAT Reading score, above and beyond that contained in the LNNB Reading scale. As both of these measures were observed to display substantial correlations with both verbal as well as general intellectual functioning, it is likely that this contribution from the Writing scale may be related to the saturation with the verbal intellectual factor. It is also noted that there is a very significant contribution from the Expressive Speech scale. As was discussed previously, this LNNB has a strong loading with the verbal intellectual factor, and it also contains a rather high number of items that demand functional reading skills for completion, such that it is as much a measure of reading ability as a measure of oral language or verbal intellectual skills.

The regression equation data for the prediction of the WRAT Spelling grade score is presented in Table 29.

Table 29

Multiple Regression Analysis of LNNB to Predict WRAT Spelling

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Writing	-0.158	0.009	298.519	0.000

Constant term in equation = 15.535

Multiple correlation R = 0.847; R^2 = 0.717

This regression analysis indicated that the best LNNB predictor of the WRAT Spelling score was the Writing scale, and that this scale variable alone was sufficient to account for about 72 percent of the variance in the sample distribution of WRAT Spelling scores. Spelling ability, as measured by the WRAT Spelling test, appears to be a more specific skill than reading, and has fewer variables contributing to its variance.

The multiple regression equation data for predicting the WRAT Arithmetic grade score are presented in Table 30.

Table 30

Multiple Regression Analysis of LNNB to Predict WRAT Arithmetic

LNNB Variable	Beta Coef.	S.E.	F of Beta	P
Arithmetic	-0.035	0.006	36.242	0.000
Writing	-0.026	0.010	7.327	0.001

Constant term in equation = 8.267

Multiple correlation R = 0.720; R^2 = 0.519

The results indicate that a weighted combination of the LNNB Arithmetic plus the Writing scales was the most successful in predicting the WRAT Arithmetic grade score, with a multiple correlation coefficient that accounted for approximately 52 percent of the sample variance on WRAT Arithmetic. These scale weightings are consistent with expectations and suggest that both specific arithmetical skills as well as language arts skills are required to complete the written calculations of the WRAT Arithmetic test.

Consideration of the results from these regression analyses would indicate that the LNNB is a relatively robust instrument with respect to evaluating academic achievement, with very high correlations with reading and spelling and a moderately high multiple correlation with arithmetic attainment. In each case it is observed that the LNNB academic scale that corresponds to the WRAT attainment area is the first (or in the case of spelling the only) variable to enter the regression equation, a finding that supports the validity of the LNNB academic scales as measures of the skills they purport to assess. It is also noted that the multiple correlations obtained are generally higher than those typically reported from studies using intellectual test scores to predict academic achievement, which are generally observed to result in correlations around the 0.50 level (Matarazzo, 1972). This would appear to be related to the inclusion of measures of both fluid intellectual abilities as well as more crystallized measures of actual academic skills in the LNNB.

With respect to the prediction of Hypothesis 19 that highly significant predictions of WISC-R and WRAT variables would be made on the basis of the LNNB clinical scale data, it is observed that all three of multiple correlations with the WRAT grade score variables were highly significant and exceeded the arbitrarily chosen cut-off of 0.70. With respect to the WISC-R intellectual variables, the multiple correlations with both the Verbal and Full Scale I.Q. were in excess of the 0.70 level, while the multiple correlation with Performance I.Q. fell substantially lower than this criterion level. It is therefore concluded that the LNNB represents a highly significant measure of both academic attainment and verbal and general intellectual functions, but that it is relatively deficient as a measure of nonverbal intellectual abilities.

Comparison of the LNNB and LNNB-A Multiple Regression Analyses.

A similar series of multiple regression analyses to predict WISC-R I.Q. scores and WRAT grade scores was also conducted, using the raw score totals for each of the LNNB-A clinical scales. These regression analyses yielded equations that were highly consistent with those from the full LNNB. Verbal I.Q. was found to be best predicted from the LNNB-A by a weighted combination of the Intellectual Processes and Expressive Speech scales, Performance I.Q. was also found to be predicted by the single variable of the Motor scale, and the WISC-R Full Scale I.Q. was predicted from a combination of Intellectual Processes, Expressive Speech, and Motor scales. The WRAT grade scores were observed to also show a relatively consistent pattern of predictor loadings, with the WRAT Reading grade score being predicted by the weighted combination of LNNB-A Reading, Writing and Expressive Speech, WRAT Spelling was predicted by the combination of LNNB-A Writing and Reading scales, and the WRAT Arithmetic grade score was best predicted by a combination of LNNB-A Arithmetic, Writing and Intellectual Processes scales. The obtained multiple correlation coefficients for the LNNB-A were essentially equivalent to those from the LNNB regression analyses, and are presented in Table 31 for comparison purposes.

Table 31

Multiple Correlation Coefficients for LNNB and LNNB-A

Predictor Variable	LNNB		LNNB-A	
	R	R^2	R	R^2
WISC-R:				
Verbal I.Q.	0.766	0.587	0.765	0.586
Performance I.Q.	0.632	0.399	0.568	0.323
Full Scale I.Q.	0.751	0.564	0.763	0.582
WRAT:				
Reading Grade Score	0.901	0.811	0.901	0.812
Spelling Grade Score	0.847	0.717	0.895	0.801
Arithmetic Grade Score	0.720	0.519	0.737	0.543

As can be seen from the data in Table 31, the abbreviated LNNB-A battery has maintained a robust level of validity with respect to being able to predict both intellectual as well as educational attainment variables. The multiple correlations for the WT [illegible] equivalent between the two batteries with respect to predicting both the Verbal and Full Scale I.Q.'s, but a small reduction in the correlation with Performance I.Q. is observed. With respect to predicting educational attainment, the LNNB-A is observed to yield an identical level of predictive validity with respect to predicting the WRAT Reading grade score, and slightly higher multiple correlations are observed between the LNNB-A and the WRAT Spelling and Arithmetic grade scores than between the LNNB and the same W[illegible] test variables.

It appears that the substantial reduction in test length for the LNNB-A has had insignificant effects on the validity of this abbreviated battery as a measure of general intelligence and academic attainment. The only area in which the abbreviated battery displays a significantly lower multiple correlation coefficient is in the prediction of the Performance I.Q. score, and this is observed to be the weakest area of predictive validity for both the LNNB and LNNB-A batteries. It is concluded that neither the LNNB or the LNNB-A are fully adequate as measures of the nonverbal intellectual skills tapped by the WISC-R Performance subtests, and that the abbreviation process has only minimally contributed to a further reduction in this area of relative weakness. The reduction in the length of the battery as achieved on the LNNB-A could, however, allow for the incorporation of additional measures of nonverbal intellect while maintaining the brevity of the overall assessment process, thereby improving upon the coverage of skills without adding to the total time required for an assessment.

Identification of Subgroups of Students Defined A Priori

An additional approach to the evaluation of the utility of the LNNB for educational assessment purposes was to explore the ability of this neuropsychological battery to identify individuals with varying patterns of academic performance and the presence of different patterns of academic achievement deficit. This was done through an application of a linear discriminant analysis to attempt to identify subjects within the sample that had different patterns of academic achievement on the WRAT. From the total sample of 120 subjects, a total of 114 were identified that could be classified a priori into one of three subgroups consisting of: 1) A specific learning disability in arithmetic, 2) A global learning disability involving both language arts and arithmetic, or 3) Normal learner, with no significant deficits in either language arts or arithmetic. The groups were defined on the basis of the WRAT standard scores, with membership in Group I (specific arithmetic disability) defined as displaying a standard score of less than 80 on the WRAT Arithmetic test, with both Reading and Spelling falling at or above a standard score of 80. Group 2 (global learning disability) membership was assigned to any individual who displayed a combination of significant delays in academic attainment in both language arts and arithmetic, with WRAT standard scores below 80 in both Arithmetic and either or both of the Reading and Spelling tests. Cases were assigned to Group 3 (normal learners) if none of the WRAT standard scores fell below the 80 point level. Four cases from the original pool of 120 were eliminated due to missing data that prevented the calculation of equivalent scores for both the LNNB and LNNB-A, and two additional cases were eliminated as not classifiable into any of these three defined groups.

Following assignment of all cases to these a priori group classifications, the scores from the eleven LNNB clinical scales were utilized in a linear discriminant analysis to attempt assignment to three subgroups. The results of this discriminant analysis for the LNNB clinical scale data is presented in Table 32.

Table 32

Classification of Adolescent Cases by LNNB Data

Group	Linear Discriminant Classification: Group 1	Group 2	Group 3	
1	27	7	8	N = 42
2	6	43	0	N = 49
3	4	2	17	N = 23
	N = 37	N = 52	N = 25	Total = 114

On this 3 x 3 classification matrix it is readily observed that the majority of the classifications derived from the linear discriminant analysis of the LNNB clinical scale data were consistent with the actual group assignments as specified a priori by the WRAT academic attainment pattern. Applying a test for discriminatory power as described by Press (1982), the relationship between the two classification schemes is observed to be highly significant, with an obtained chi square of 94.78 (df = 2, P < .001). An overall classification accuracy (all three groups combined) of 76.3 percent was obtained. It can be concluded that the linear discriminant analysis was highly successful in classifying individuals into the three subgroups at much better than a chance level, and this confirms the prediction of Hypothesis 20.

It is noted, however, that the level of classification accuracy was rather inconsistent across the three groups, with members of Group 1 (specific arithmetic disability) being identified correctly by the discriminant analysis 27 times out of 42, for a 64.3 percent hit rate, members of Group 2 (global learning disability) bein correctly identified for 43 out of 49 cases (an 87.8 percent hi rate) and Group 3 (normal learners) receiving correc classifications in 17 out of 23 cases, for a 73.9 percent hit rate. The patterning of erroneous classifications was also observed to b quite different for the three a priori groups, with members of th arithmetic disability grouping having an essentially equa probability of being misidentified as either globally impaired o normal learners, all misidentifications for the global learnin disability group were misassignments to the arithmetic disabilit group, and from the normal learners group there were twice as man misclassifications as arithmetic disabled than as globally impaired.

The patterning of these classification errors would b consistent with general expectation, as the three a priori group could be viewed as falling along a continuum of learning deficits with the normal group functioning the highest, the global disabilit group the lowest and the selective arithmetic disability group a falling at some intermediate point between these two extremes. Within such a view, the fact that the intermediately placed grou has the lowest hit rate of the three a priori groups could b anticipated, as well as the relatively equal probability o misassignments of individuals from this intermediate group to eithe of the surrounding groups. There is, in contrast, relatively littl overlap between the globally impaired and the normal learners group and only two subjects out of the total of 72 cases (less than percent) in these two extreme groups were misassigned to the other's classification.

These results would indicate that the data from the LNNB can be utilized with a relatively high level of confidence in identifying the presence of a learning disorder, as externally defined by the presence of significant delays in attainment on the WRAT. Considering both a priori learning disability groups together and contrasting them to the normal learner group, the overall hit rate in detecting the presence of a learning disability (unspecified as to type) with the LNNB would be 91.2 percent, with only 8 out of the total of 91 learning impaired cases being classified as normal. It is also observed that none of the globally impaired learners (the most severe category of learning problems) would be misclassified as being normal in learning abilities, but that a substantial number of cases (26 percent) with normal academic attainment would be classified as having learning deficits, with most such misdiagnoses falling into the less severely impaired arithmetic disability subgroup. It would therefore appear that the use of the LNNB as a screening instrument for the identification of learning disabilities would result in relatively few false negative (Type I) errors, but a substantial number of false positive (Type II) errors in the adolescent population.

A similar discriminant analysis was also conducted using the raw score totals from the shortened LNNB-A, against the same sample of cases that had been assigned a priori to the learning ability subgroups. The results for this linear discriminant analysis using the LNNB-A to predict group classification is presented in Table 33.

Table 33

Classification of Adolescent Cases by LNNB-A Data

Group	Linear Discriminant Classification Group 1	Group 2	Group 3	
1	28	5	9	N = 42
2	5	43	1	N = 49
3	5	2	16	N = 23
	N = 38	N = 50	N = 26	Total = 114

As with the classification matrix from the linear discriminant analysis of the full LNNB clinical scores, this LNNB-A classification matrix also displays a highly significant relationship between the a priori group assignments and the classifications generated by the discriminant function, with an equal number of total correct assignments displayed across the diagonal (chi square = 94.78, df = 2, $P < .001$). Overall classification accuracy using this abbreviated neuropsychological battery remains unchanged (76.3 percent for all three groups combined), with a 66.7 percent hit rate for the arithmetic disability subgroup, an 87.8 percent hit rate for the global disability subgroup and a 69.6 percent hit rate for the normal learners group. The prediction of Hypothesis 21 is therefore confirmed by this finding.

As with the classifications based on the full LNNB, the majority of the misassignments were based upon confusions of cases from the intermediate severity arithmetic disability subgroup with either the global or the normal groups, and there was very little confusion or misclassification errors between the global and normal groups. Consideration of the ability of the LNNB-A to identify the presence of a learning disorder and to discriminate between cases with a learning disability (unspecified as to type or severity) and normal learners indicated that this broader classification could be completed with an overall hit rate of 89 percent. Also consistent with the LNNB classification results, there remained a tendency to display relatively few false negative errors while having a substantial number of false positive errors in this adolescent sample.

The results of both linear discriminant analyses would suggest that the LNNB (as well as the abbreviated battery derived in this study) is capable of identifying the presence of a learning disability or delay with a relatively high degree of accuracy, with few cases with significant learning disorders missed or classified as normal learners. A substantial number of cases of essentially normal academic attainment will, however, be classified as learning disordered on the basis of either the LNNB or LNNB-A data. It is felt that in an educational context the potential ramifications of a misclassification as having a potential learning disorder (and thereby probably initiating some further explorations or remedial interventions) would be a less serious error for the individual case than an error of failing to detect the presence of a significant learning disability where one is in existence, and thereby reducing the individual's probability of receiving needed remediation and assistance. While the relative significance or importance of Type I and Type II classification errors must be decided within the context of the particular application or setting in which the instrument is being utilized, the apparent tendency of this instrument to minimize the occurrence of a Type I error (false negative) at the expense of some increased occurrence of Type II errors (false positives) would be considered to be appropriate for a general educational framework.

Cluster Analysis of the LNNB

To explore the ability of the LNNB to identify more homogeneous and naturally occurring subgroups within the sample of adolescents tested, the clinical summary scale scores were utilized as a basis for clustering, using the SPSSx Cluster subprogramme. On the basis of the data from the eleven basic clinical scales, the total sample of 116 subjects was subjected to a step-wise clustering process, using an hierarchical agglomerative algorithm that defined clusters on the basis of maximizing the between groups distance. This procedure produced successively larger cluster groups by joining together individual cases that had the highest similarity between their profiles, and which were maximally dissimilar to other individuals or already formed clusters, with the process continuing until eventually all individuals were combined into a single cluster group. For the purposes of this study, a distance measure of similarity (squared Euclidean distance) was selected as the primary measure due to the expectation that profile elevation as well as shape would have a significant impact on external predictor variables of intelligence and academic attainment.

Cluster analysis is a general term that groups together a large number of alternative grouping procedures, and the resultant clusters formed by the various algorithms can be quite different and show little correspondence to each other (Bartko, Strauss, & Carpenter, 1971). For this reason, a recommended practice has been to attempt clustering using various methods or algorithms, and to select for analysis cluster groupings that are replicated across different clustering methods, as these would then represent the most stable or acceptable cluster solutions (Joschko & Rourke, 1985; Wishart, 1975).

To evaluate the appropriateness of various clustering methodologies for this data sample, an initial procedure was to utilize the SPSSx Cluster programme to cluster analyze the LNNB clinical scale data set using five different methods: 1) Average linkage between groups, 2) Average linkage within groups, 3) Ward's minimum variance method, 4) Centroid clustering, and 5) Median clustering. Of these methods, the average linkage between groups is the SPSSx default method, as it has been shown to be applicable to a very wide range of applications. To explore the stability of the derived cluster solutions, both the size as well as individual case assignments for each of the derived cluster groups was examined. Table 34 presents the data for cluster group size for each of these methods, and the degree of concordance (percentage of overlap) of group membership between the average linkage between groups method and the other applied methods is also presented. For reasons outlined below, these comparisons were based upon a three cluster solution, as being the most appropriate number of clusters in this data set.

Table 34

Concordance Across Different Clustering Procedures for LNNB Data

Cluster Analysis Procedure	Cluster Groups					
	Group 1		Group 2		Group 3	
	N	(%)	N	(%)	N	(%)
Ave. Link. Between	62	(100)	50	(100)	4	(100)
Ave. Link. Within	63	(97)	49	(96)	4	(100)
Ward's Method	72	(98)	32	(62)	4	(100)
Centroid	111	(100)	1	(2)	4	(100)
Median	111	(100)	1	(2)	4	(100)

On the above table it is apparent that the correspondence between the cluster solutions for the various methods applied varied considerably, but with considerable overlap between the solutions. The third cluster group, which was composed of only four members in each solution was observed to remain completely invariant across all four methods, and the members of the first cluster group continued to be grouped together with a very high level of consistency across all solutions (with the addition of further members in several alternative methods). Cluster group 2 was observed to display an almost identical membership across the two average linkage methods, but substantially reduced concordance across the other methods. Inspection of the cluster solutions for both the Centroid and Median methods indicated that while identical cluster assignments had been derived by both of these closely related methods, the cluster solution was composed of essentially one extremely large group with very few assignments to other groups, a pattern that is interpreted as reflective of a chaining process, and indicative of an unsuccessful clustering for these procedures. Based upon the observation of an extremely high degree of concordance between the average linkage between groups and the average linkage within groups methods, it was decided to accept the use of the default method of average linkage between groups for the subsequent analysis of derived cluster groups.

An initial focus of the interpretation of the selected cluster analysis was to determine the most appropriate number of clusters that needed to be formed to allow for the most interpretable solution to the data set. This determination was primarily guided by the examination of the clustering coefficient metric (which reflects the overall level of distance differences within the clusters) that was calculated for each stage of the clustering process, using the guideline that the optimal cluster solution can generally be observed to lie at a point just before the clustering coefficient displays a sudden increase that is out of proportion to previous increments (Everitt, 1974; Wishart, 1975). Such a sudden increase in the value of the clustering coefficient would indicate that two quite dissimilar groupings have been forced together by the agglomeration process, resulting in a new group with significantly reduced internal homogeneity (Morris, Blashfield & Satz, 1981). While the computer output provides a reported clustering coefficient for all 116 steps in the clustering process, only the cluster coefficients for the last ten clustering steps with be presented here (see Table 35) for the sake of brevity.

Table 35

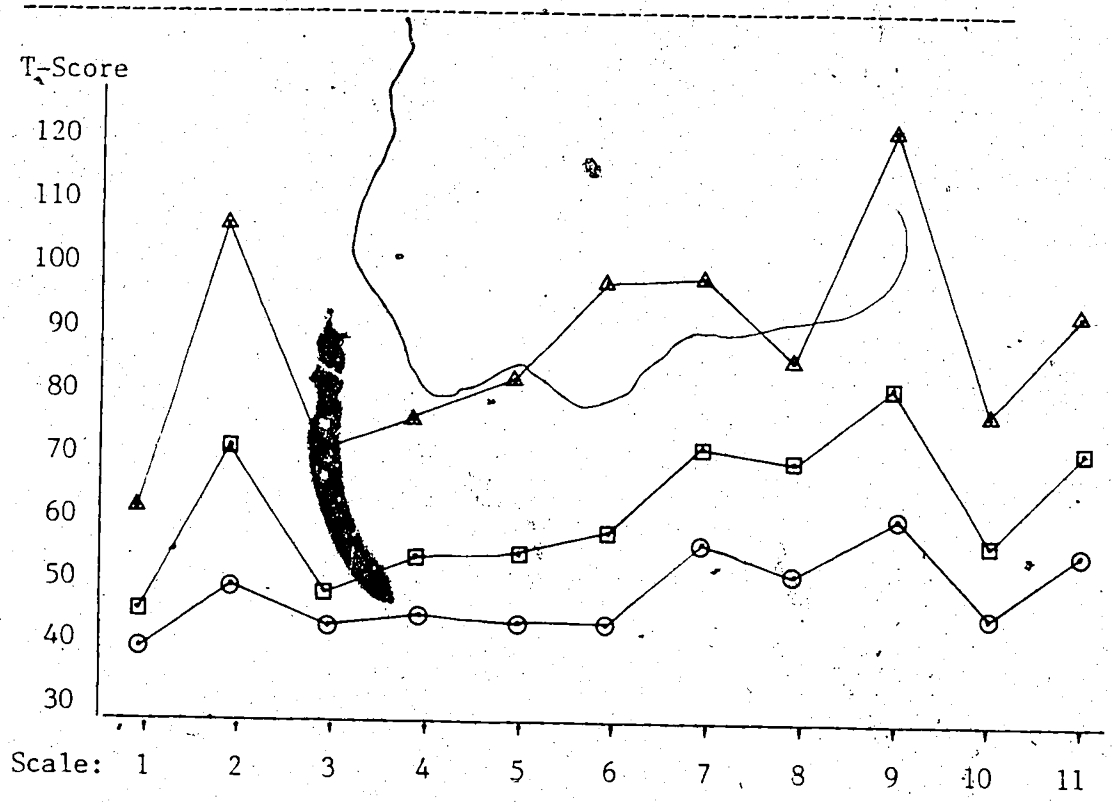
Clustering Coefficients for Adolescent LNNB Data

Number of Clusters	Clustering Coefficient	Percent Change
10	377.552	0.5
9	381.964	1.2
8	382.666	0.2
7	430.227	12.4
6	437.800	1.8
5	498.511	13.9
4	545.489	9.4
3	618.578	13.4
2	1019.243	64.8
1	2574.689	152.6

From this table it is seen that in the step between the three cluster solution and the two cluster solution there is a dramatic and sudden increase in the percentage of change in the cluster coefficient, followed by a further extreme increase as the last two groups are forced together. This pattern of discontinuity in the cluster coefficient values would strongly point to the presence of three clusters within the data set as being the optimal solution. This conclusion was also corroborated by plotting the mean LNNB clinical scale scores for the derived cluster groups for the five, four and three cluster solutions. For both the five and four cluster solutions there remained substantial overlap and crossing over of the plotted mean values for the subgroups, while on the three cluster solution (see Figure 1) a clear discrimination between the three groups was obtained, with no overlap or crossing over of the mean values on the LNNB variables for the three groups.

Figure 1

LNNB Clinical Scale Mean Score Values for Three Cluster Solution



Scale 1 = Motor; Scale 2 = Rhythm; Scale 3 = Tactile;
Scale 4 = Visual; Scale 5 = Receptive Speech;
Scale 6 = Expressive Speech; Scale 7 = Writing;
Scale 8 = Reading; Scale 9 = Arithmetic;
Scale 10 = Memory; Scale 11 = Intellectual Processes

Cluster Group 1 (N = 62) = ○
Cluster Group 2 (N = 50) = □
Cluster Group 3 (N = 4) = △

From the data displayed in Figure 1 it is readily apparent that the three cluster groups observed in this mixed sample of adolescents with educational, emotional and behavioural problems are all quite similar to each other with respect to the average LNNB profile shape or configuration, but with significant differences observed on the mean profile elevation between the three groups. A

MANOVA of the LNNB clinical scale average T-Score values for the three cluster groups indicated a highly significant overall difference between the three group profiles, with an obtained Wilks Lambda value of 0.164 (F = 13.79, P < .001). Univariate F tests for each clinical scale individually also displayed highly significant differences for all clinical scales between the three groups, as summarized in Table 36.

Table 36

Univariate F Tests For LNNB Clinical Scale Variables

Variable	Hyp. SS	Error SS	Hyp. MS	Error MS	F	P
Motor	2942.6	5871.6	1471.3	52.0	28.3	0.000
Rhythm	17930.5	17440.5	8965.3	154.3	58.1	0.000
Tactile	2820.0	6920.3	1410.0	61.2	23.0	0.000
Visual	4215.7	6323.5	2107.9	56.0	37.7	0.000
Rec. Sp.	7546.0	9264.4	3773.1	82.0	46.0	0.000
Exp. Sp.	10979.0	5428.7	5489.5	48.0	114.2	0.000
Writing	8588.5	6162.7	4294.2	54.5	78.7	0.000
Reading	11094.8	7572.0	5547.4	67.0	82.8	0.000
Arithmetic	22476.1	17639.6	11238.0	156.1	72.0	0.000
Memory	5706.8	8944.7	2853.4	79.2	36.1	0.000
Int. Proc.	6147.4	9109.4	3073.7	80.6	38.1	0.000

D.F. = (2,113)

Group 1 is characterized by a generally low profile elevation (mean T-Score elevation = 48.81) with relatively limited variability or scatter on the average profile, with a maximum range of only 21 points between the lowest clinical scale T-Score of 40 (on the Motor scale) and the highest T-Score of 61 on the Arithmetic scale. Standard objective rules for the interpretation of the LNNB would indicate that this group profile is clearly within normal limits even with respect to the adult level normative standards the T-Scores are based upon, and no indication of any specific neuropsychological abnormality emerges for this group as a whole. Calculation of the LNNB Impairment Index, as described by Johnson, Moses, and Bryant (1984) for these mean T-Score values would result in an objective classification of this group profile as falling within the Intact or normal range.

Group 2 from this cluster solution is typified by a profile of similar shape or configuration, but with a significantly higher average profile elevation (mean T-Score elevation = 61.77) and a scatter that is somewhat greater than that of Group 1, with a range of 34 points between the lowest average scale elevation (on the Motor scale) and the highest (on Arithmetic). Utilizing published guidelines for LNNB profile interpretation as presented in the battery's manual (Golden, Purisch & Hammeke, 1985), it is noted that this group profile displays a total of three elevations that exceed the T-Score level of 70, that there is a larger than normal degree of profile scatter or range between the lowest and highest clinical scale elevation, and that using the overall sample age and years of educational exposure to calculate a critical level (the highest T-Score that would be considered to be normal for the age and educational level) this critical level would be set at 62, and there would be a total of 5 significant elevations beyond this level. All of these observations would consistently rate this profile as being abnormal and indicative of significant neuropsychological dysfunction. A further recommended interpretation procedure is to consider the number of clinical scale elevations that fall at or above a level that is 10 T-Score points below the calculated critical level, with observation of more than two thirds of the scales falling in this range being suggestive of brain impairment. Using this procedure, it is noted that 9 out of the 11 scales fall within this range, and this would also be considered as a reflection of an overall group abnormality. The average profile for Group 2, when evaluated with respect to the Johnson et a[illegible]airment Index, would be rated as falling within the Mildl[illegible]ange. The results for this second cluster group wo[illegible]efore be strongly suggestive of the presence of some form of significant neuropsychological dysfunction of a mild level of impairment.

The averaged profile for Group 3 was characterized by an even more extreme profile elevation (mean T-Score = 84.5) and an even more profound level of scatter, with a range of 59 T-Score points between the lowest and highest clinical scale elevations. Consistent with the results on the other two group profiles, the lowest and highest elevations were again observed on the Motor and Arithmetic scales. This group profile would be rated as highly abnormal using standard interpretive guidelines, with all clinical scales falling at or above the calculated critical level of 62 and nine falling in excess of the 70 T-Score level. The level of impairment displayed on this group profile would be rated as as falling in the Marked range according to the objective standards used in calculating the Impairment Index.

The obtained mean profiles for these three cluster groups would therefore reflect the presence of three subgroups within this adolescent population, which are separated primarily on the basis of level of performance or impairment on the LNNB clinical scales. These groups can be classified according to established rules for LNNB profile analysis as being characterized by normal neuropsychological functioning (Group 1), mild neuropsychological dysfunction (Group 2) and marked neuropsychological impairment (Group 3). This would confirm the prediction of Hypothesis 22 with respect to the identification of distinct subgroups in the sample that can be observed to have significant differences in their LNNB profiles. The overall profile shape is, however, remarkably similar across the three groups, with all three subgroups displaying a similar pattern of relative elevation (i.e., poorer performance) on the Arithmetic scale, along with relative elevations on the Rhythm, Writing, and Intellectual Processes scales, contrasted with lower elevations (less impairment) on the Motor, Tactile, Visual and Memory scales. The correlations between these three scales are displayed in Table 37.

Table 37

Intercorrelations of the Mean LNNB Profiles for Three Clusters

Group Profiles	Cluster 1	Cluster 2	Cluster 3
Cluster 1	1.00	0.94	0.81
Cluster 2		1.00	0.91
Cluster 3			1.00

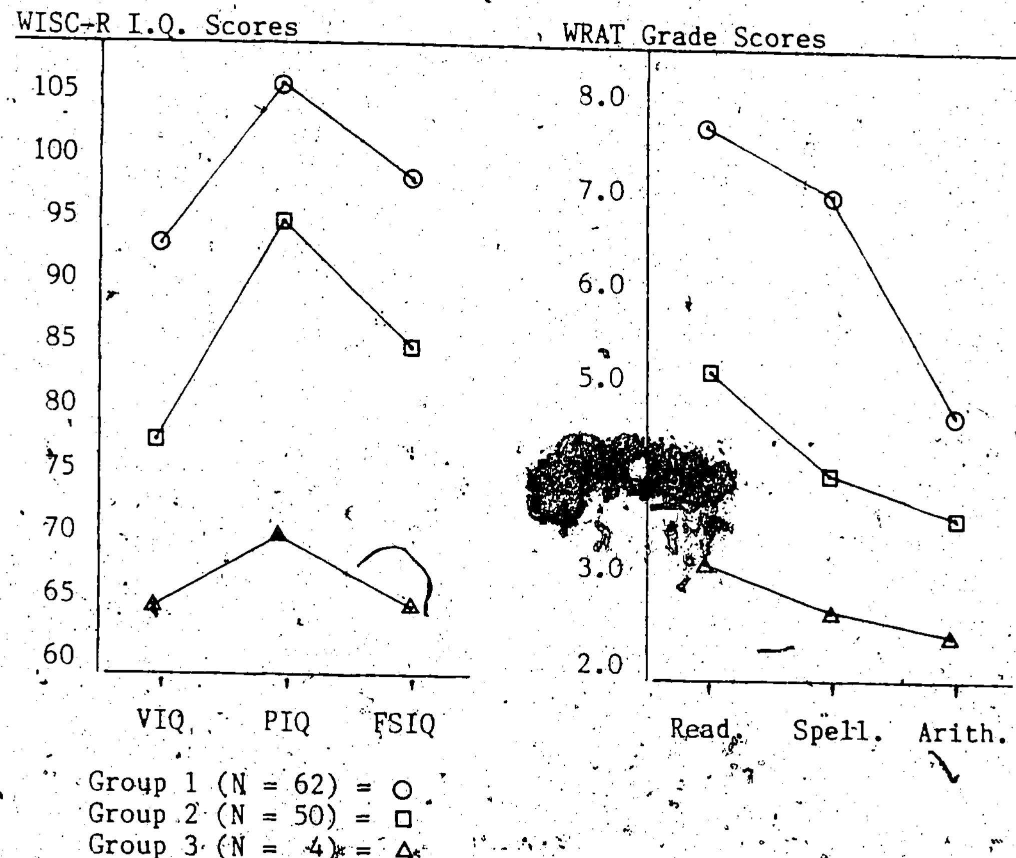
As can be seen in Table 37, all three group profiles are strongly intercorrelated, with all correlations being significant at beyond the .01 level. The correlation strength is also observed to covary with the distance seperation between the groups, with the correlation between the two most extreme groups falling substantially lower than between the intermediately located Group 2 and either of the other groups. This would reflect the observed tendency for the profile scatter or range to increase along with the profile elevation, such that Group 3 displays a much higher level of variability or range in the individual scale elevations than did Group 2, which was also more variable than group 1. Inspection of the profiles for these three groups indicated that this increased variability was primarily related to the dramatic increase on the Arithmetic and Rhythm scales, which displayed differences of 52 and 60 T-Score points respectively between Groups 1 and 3, as contrasted to the relatively lesser differences on measures of basic motor and sensory functioning, with differences of 22 points on the Motor scale, 26 points on the Tactile scale and 28 points on the Visual scale for these same groups. Other LNNB measures of academic attainment and intellectual skills displayed differences of intermediate size between these extremes, with differences of approximately 31 points observed for both the Writing and Reading scales, as well as on the Memory and Intellectual Processes scales.

It is apparent that with respect to the T-Score elevations for the clinical scales of the LNNB, these three identified cluster groups differ primarily with respect to the overall profile elevation, but secondarily with respect to a pattern of increased profile variability or scatter, with the most significant deviations occurring on measures of academic achievement and higher cognitive functions, while measures of more basic motor speed and coordination as well as basic sensory/perceptual functioning displayed somewhat less extreme differences between the three groups. THis pattern of differences between the three objectively defined clusters or subgroups would be consistent with the previously discussed factorial structure of the LNNB, in that the high loading of all the clinical scales on the first factor (which was identified as a general ability or level of performance factor) would be expected to work towards separating individuals or groups based upon their overall or general level of functioning across all the scales. The effect of the identified group factors (Verbal Intelligence, Academic Attainment and Perceptual/Motor factors) would be expected to be considerably less powerful, and to result in less extreme differences in the shape or variability of the profile, in addition to the marked differences in overall profile elevation.

To explore the external validity of the derived cluster groupings (which were based solely upon the LNNB clinical scale elevations) the averaged scores for both the WISC-R I.Q. variables as well as the grade score variables from the WRAT were examined for each of the identified cluster groups. These external variables were found to display clear seperation between the three groups, and their mean value scores are displayed graphically on Figure 2.

Figure 2

Mean I.Q. and WRAT Grade Scores for Three Cluster Groups



On the WISC-R, the three subgroups derived from the LNNB clinical scales were observed to display a clear seperation with respect to averaged intellectual levels, with no overlap or crossing over of the plotted scores occurring for the three group I.Q. profiles. A MANOVA conducted on these I.Q. variables indicated that there was a highly significant overall effect for the three cluster groups (Wilks Lambda = 0.561, F = 12.39, P < .001). There were also clearly significant differences observed for all three I.Q. score variables across the three groups, with univariate F tests indicating high degrees of significance for all three WISC-R scores (see Table 38).

Table 38

Univariate F Tests for WISC-R I.Q. Scores

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	P
V.I.Q.	7368.3	11539.3	3684.1	102.1	36.1	0.000
P.I.Q.	5513.8	23611.9	2756.9	209.0	13.2	0.000
F.S.I.Q.	6936.9	11843.4	3468.4	104.8	33.1	0.000

Consistent with the results displayed on the plotted LNNB clinical scale profiles, these three groups were found to display a clear differentiation with respect to level of functioning in the WISC-R, with Group 1 (previously identified as displaying normal or intact neuropsychological functioning) scoring consistently within the mid-average range on all three WISC-R I.Q. scores, Group 2 (described by the LNNB as displaying mild neuropsychological dysfunction) scoring in the borderline deficient range of functioning on the Verbal I.Q. score, in the mid-average range on the Performance I.Q. score and in the low average range on the Full Scale I.Q. score. The difference between the Verbal and Performance I.Q. scores for this group is large enough to be considered as clinically significant, and would suggest a selective deficit on verbal intellectual functions with normal nonverbal ability level, a pattern generally viewed as typical of the majority of learning disabled students. The Group 3 WISC-R I.Q. profile was observed to display significantly below average scores on all three I.Q. variables, with both the Verbal and Full Scale scores falling within the mentally deficient range, while the Performance I.Q. score fell near the bottom of the borderline deficient range. The level of the Verbal/Performance for this group would be insufficient to be considered to be clinic[illegible]ificant.

These cluster gro[illegible]-R scores confirm the prediction of Hypothesis 23 and also str[illegible]ly support the interpretation that the LNNB clustering was based primarily upon the general ab[illegible]ity fact[illegible] The consistency of the seperation between t[illegible] three groups on external I.Q. scores would present a strong level of support [illegible] validation for the meaningfulness of the resultant cluster groupings.

The group mean scores on the WRAT grade scores were also examined as a second test of external validation for the LNNB cluster groups. As seen previously on Figure 2, these academic variables also displayed a [illegible] seperation between the three cluster groups, with Group 1 demonstrating the highest overall academic attainment, Group 2 scoring somewhat lower, and the Group 3 profile being the most severely depressed of the three. A MANOVA evaluation of these three academic scores across the three subgroups indicated a highly significant difference or overall effect was present (Wilks Lambda = 0.388, F = 22.39, P = 0.000). Univariate F tests also indicated that there were highly significant differences for each of the three academic achievement test variables across the three groups (see Table 39).

Table 39

Univariate F Tests for WRAT Reading, Spelling and Arithmetic

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	P
Reading	216.4	154.2	108.2	1.4	79.3	0.000
Spelling	262.4	244.8	131.2	2.2	60.7	0.000
Arithmetic	49.0	125.2	24.5	1.1	22.1	0.000

It is apparent from both the univariate F tests as well as from a visual inspection of the group profiles in Figure 2 that these cluster groupings are highly different with respect to overall level of academic attainment as reflected by the WRAT, consistent with the prediction of Hypothesis 24. It is also noted, as well, that the three group profiles also display a relatively consistent pattern of maximal delay in academic attainment on the Arithmetic test, the best performance on the Reading test, and the Spelling test score falling intermediate between these extremes. As with the previously discussed I.Q. profiles, there is a strong degree of intercorrelation present between the WRAT grade score profiles for these three groups (see Table 40).

Table 40

Intercorrelations of WRAT Profiles for Three Cluster Groups

Group Profiles	Cluster 1	Cluster 2	Cluster 3
Cluster 1	1.00	0.88	0.78
Cluster 2		1.00	0.98
Cluster 3			1.00

The strength of these intercorrelations, coupled with the visual display in Figure 2, would indicate that the principal difference between these three cluster group WRAT profiles is with respect to the profile elevation, with little or no significant difference in the patterning of academic attainment. This would be entirely consistent with the results observed on the WISC-R I.Q. profiles for these three groups, and supports the interpretation that the three cluster groupings are differentiated primarily on the basis of their overall or general ability factor, with little significant differences between the groups in their relative strengths or weaknesses, at least on the educational variables tapped by the WRAT. The clear differentiation between the level of academic attainment for these three groups would also offer a further support for the validity and meaningfulness of the cluster solution, and adds to the validation of the LNNB as a measure of general adaptive or intellectual functioning.

A final comparison was made between the subject assignments to the three cluster groups identified objectively by the cluster analysis and the a priori assignments to the three learning ability/disability subgroups that were based upon the patterning of WRAT standard scores and which had been used in the previously discussed linear discriminant analysis. As the cluster analytic groupings were clearly defined with respect to the overall level of performance for the subjects, the three learning ability subgroups were renumbered to make the numerical assignments more consistent.

For the purpose of this comparison, the a priori subgroup that displayed no significant impairments on either the language arts or the arithmetic scales of the WRAT (the normal learner subgroup) was hypothesized to be equivalent to the neuropsychologically normal Group 1 of the cluster analysis and it was relabelled as a priori Group 1. The a priori group with selective deficits in arithmetic but no significant impairments in either reading or spelling skills was felt to represent an intermediate level of deficiency, and was relabelled to be equated with the Group 2 from the cluster analysis. The global learning disordered a priori subgroup was observed to be the most consistently and severely impaired of the three learning ability subgroups, and was therefore recoded as Group 3 for this comparison. A classification matrix for this comparison of the a priori and the cluster analytic assignments is presented in Table 41.

?

Table 41

A Priori versus Cluster Analytic Classification Matrix

A Priori	Cluster Group: Group 1	Group 2	Group 3	
Group 1	20	3	0	N = 23
Group 2	33	9	0	N = 42
Group 3	7	38	4	N = 49
	N = 60	N = 50	N = 4	N = 114

From the above classification matrix, a chi square of 0.99 was obtained, which was insignificant and suggested that the classifications were not related at beyond a chance level. Examination of the pattern of classifications summarized on this table indicated that there was essentially no clear seperation between the a priori Groups 1 and 2 (the normal learners and specific arithmetic disability groups), with both of these groups displaying large number of assignments to the first cluster group (i.e., the no neuropsychological dysfunction group), and a smaller portion of members from each to the second cluster group (mild neuropsychological dysfunction). Group 3 from the a priori assignments (global learning disorder) were observed to be assigned predominantly to the second cluster, along with a small percentage to the first cluster as well as to the third (marked neuropsychological impairment) cluster group. It is apparent that while the cluster analytic groupings are not consistent with the groupings of learning disorder type based upon the WRAT profile pattern, there is clearly some relationship between the cluster grouping and the presence of a learning disorder. Collapsing the two a priori learning disability subgroups into a heterogeneous group of unspecified learning disordered cases and combining the two neuropsychologically dysfunctional cluster groups to produce a 2 X 2 table of normal versus impaired functioning yielded a considerably more interpretable classification matrix, as presented in Table 42.

Table 42

2 X 2 Classification Matrix

A Priori Class	Cluster Analytic Classification: Normal	Impaired	
Normal	20	3	N = 23
Impaired	40	51	N = 91
	N = 60	N = 54	N = 114

On this condensed classification matrix, in which both the a priori as well as the cluster analytic groups are collapsed to yield a dichotomous rating of normal versus impaired, there was an obtained chi square of 6.88, which was significant at beyond the 0.01 level. Thus, while there is a lack of clear association of the cluster memberships with the specific learning subtypes as defined by the pattern of learning delay on the WRAT, there was a significant relationship between the cluster groupings and the presence of a learning disorder, regardless of type. This would appear to be a reflection of the sensitivity of the LNNB to the overall level of performance (the first, unrotated principal component factor) and the relative lack of sensitivity to the patterning of relative strengths and weaknesses that could define a specific subtype of learning disability.

In addition to the clustering based upon a distance measure of profile similarity (which emphasizes the effect of profile elevation) an attempt was also made to cluster the same data using the cosine option in the SPSSx package, to explore the ability of the data to be organized on the basis of correlation as a similarity metric. This was attempted as the use of correlation would negate the impact of overall profile elevation, and would define similarity on the basis of the profile patterns. As with the use of a squared Euclidean distance metric, a clustering coefficient is also printed for each step of this clustering process, but in this case it represents the averaged intercorrelation of the cluster formed at each step, with reduction in the size of this internal correlation reflecting the forcing together of dissimilar cases or cluster groups. The cluster coefficient for the last ten steps of this cosine based cluster solution are presented in Table 43.

Table 43

Cosine Clustering Coefficients for Adolescent LNNB Data

Number of Clusters	Clustering Coefficient	Percent Change
10	.982716	0.10
9	.982660	0.01
8	.981287	0.14
7	.981079	0.12
6	.979422	0.17
5	.979329	0.01
4	.977180	0.22
3	.975353	0.19
2	.973105	0.23
1	.967428	0.58

From an inspection of this table it is apparent that there is a much less pronounced change in the size of the clustering coefficient across the span presented than with the same table for the squared Euclidean distance coefficient table. It is also apparent that the only increase that stands in contrast to the previous changes in coefficient value was on the last step when the final two groups were merged into a single cluster. This might be taken to represent a two cluster solution, but inspection of the cluster memberships indicated that these two clusters were actually composed of one large cluster of 115 cases and a second cluster of only a single case. It is apparent that this solution does not represent two true clusters, but rather a single cluster with one outlier that does not correlate with any of the other cases and which is only brought into a cluster group at the last stage when all cases are forced into a single cluster representing the entire sample population.

Inspection of the group memberships under the five, four and three cluster solutions also reflected the presence of a single massive cluster with a few relatively insignificant groupings of slightly different profile patterns, with each step resulting in one of the smaller groups being merged into the main large cluster. Thus, at the five cluster solution stage the first cluster was composed of 92 cases (81 percent of the sample), Group two had 10 cases, Group 3 had 6, Group 4 had 7, and Group 5 was composed of the single outlier. The four cluster solution displayed Group 1 swelling to 99 cases, while Groups 2, 3 and 4 had 10. 6, and 1 cases respectively. The three cluster solution resulted in Group 1 increasing to 105 cases (92 percent of the sample), with ten cases remaining in Group 2 and the single case in Group 3.

This pattern of clustering represents a failure to obtain a meaningful cluster solution using this correlational metric of profile similarity. It is an example of a "chaining" phenomenon in which the vast majority of members are grouped into a single cluster with remaining cluster groups containing only a few members, and it is a problem that is commonly encountered when the variable entered into the clustering process are highly intercorrelated (Morris, Blashfield, & [illegible] 1981). This result would indicate that, when the effect [illegible] vation is removed by use of a correlational similari[illegible] high level of intercorrelation between the LNNB [illegible] primary or level of performance factor) acts [illegible] nto a single grouping, and that there is insuff[illegible] ounted for by the group factors to result in mean[illegible] reliable subgroups with distinct profile patterns.

This failure to obtain a valid solution to the clustering using a correlational approach was consistent with a recently published study conducted by Goldstein, Shelly, McCue and Kane (1987), who used cluster analysis procedures on LNNB clinical scale data from 117 neuropsychiatric and learning disabled adults, and who found that the use of a correlational similarity metric resulted in difficulties in obtaining a stable clustering solution, particularly with average linkage within groups methodology. The use of a squared Euclidean distance metric was adopted for their study, and average linkage within groups was applied as the clustering method. The result of this clustering process was to identify four distinct subgroups that varied with respect to both level of performance (as reflected in average profile elevation) as well as profile pattern or shape (the patterning or relative strengths and weaknesses within each subgroup).

These authors reported that while in their study they were able to identify four clusters arising from their sample pool (which was composed of 23 cases of predominantly left hemisphere brain damage, 25 cases of predominantly right hemisphere brain damage, 26 cases of diffuse brain damage, and 40 learning disabled adults without brain damage), these clusters did not demonstrate any clear relationship between the cluster membership and actual a priori diagnosis. They concluded: "if one agrees that concordance between cluster membership and actual diagnostic group is an appropriate index of the external validity of the clustering solution, then it would appear that the present solution does not have satisfactory external validity" (p. 224). Further discussion by these authors also noted that although they had specifically selected patient groups that would maximize the likelihood of distinct patterns of neuropsychological impairment: "A striking result of the study was the failure of the LNNB to lateralize, that is, to form clusters corresponding with the presence of left hemisphere, right hemisphere, or diffuse brain damage" (p. 228). They also indicated that the patients in their sample with right hemispheric brain damage frequently displayed normal LNNB profiles, while the left hemispheric brain damaged patients more often produced abnormal LNNB profiles, but with profiles that frequently resembled those of diffuse brain injured patients. They concluded that: "the LNNB is differentially sensitive to left and right hemisphere brain damage with the result that left hemisphere cases appear to show more deficit on it than do right hemisphere cases" (p. 229). This biasing of the LNNB to tap left hemispheric functions was related to the selection of items for the LNNB from the original Christensen (1975) investigation items, and Christensen was quoted from her monograph as stating that: "It is to be remembered that the whole investigation is mainly considered to evaluate the functions of the left (dominant) hemisphere" (Christensen, 1975, p. 11).

The current study's results would appear to be in partial agreement with that of Goldstein et al. (1987) in that application of a clustering methodology using squared Euclidean distance resulted in a clear solution that appeared to be internally consistent and reliable, and that attempts to use a correlational index of similarity resulted in difficulties with obtaining a valid cluster solution. In both studies there were clear differences in the level of performance between the identified cluster groups, and there was also a relative lack of correspondence between the a priori diagnostic classifications and the cluster memberships, such that the subgroups defined by the clustering procedures were not externally validated with reference to the external criteria of diagnostic classification. A further similarity between the results of the two studies is that both suggest that the LNNB is primarily sensitive to the verbal or language deficits that are most likely to arise from left cerebral hemisphere damage or dysfunction, and that the sensitivity to or coverage of the nonverbal cognitive functions mediated by the right hemisphere is much less thoroughly assessed. The failure of the LNNB to be able to discriminate between the left and right hemisphere patient groups in the Goldstein et al. (1987) study would be quite consistent with the previously discussed factor analytic results that suggested that this test battery has predominant loadings on verbal intellectual and language functions and much more limited coverage of the nonverbal intellectual functions.

A difference that is observed between the current results and those of the Goldstein et al. (1987) study is that while they were able to identify four distinct subgroup LNNB profiles that differed significantly in the shape or pattern of relative strengths and weaknesses, all three of the current study's cluster groups displayed essentially similar profile patterns, with the only significant difference in profile shape being that of an increasing degree of scatter or exaggeration of the differences between the strengths and weaknesses for the neuropsychologically dysfunctional groups. This is most probably related to subject sample differences, with the Goldstein et al. study selecting a purposefully heterogeneous sample of brain damaged and learning disabled adults, while the current study was restricted to the use of adolescent cases of generally milder levels of severity. The report of a similar methodology applied to a radically different sample group resulting in distinct cluster groups that had differences in profile pattern as well as general elevation would suggest that the failure to obtain such differences in profile configuration on the current study may be primarily related to a restricted range of clinical impairments in the adolescent sample.

Cluster Analysis with the LNNB-A

The ability of the abbreviated battery to subgroup the adolescent cases in this sample into meaningful clusters was also explored. To maintain a direct comparability with the results of the cluster analysis of the full LNNB, it was decided to apply the same methodology (using an average linkage between groups algorithm) as well as the same similarity metric (squared Euclidean distance) to the LNNB-A raw score data.

As with the analysis of the LNNB clustering results, an initial evaluation was to determine the most appropriate number of clusters that should be selected to most accurately describe this data set. As noted previously, the examination of the clustering coefficient allows for an objective approach to this decision, and the tabulated values of the clustering coefficient for the last ten stages of the LNNB-A clustering process are presented in Table 44.

Table 44

Clustering Coefficients for Adolescent LNNB-A Data

Number of Clusters	Clustering Coefficient	Percent Change
10	377.552	0.5
9	381.964	1.2
8	382.666	0.2
7	430.227	12.4
6	437.800	1.8
5	498.511	13.9
4	545.489	9.4
3	618.578	13.4
2	1019.243	64.8
1	2574.690	152.6

As can be seen in the above table, it is apparent that there is a dramatic discontinuity in the increase in the clustering coefficient between the three and two cluster solutions, followed by an even more extreme expansion in the coefficient at the last stage when all cases are forced into a single cluster. This pattern is quite consistent with that of the analyses for the LNNB, and would indicate that a three cluster solution remains the best alternative for the data set even while using the abbreviated battery scores.

The cluster sizes and memberships were next examined and contrasted with those obtained from the same clustering procedures using the full LNNB T-Score data. The cluster groups obtained using the LNNB-A displayed a considerable degree of overlap with those derived from the LNNB, with the notable exception of the second group (the middle placed group with an intermediate level of impairment on the LNNB data), which was considerably reduced in size, with a rather large number of cases from this middle group being reassigned to the first cluster. Table 45 presents the data regarding cluster size for these two analyses, as well as the degree of overlap between the obtained cluster groups for the two test forms, expressed as a percentage of cases from the LNNB analysis that were located in the same cluster in the LNNB-A analysis.

Table 45

Comparison of Cluster Groups From LNNB and LNNB-A Data

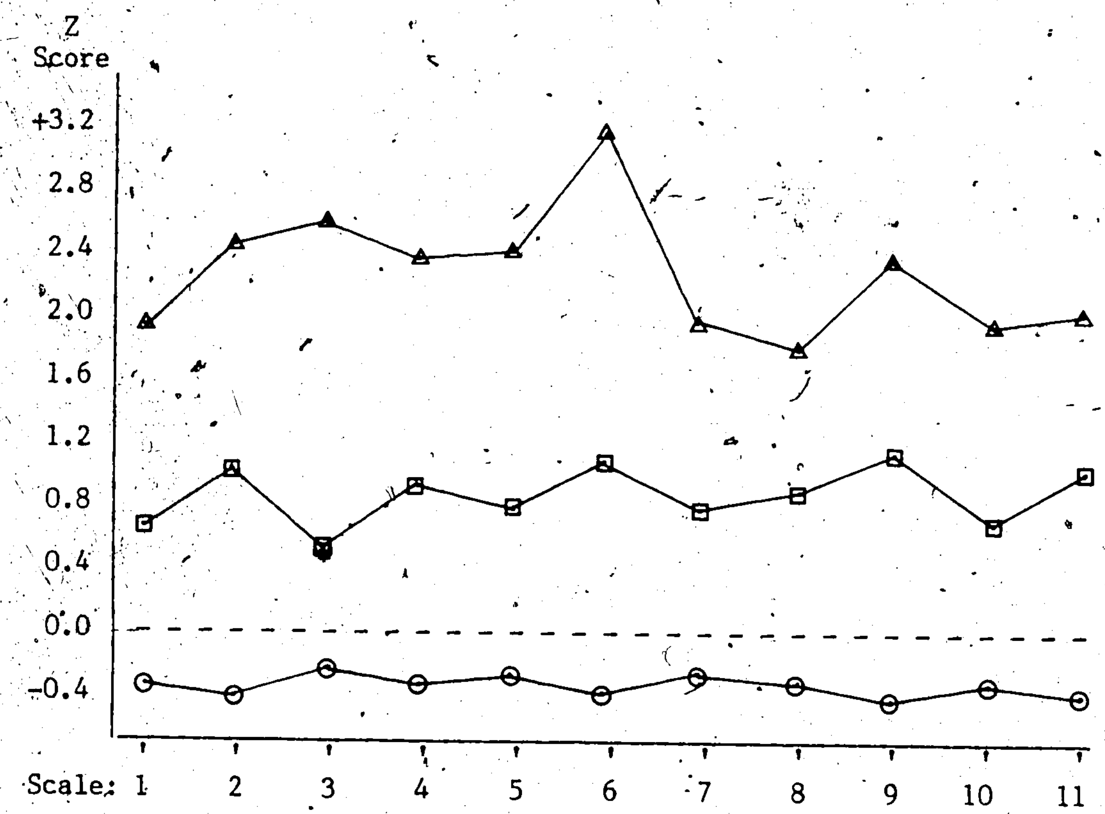
Test Form	Cluster Groups					
	Group 1		Group 2		Group 3	
	N	(%)	N	(%)	N	(%)
LNNB	62	(100)	50	(100)	4	(100)
LNNB-A	91	(100)	21	(42)	4	(100)

From this table, it is observed that for the LNNB-A both cluster Groups 1 and 3 contained all of the members from the same clusters derived from the full LNNB, with none of the members from these clusters being reassigned to a different grouping. The membership of cluster Group 2 was, however, radically reduced, with a large number of cases being reassigned to the first cluster, but none to the third group. Thus, the abbreviated battery is observed to maintain a good discrimination between the two most extremely situated groups, with no confusion introduced with respect to classifying or grouping together cases that are representatives of either the least impaired (Group 1) or most impaired (Group 3) on the LNNB-A. The pattern of altered group assignments on the LNNB-A compared to the LNNB would suggest that the abbreviation has acted to reduce the observed differences between some cases that would be classified (on the basis of the full LNNB) as mildly impaired and normals. This would be consistent with expectation, as the procedures employed to abbreviate the LNNB acted to remove a substantial number of items that displayed large correlations with many of the LNNB clinical scales (i.e., the items correlating higher with scales other than the one the item was assigned to) such that a reduction in the number of items sensitive just to the general level of impairment (as contrasted to more specific functions reflected by the separate clinical scales) would be expected. The expected net result from such an abbreviation procedure would be to decrease the sensitivity of some scales (especially those with large numbers of the nonspecific items in the original version) to the general impairment level, and result in lowered overall differences between the profiles of individuals with similar types of problems that differ principally in level or intensity of the symptoms or impairments. The use of a similarity metric based on distance (to take the profile elevation into account) would therefore be likely to group more persons together due to the reduction in the size of the differences between the profile elevations.

The patterns of the LNNB-A clinical scale elevations for the identified clusters were next investigated, and the group mean profiles for the three clusters are presented in Figure 3. As normative data for this abbreviated battery is not yet available, the profiles were plotted in this figure based upon a Z score transformation, using the mean and standard deviation for each clinical scale for the total sample as a reference.

Figure 3

Mean Z Scores for LNNB-A Clinical Scales for Three Clusters



Scale 1 = Motor; Scale 2 = Rhythm; Scale 3 = Tactile;
Scale 4 = Visual; Scale 5 = Receptive Speech;
Scale 6 = Expressive Speech; Scale 7 = Writing;
Scale 8 = Reading; Scale 9 = Arithmetic
Scale 10 = Memory; Scale 11 = Intellectual Processes

Cluster Group 1 (N = 91) = ○
Cluster Group 2 (N = 21) = □
Cluster Group 3 (N = 4) = △

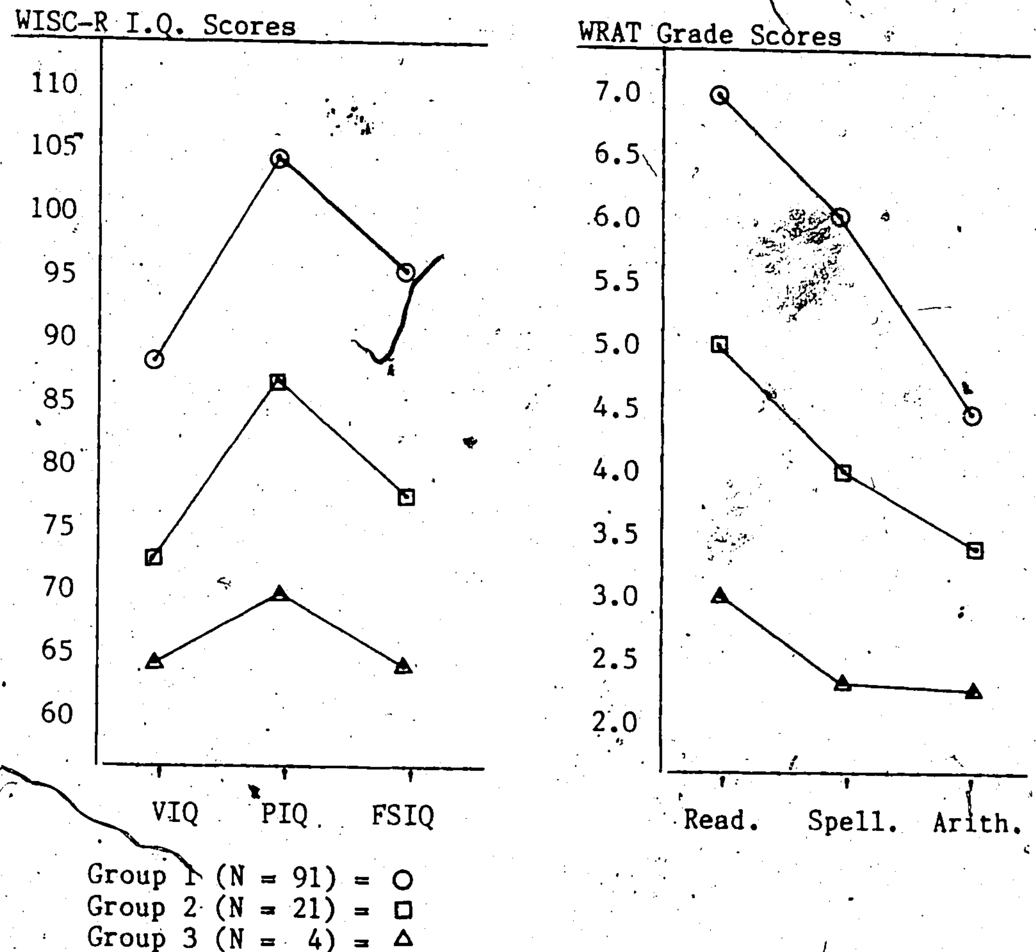
From an inspection of the displayed profiles, it is apparent that there is a clear seperation between the three group profiles, with Group 1 displaying scores that fall consistently below the mean value for the total sample (mean profile elevation for group 1 = -0.30), and with relatively little variability in the profile pattern. Group 2 displays an overall profile elevation (mean = +0.86) that is approximately one standard deviation higher than that of Group 1, and with a slightly increased level of variability displayed. Group 3 displays the most significantly elevated of the three group profiles (mean elevation = +2.23), and a dramatically greater level of inter-scale variability is observed for this group. The relative elevations of these three group profiles from the LNNB-A are quite consistent with those from the cluster groups derived from the LNNB, and point to the presence of clearly significant differences between the three groups with respect to the general level of performance displayed for each group. This finding confirms the prediction of Hypothesis 25 with respect to finding significant differences in the subgroup LNNB profiles.

What is also apparent on these profiles is that there appears to be some significant differences in the profile shapes for the three groups, such that the three groups could be viewed as displaying distinct patterns of neuropsychological abilities or deficits. As Group 1 was the largest of the three subgroups, and these Z scores are based upon the comparison to the mean and standard deviation from the total sample, these patterns of differences can only be interpreted as demonstrating relative differences in performance level across the three groups as compared to the standards set primarily by Group 1, and the identification of areas of strength and weakness as contrasted to a nonclinical or "normal" population is not possible. The identification of distinctly different profile shapes for the identified subgroups from this cluster analysis would, however, be encouraging and suggestive that further research with this abbreviated battery may be successful in identifying meaningful profile configurations that are specific to different clinical subpopulations.

As an external validation procedure for the LNNB-A cluster analysis, the mean elevations for the WISC-R I.Q. scores and WRAT grade scores were investigated for the three cluster groups. These data are presented in graphically in Figure 4.

Figure 4

WISC-R I.Q. and WRAT Grade Score Profiles for LNNB-A Clusters



Inspection of the plotted profiles for the WISC-R I.Q. scores indicates that the three groups, defined on the basis of LNNB-A data alone, were quite different with respect to their overall level of intellectual abilities. Group 1 displayed the highest overall intellectual ability, with an overall or Full Scale I.Q. that fell at an average level (mean F.S.I.Q. for group 1 = 95.4). There was a mildly significant Verbal/Performance I.Q. score discrepancy displayed for Group 1, with the Verbal I.Q. falling just below the mid-average (mean V.I.Q. score = 89.2) while the Performance I.Q. fell solidly in the mid-average range (mean P.I.Q. score = 103.7), a difference of almost 15 points. This is relatively consistent with the patterning of I.Q. scores for the first group from the LNNB cluster analysis, with that group displaying a similar discrepancy of 12 points between the Verbal and Performance I.Q. scores in favor of the Performance I.Q. score.

Group 2 from this LNNB-A cluster analysis displayed an I.Q. score profile that placed the overall ability as falling in the upper end of the borderline deficient range (mean F.S.I.Q. = 78.4), and a significant Verbal/Performance discrepancy of 15.5 points was observed between the mean Verbal I.Q. value of 72.3 and the Performance I.Q. of 87.8. The profile shapes for Groups 1 and 2 are almost identical, and the only significant difference in I.Q. scores between these two groups was on the level of the profile elevations.

Group 3 from this cluster analysis of the LNNB-A was identical to that obtained in the clustering of the full LNNB data. The overall intellectual functioning level for this small cluster of cases fell in the mentally deficient range of ability (mean F.S.I.Q. = 65.0) and there was no significant discrepancy between the elevations observed on the Verbal I.Q. score (65.0) and the Performance I.Q. score (70.8).

A MANOVA was conducted on these intellectual score values and a highly significant overall effect was observed, indicating that these three groups were significantly different with respect to the level of intellectual abilities (Wilks Lambda = 0.553, F = 12.74, P < .001). One way analyses of variance also indicated that the three groups differed significantly on all three I.Q. score variables, as summarized in Table 46.

Table 46

Univariate F Tests for WISC-R I.Q. Scores for LNNB-A Groups

Variable	Hyp. SS	Error SS	Hyp. MS	Error MS	F	P
V.I.Q.	6577.28	12330.27	3288.64	109.12	30.14	0.000
P.I.Q.	7835.87	21289.75	3917.93	188.40	20.80	0.000
F.S.I.Q.	7669.25	11110.95	3834.63	98.33	39.00	0.000

The patterning and relative elevations of the I.Q. score profiles for these three groups is remarkably consistent with that of the three groups derived from the analysis of the LNNB data, despite the significant alteration in the size of the first two groups. As with the I.Q. score profiles for the LNNB derived groups, these profiles display a high degree of similarity of profile shape or configuration between the first and second groups, with substantial Verbal/Performance discrepancies in which the Performance score is higher than the Verbal I.Q. score. The significant intellectual difference between these two groups was observed to be restricted to level of overall ability, with approximately the same amount of Verbal/Performance discrepancy observed for both groups. The WISC-R profiles for each of these groups would be similar to that most commonly observed among most learning disabled youths, with nonverbal intellectual skills falling substantially higher than verbal skill levels. The third group is identical for both cluster analyses and indicates that there were a small number of cases in the sample that were distinct from the balance of the adolescents by virtue of having rather uniformly retarded intellectual skills, both verbal and nonverbal. This third cluster profile is distinct from the other two with respect to having an insignificant difference or discrepancy between it's Verbal and Performance I.Q. scores.

These results would indicate that the LNNB–A cluster solution has external validity, in that the use of the LNNB–A raw score data was successful in separating out three groups of cases from the total sample that were then seen to be significantly different from each other on an alternative measure that was not used in the original clustering process. The prediction of Hypothesis 26 of significant differences in WISC–R I.Q. scores for the three groups derived from the LNNB–A was confirmed.

The educational attainment data presented graphically in Figure 4 also demonstrated that the three derived cluster groups were different from each other in terms of their academic achievement levels, with Group 1 having the highest level of performance on all three WRAT test grade scores, Group 3 having the most severely delayed levels of attainment, and Group 2 falling intermediate to the other two groups. Relatively little difference in the profile patterns is apparent between the three groups, with all three groups displaying the highest score on the Reading grade score and the lowest on Arithmetic, and the differences are primarily related to differences in overall profile elevation. Group 3 was, however, displayed slightly less variability on the WRAT scores, and a relatively consistent level of academic deficiency was suggested, contrasted to the tendency for Groups 1 and 2 to display a relatively greater level of deficit in Arithmetic than in the language arts tests of Reading and Spelling. A MANOVA on these academic score values for the three groups indicated a highly significant overall or group effect (Wilk's Lambda = 0.652, F = 8.82, P < .001). Univariate comparisons of the three groups for each WRAT test variable also indicated that all three WRAT scores were significantly different between the three clusters, confirming the prediction of Hypothesis 27.

Table 47

Univariate F Tests for WRAT Grade Scores for LNNB-A Clusters

Variable	Hyp. SS	Error SS	Hyp. MS	Error MS	F	P
Reading	116.32	254.27	58.16	2.25	25.85	0.000
Spelling	110.40	396.30	55.20	3.51	15.74	0.000
Arithmetic	35.69	138.51	17.85	1.23	14.56	0.000

From the data presented graphically in Figure 4 as well as from the tabulated univariate F tests, it is concluded that the three clusters derived from the LNNB-A data represent subgroups among the total sample that have significant differences between them with respect to both intellectual as well as academic attainment levels. As with the clusters derived from the full LNNB data, these groupings are observed to differ from each other on the external predictors primarily on the level of performance, with differences in the patterning of intellectual ability or academic attainment only observed for the rather small third cluster group. The failure to obtain significant differences with respect to the patterns of either intellectual or academic abilities between Groups 1 and 2 would be seen as reflecting either a lack of sensitivity or coverage in the retained items from the LNNB to detect such differences in the sample, or a problem of inadequate sample composition, such that there were too few cases of radically different patterns of academic learning deficit to allow for the generation of reliable and significantly distinct learning disability subgroups.

The fact that the four cases that were grouped into the third cluster for both the LNNB and the LNNB-A cluster analyses were observed to display both a difference in the overall profile elevation as well as profile shape on the WISC-R would support an interpretation that a large portion of the difficulty could be attributed to the lack of sampling coverage. This third group is observed to be different from the balance of the sample in several important ways, with its LNNB-A profile displaying a substantially greater level of scatter or variability than either of the other two cluster groups, it's WISC-R profile displaying the least variability with a uniformly low level of ability across the three I.Q. scores, and its academic attainment levels also displaying a pattern of rather uniformly low achievement in all three areas tapped by the WRAT. It is apparent that with respect to identifying at least this one subgroup from the total sample, both the LNNB and LNNB-A were equally successful in terms of displaying a stable, reliable, and meaningful subgroup in the sample, with this group representing those cases that had a generalized level of intellectual retardation, uniformly severe academic attainment, and a high level of apparent neuropsychological impairment. The other two cluster groups identified are less distinct, and display similar patterns of intellectual and academic ability, but at different levels of general functioning.

As noted previously in the discussion regarding the linear discriminant analysis of the LNNB data, the obtained sample for this study was derived from consecutive referrals for clinical neuropsychological evaluation on adolescents, and that the obtained sample composition was undoubtedly skewed by this referral process. As a consequence, an atypical proportion of the sample was observed to display either learning deficits restricted to arithmetic or global learning delays in which all subject areas (including arithmetic) were impaired. From the total sample of 116 cases, only two cases were obtained of a selective learning disability in the language arts area with normal arithmetic attainment level. Thus, there would be an expected high degree of overlap between the deficits that were displayed by the two subtypes of learning disorder represented in this sample, with deficits in intellectual and academic skills needed for arithmetic achievement being common to both groups, while expected or hypothesized differences between these two types of learning disorders would need to be restricted to skills that are essential for language arts attainment but not for arithmetic.

The failure to obtain a sizable number of cases with selective learning disabilities in the language arts area and normal arithmetic ability would be seen as restricting the opportunities to display a dissociation between the learning disability subgroups, and likely played a large role in the lack of obtained differences in intellectual and academic ability patterns between the first two cluster groups. Consequently, the cluster analysis of the LNNB-A data can be considered to have been at least partially successful in clearly discriminating between the cases with a global or general intellectual retardation and those with either normal abilities or learning disabilities, but the current clinical case data are inadequate to allow for a full test of the LNNB-A's ability to discriminate between cases with significantly different subtypes of selective learning disability.

CHAPTER VI

DISCUSSION

The primary goal of this investigation was to explore the utility of the Luria - Nebraska Neuropsychological Battery for use with an adolescent population and with a focus upon its potential as an instrument to evaluate the neuropsychological factors contributing to academic learning disabilities. To this end, it was necessary to investigate the basic psychometric properties of reliability and item to scale consistency of the battery, as well as conduct evaluations of the construct and concurrent validity of the battery as a measure of academically relevant abilities (both intellectual and achievement). A preliminary investigation of the ability of the LNNB to discriminate between predefined groups of adolescents with differing patterns of academic attainment as well as to empirically group adolescents into more homogeneous subgroups or clusters was also conducted. In this section, the results of this research will be interpreted with respect to the stated research hypotheses, and with respect to theoretical issues underlying the hypotheses. Limitations of the study, as well as theoretical and pragmatic implications will also be addressed.

Hypothesis 1 predicted that the LNNB would display internal consistency estimates of reliability that were equivalent to similar values derived from adult studies. The obtained results supported this prediction, with internal consistency estimates falling at a level that was intermediate between those obtained from normal adults and adult patient groups with more severe neurological or psychiatric disturbances. The level of internal consistency obtained in this study is felt to be quite adequate for research and clinical applications, particularly in view of the effects of a somewhat restrictive range of functional impairments as well as very restricted range of age and educational exposure in this sample.

Hypothesis 2 predicted that the clinical scales from the LNNB would display levels of item to scale correlation consistency that would be equivalent to those reported for a similar investigation using an adult sample. This prediction was not supported by the current study, as a significantly lower percentage (66.9%) of the LNNB items in this study were seen to correlate maximally with their assigned scale than on the previously reported study of Golden, Fross and Graber (1981), in which their adult sample displayed a 92.9 percent item to scale consistency. This finding is viewed as a potentially serious challenge to the construct validity of the individual scales of the LNNB, as the procedure of summing the individual item scores for each scale and then contrasting the relative elevations of the clinical summary scales is dependent upon the individual scales being distinct from each other in some significant manner. To the extent that an item correlates higher with a scale other than its own, the item can be viewed as having a contamination effect on the scale summary score, and it would act to inflate the intercorrelation of the scales. Differences between the summary scale score elevations would thereby tend to be reduced, and the interpretability of the individual scales would be weakened by the inclusion of large numbers of such cross-correlating items. The observation of a dramatically higher percentage of such items in the results from this adolescent population would bring the construct validity of the LNNB for such an age range into question, and suggested a need for pursuing an abbreviated version of the battery to reduce the level of such cross-correlating items.

Hypothesis 3 predicted that the derivation of such an abbreviated battery would be possible without diminishing the internal consistency of the clinical scales. Hypothesis 4 also predicted that such an abbreviated form would have an improved level of item to scale consistency that would be comparable to that reported for the LNNB in adult populations. The results obtained in this portion of the study indicated that elimination of all items that displayed correlations with other scales that exceeded that with their own would adversely affect the internal consistency of three clinical scales, and as a small percentage of such cross-correlating items has been recognized and accepted in the adult studies with the LNNB, modification of the exclusionary criteria was explored, resulting in the retention of several items that had high correlations with their assigned scale but equally high or marginally higher correlations with other scales. This resulted in the identification of 189 items that comprised an abbreviated battery (designated the LNNB-A) that had equal or higher internal consistency than the full 269 item LNNB battery, and which had an item to scale consistency that was equivalent to the level previously reported for the adult population, with 92.6 percent for the LNNB-A compared to 92.9 percent for the adult sample in the Golden et al. (1981) study. The predictions of both Hypotheses 3 and 4 were confirmed by these results for the LNNB-A, and indicated that the item selection for the LNNB-A represented an adequate compromise between the demands of scale length to maintain reliability and item to scale consistency to maintain construct validity.

Hypothesis 5 predicted that there would be a strong primary factor related to the level of performance that would be represented in all the LNNB summary clinical scale scores. This prediction was based upon the neuropsychological literature that indicates that in children the effects of neurological damage is particularly striking on the level of overall intellectual development (c.f., Reitan & Davison, 1974). As the clinical sample from which this study was derived was composed primarily of youths with long-standing difficulties with learning and/or emotional/behavioural disturbances, it was hypothesized that any significant neuropsychological dysfunctions present in these cases would likely be of a chronic/static nature, and that interferences with general development or functioning level would be expected. The previous research which also indicated that there was a high level of correlation with intelligence for all of the LNNB clinical scales was also supportive of this prediction.

The results from the principal components factor analysis of the LNNB data strongly supported this prediction, and only a single eigenvalue in excess of unity was obtained, and with all factor loadings with this initial principal component being highly significant. This result supported the interpretation that the LNNB clinical scales are very sensitive to the overall or general level of adaptive functioning or intellectual ability, and the consistency of these factor loadings across the various LNNB scales may be viewed as supporting the view that neuropsychological impairments in children and adolescents are likely to be reflected primarily in deficits in general level of functioning.

Hypothesis 6 predicted that in addition to the loadings on an overall general ability factor, there would be interpretable group factors emerging from the LNNB data that would display neuropsychologically meaningful dimensions in the data particularly with respect to the identification of specific verbal and nonverbal abilities. Examination of the eigenvalue structure of the LNNB data indicated that these eigenvalues were essentially similar to those reported for the Wechsler intelligence scales, and that a three factor model would be supported by the data. Three group factors were identified after Varimax rotation to simple structure, and were interpreted as a Verbal Intelligence factor, an Academic Attainment factor, and a nonverbal Perceptual/Motor factor. The scale compositions of these factors were consistent with the expected occurrence of factors reflecting left hemispheric abilities (Verbal Intelligence), right hemispheric abilities (Perceptual/Motor) and an interactive effect of both underlying neuropsychological competencies and environmental experiences and opportunities to develop crystallized skills (Academic Attainment). The prediction of Hypothesis 6 was confirmed by the obtained results of this factor analysis.

Hypotheses 7 and 8 indicated that the abbreviated LNNB-A should also display a strong loading for all scales on the first principal component and that in a manner equivalent to the LNNB there would also be interpretable group factors that were neuropsychologically meaningful. These predictions were both confirmed by the principal component factor analysis of the LNNB-A clinical scale data, with a high loading on an initial factor as well as identification of three group factors being supported by the eigenvalue structure. Hypothesis 9 also predicted that there would be a correspondence between the factor structure of the LNNB and LNNB-A, and this was also confirmed, with the identification of equivalent Verbal Intelligence, Academic Attainment and Perceptual/Motor factors for the LNNB-A. This was interpreted to mean that the process used in abbreviating the battery to create the LNNB-A had been successful in preserving the dimensional structure of the battery, and had not significantly reduced the functional coverage of the battery.

Hypothesis 10 predicted that the LNNB Intellectual Processes scale would correlate higher with the WISC-R I.Q. scores than any of the other clinical scales from the LNNB. This prediction was observed to be confirmed with respect to the correlations with the Verbal I.Q. score, but it was not found to hold for either the Performance I.Q. nor the Full Scale I.Q. scores, which displayed higher correlations with other LNNB clinical scales. This finding resulted in a reappraisal of the meaning of the Intellectual Processes scale, and it was interpreted as reflecting principally verbal intellectual abilities, rather than being a measure of general (both verbal and nonverbal) intelligence.

Hypotheses 11, 12, and 13 predicted that the LNNB academic scales of Writing, Reading and Arithmetic would display stronger correlations with the equivalent WRAT grade scores (Spelling, Reading and Arithmetic) than would any other of the LNNB clinical scales. These predictions were all supported by the obtained results, but the size of the difference in the correlations between the LNNB Writing and Reading scales and the WRAT Spelling grade score was extremely small. The results were interpreted as supporting the concurrent validity of these academic scales from the LNNB, while noting the strong intercorrelation between the language arts variables for both the LNNB and the WRAT made it impossible to achieve a high level of differentiation between measures of reading and spelling abilities.

Hypothesis 14 predicted that the abbreviated LNNB-A Intellectual Processes scale would also display significant and strong correlations with the WISC-R I.Q. variables. The results confirmed this prediction with respect to the correlations with both the Verbal and Full Scale I.Q. scores, but the Performance I.Q. score was observed to display stronger correlations with scales other than Intellectual Processes. These results confirmed that the abbreviation process had not changed the essential bias of the LNNB toward coverage of verbal functions, and the Intellectual Processes scale of the LNNB-A was interpreted to be principally a measure of verbal rather than general intellectual functioning.

Hypotheses 15, 16, and 17 predicted that the LNNB-A academic scales would display stronger correlations with the equivalent WRAT grade scores than any other LNNB clinical scales. These predictions were confirmed for the correlations with the WRAT Reading and Arithmetic scores, but the WRAT Spelling grade score was observed to correlate equally with both the LNNB Writing and Reading scales. The failure of the LNNB Writing scale to display a higher correlation with WRAT Spelling than the LNNB Reading scale was interpreted as a reflection of the high degree of functional overlap between the two language arts areas, and the intercorrelation between the WRAT Spelling and Reading tests of 0.87 (which is higher than the correlations between the LNNB and WRAT variables) was noted as a limiting factor in obtaining a clear differentiation between these two language arts areas.

Hypothesis 18 predicted that multiple regression formulae would be identified that would allow for predicting both intellectual as well as academic achievement scores from the LNNB clinical scale data. An arbitrary standard of a multiple correlation of 0.70 was selected as reflective of a high level of predictive validity for these formulae. The results confirmed these predictions for all three academic achievement variables, and for the Verbal and Full Scale I.Q. variables from the WISC-R. Only the multiple correlation with the Performance I.Q. score failed to exceed the 0.70 level, but this correlation was observed to fall at a level (0.632) that was still quite significant on a statistical level. These multiple regression formulae were also observed to display inclusion of clinical scales and regression weights that were consistent with expectation, with the Verbal I.Q. being best predicted by a combination of the Intellectual Processes, Expressive Speech and Rhythm scales, the Performance I.Q. by the LNNB Motor scale, and the Full Scale I.Q. by a combination of the Intellectual Processes and the Motor scale. WRAT Reading was predicted best by a weighted combination of the LNNB Reading, Writing and Expressive Speech scales, the WRAT Spelling score was adequately predicted by just the LNNB Writing scale, and the WRAT Arithmetic score was predicted by a combination of the LNNB Arithmetic and Writing scales.

These multiple regression analysis results supported the validity of the LNNB as an instrument that can be used to predict academic achievement and some aspects of intellectual ability with a high degree of accuracy. Only nonverbal intelligence was found to be somewhat weakly represented in the LNNB, and this is consistent with the theoretical biases of A. R. Luria, whose research was particularly focused on studying the verbal functions of the language dominant left hemisphere, and the manner in which it is responsible for controlling and directing the activities of the whole brain.

Hypothesis 19 predicted that multiple regression analyses using the LNNB-A to predict intellectual and academic abilities would also result in highly significant levels of predictive validity. As with the results from the full LNNB, highly significant multiple correlations were obtained between the LNNB-A and all three WRAT academic scores as well as between the LNNB-A and the Verbal and Full Scale I.Q. scores from the WISC-R. Only the multiple correlation with Performance I.Q. failed to exceed the 0.70 criteria level, and this confirmed that the abbreviated battery was similar to the full LNNB in its relative weakness as a measure of nonverbal intelligence.

Hypothesis 20 predicted that the use of the LNNB clinical scale data in a linear discriminant analysis would allow for the discrimination between a priori defined cases with differing patterns of academic functioning on the WRAT. This prediction was confirmed at a high level of significance, with a 76.3 percent hit rate for identifying the specific pattern of academic performance, and a 91.2 percent accuracy rate for the less specific discrimination between normal and learning disabled (unspecified as to type). This finding was interpreted as indicating that the LNNB has a good potential for practical application within an educational context, and it was also noted that it had a particularly low false negative (Type I) error rate.

Hypothesis 21 predicted that the LNNB-A clinical scale data would also be successful in predicting the pattern of academic functioning on the WRAT. This prediction was also confirmed by the results of the study, and an identical overall hit rate was obtained for the LNNB and the LNNB-A, indicating that the abbreviation process had not significantly altered the battery's coverage of functions important to educational achievement. As with the LNNB, the LNNB-A linear discriminant analysis also displayed a tendency to make a higher number of false positive errors in assigning normal learners to one of the learning impaired groups, while very few of the most severely impaired cases were misassigned to either the intermediate or the normal group. This tendency to make a higher number of false positive errors, which would likely act to precipitate further assessment or intervention for the individual case, was felt to be more acceptable in an educational context than a test that made a larger number of false negative errors (failing to recognize the existence of a real problem) which would act to reduce the probability of the individual receiving additional help or remediation.

Hypothesis 22 predicted that the application of clustering methodology to the LNNB clinical scale data would result in the identification of distinct subgroups within the sample that had significant differences in their LNNB clinical scale profiles. The clustering process operates to group cases together on the basis of the similarities in their profiles, and the finding of reliable and stable subgroups within the data through the clustering process is a reflection of the existence of distinct subtypes of cases in the sample, rather than of a homogeneous collection of essentially similar cases. The finding that the use of two different clustering methods resulted in almost identical group memberships was observed to confirm the stability or reliability of the cluster solutions, and for at least one subgroup (the most severely impaired group) all applied methods resulted in identical cluster formations. The prediction of significant differences between the group profiles was tested by use of MANOVA, and a highly significant overall effect was observed, and univariate F tests confirmed that all eleven clinical scales were significantly different between the three groups.

Hypotheses 23 and 24 predicted that the derived cluster groups from the analysis of the LNNB data would be found to show significant differences on intellectual and academic test variables that were not utilized during the clustering process. Both of these predictions were confirmed with a high degree of significance, both with MANOVA evaluation of overall effects as well as univariate F tests that confirmed that all of the I.Q. and WRAT grade scores were significantly different across the three groups. These results confirmed that the application of a clustering procedure had identified subgroups of cases on the basis of the LNNB scores that had highly significant differences on external validation measures, indicating that the derived clusters were displaying educationally relevant differences in both intellectual abilities as well as academic attainment. This was interpreted as supporting the validity of the LNNB as an instrument for educational evaluation purposes.

Hypothesis 25 predicted that the LNNB-A clinical scale data would also be successful in clustering the sample into discrete subgroups that would display significant differences on their LNNB-A profiles. This prediction was tested by MANOVA and a highly significant difference was found between the three cluster groups with respect to their overall profiles. Univariate F tests also confirmed that significant differences were obtained for all eleven clinical scales.

Hypotheses 26 and 27 predicted that the cluster groups derived from the LNNB-A data would also display significant differences on their WISC-R and WRAT profiles, and these predictions were confirmed by MANOVA evaluation of the overall effect, and univariate F tests confirmed significant differences for all of the separate intellectual and academic achievement test variables. These results confirmed that the cluster solution derived from the LNNB-A data had been successful in identifying three subgroups within the sample that were significantly different not only on the basis of their LNNB-A profiles but also on intellectual and academic achievement test variables that were not incorporated into the clustering process. This would validate the significance of these derived clusters, and support the interpretation that the differences between the cluster groups on the LNNB-A were reflecting real differences between cases that were of significance for other aspects of their lives, particularly academic achievement and intellectual functioning. These external validations of the LNNB-A cluster solution would therefore also reflect upon the validity of the LNNB-A as an instrument that is capable of measuring aspects of neuropsychological functioning that are relevant to educational and intellectual abilities.

Overview of the Results

Reliability Evaluations

The results of this study have, overall, been observed to largely confirm a large number of the predictions made in the hypotheses. It is observed that predictions made regarding the verification of the reliability (internal consistency) of the clinical scales of the LNNB as well as the LNNB-A have been upheld, and the level of reliability of these batteries is observed to be essentially equivalent to that previous established for the LNNB with use with adult populations, given the nature of the current sample and the type of internal consistency metric chosen. The establishment of the reliability of this battery for an adolescent age population has, to the best of this writer's knowledge, never been published to date, despite the recommendations of the test authors and publisher that the LNNB can be effectively applied to cases down to the age of 13 years (Golden, Purisch, & Hammeke, 1985).

As indicated in the <u>Standards for Educational and Psychological Tests</u> (APA, 1974), it is essential for the procedures and samples used in the determination of reliability coefficients to be described fully and for tests that claim applicability to a wide range of age or grade levels to have reliability reported specifically for the various grade or age levels recommended. This study would therefore expand the literature on the LNNB in providing a replication of the initial reliability evaluation for the lower end of the age range recommended by the test authors. The obtained reliability estimates for both the LNNB and the LNNB-A (average Alpha coefficients of 0.79 with a range of 0.71 to 0.87 for both forms of the battery) are judged by this investigator to be sufficient for the use of this battery in research and applied settings, particularly in light of the limiting factors of the somewhat restricted range of age and levels of impairment displayed in this clinical sample.

<u>Validity Evaluations</u>

In addition to the establishment of the basic psychometric property of reliability, it is also essential for the user of a particular test to have some appropriate information regarding the validity of the test or instrument. Validity has been defined as: "what the test measures and how well it does so" (Anastasi, 1968, p.

99) and it has also been stated that: "validity refers to the appropriateness of inferences from test scores or other forms of assessment. The many types of validity questions can, for convenience, be reduced to two: (a) What can be inferred about what is being measured by the test? (b) What can be inferred about other behaviour?" (APA, 1974, p. 25). The validation of a test represents an inferential process in which the investigator is called upon to make informed judgements regarding the value or utility of the instrument in the context or setting that the test is to be applied to:

> "It is important to note that validity is itself inferred, not measured. Validity coefficients may be presented in a manual, but validity for a particular aspect of test use is inferred from this collection of coefficients. It is, therefore, something that is judged as adequate, or marginal, or unsatisfactory." (APA, 1974, p. 25).

Validity is also categorized into four discrete but interdependent forms: content validity, construct validity, and criterion related validities (predictive validity and concurrent validity). Content validity involves the examination of the test content to assess whether the instrument covers a sufficiently representative sample of the behavioural domain the test purports to assess. With respect to this aspect of validity, it is relevant to note that the stated objective of the authors of the LNNB was to establish a standardized version of the Luria neuropsychological investigation process, to allow for replication of results across settings and to allow for the experimental investigation of the effectiveness of Luria's procedures for the identification of brain damage. The LNNB has also been described by its authors as: "a multidimensional battery designed to assess a broad range of neuropsychological functions... its primary purpose is to diagnose general and specific cognitive deficits, including lateralization and localization of focal brain brain impairments, and to aid in the planning of rehabilitation programs" (Golden et al., 1985, p. 1). From this description of the intents of the battery, it can be asserted that the identified behavioural domain to be measured by this battery would include all aspects of neuropsychological functioning, particularly cognitive skills, that are susceptible to the effects of brain damage, and which would be important to be be considered in a rehabilitation planning process.

With respect to the functional areas that are considered to be important within the context of a comprehensive neuropsychological evaluation, there have been a number of alternative categorical schemes proposed over the years, including the functional labels proposed by Luria (1966, 1973) and operationalized into specific test procedures by Christensen (1975), as well as by other researchers and authors such as Lezak (1976) and Adams, Rennick, and Rosenbaum (1975). Swiercinsky (1978) compared a series of such alternative conceptual schemes and outlined a relatively comprehensive listing of major functional areas that have been described by various independent researchers as important facets of neuropsychological functioning: 1) Lateral dominance, 2) Motor functioning, 3) Auditory perceptual organization, 4) Tactile perceptual organization, 5) Spatial perceptual organization, 6) Language skills, 7) General information processing, 8) Memory processes, 9) Attention and concentration, and 10) Education and experience. He also noted that several of these categories were subject to further subdivision into more specific subskill areas, such as the separation of motor functioning into distinct factors of motor strength, gross motor speed and fine motor coordination or the division of spatial perceptual organization into non-visual and visual spatial perception.

Adoption of the conceptual scheme of Swiercinsky (1978) as a guideline for the definition of the behavioural domain to be assessed in a neuropsychological evaluation would allow for some evaluation of the content validity of the LNNB. With this in mind, it is noted that the initial construction of the LNNB was based upon the published items from the Christensen (1975) neuropsychological investigation, and that there was a systematic attempt to include items that were representative of the majority of these functional areas. The structured and standardized version of these Luria tests as embodied in the LNNB is seen to include items within each of these conceptual areas, with the one exception of specific tests for lateral dominance. With respect to the issue of breadth of coverage or the content validity, the LNNB would be rated as appearing to display a sufficient coverage of the functional areas important to the comprehensive neuropsychological evaluation.

However, the issue of content validity is also based not only on the presence of items that appear to tap the specified functions, it is also important to determine if the level of representation of the functions is appropriate:

> "The content area must be systematically analyzed to make certain that all major aspects are adequately covered by the test items, and in the correct proportions. For example, a test can easily become overloaded with those aspects of the field that lend themselves more readily to the preparation of objective items.... Moreover, content validity depends on the relevance of the individual's test responses to the behavior area under consideration, rather than on the apparent relevance of item content. Mere inspection of the test may fail to reveal the processes actually used by subjects in taking the test." (Anastasi, 1968, p.100).

With these considerations in mind, the content validity of the LNNB would be open to some debate, as the previous critiques of this battery have noted that difficulties with the construct validity of some scales and the manner in which language functions may contaminate many of the ostensibly nonverbal scales (such as the Tactile and Visual scale) would make it difficult to accept at face value that the scales are actually measures of these specific functions. During the course of this study it has been observed that the LNNB clinical scales display extremely high intercorrelations and that there is a strong loading for all scales on a general ability factor. Furthermore, it has repeatedly been demonstrated on the correlational investigations as well as the multiple regression analyses that the LNNB has a considerably higher loading on the verbal than on a nonverbal intellectual factor, and that all of the clinical scales correlate with a high level of significance with the intellectual variables of the WISC-R. This would suggest that the content of the LNNB has been strongly biased toward the verbal intellectual realm, and that a disproportionate number of items have been included that are dependent upon verbal language mediation or higher level verbal intellectual skills. While there are items that appear to tap nonverbal capabilities, the results of this study would suggest that they are insufficient in number and/or discriminative power to allow for a balanced representation of all the cognitive functions (both left and right cerebral hemisphere functions) required for a fully comprehensive neuropsychological assessment. This interpretation from the current adolescent data would also be supported by the findings of the Goldstein, Shelly, McCue and Kane (1987) cluster analytic study, in which those investigators reported that the LNNB failed to lateralize between left and right hemisphere damage, and that this battery was relatively insensitive to the effects of right hemispheric brain damage.

Construct validity has been defined as: "the extent to which the test may be said to measure a theoretical construct or trait" (Anastasi, 1968, p. 114). This aspect of the test validation process is dependent upon demonstrating that the test variables relate in some predictable manner to theoretical concepts or constructs relevant to the behavioural domain, such as intelligence, verbal fluency, short term memory, visual-spatial skills, etc. The construct validity of a test is generally established through the accumulation of research results that support the interpretation of the test variables in a specific manner.

One approach to the issue of construct validity is that of examination of the internal consistency or item-discrimination coefficients in which the degree to which each item correlates with the total score of the test (or scale) is assessed, and items retained in the test only if they correlate reliably with the total score. As noted previously, the LNNB clinical scales were found to have reasonably robust levels of internal consistency (Alpha coefficients). In addition, Golden, Fross, and Graber (1981) also examined the extent to which items on the LNNB correlated with their own scale as contrasted to other clinical scales on the battery, and reported that approximately 93 percent of the LNNB items correlated higher with their assigned scale than with any other clinical scale. In this study, this level of item to scale correlation consistency was not found, and only a 67 percent of the LNNB items were found to display such consistency. This would suggest that while there is an apparently adequate level of internal consistency within each scale for the summary score to be reliable, there is a high degree of overlap in functional content across the scales, so that the validity of identifying scales as measures of discrete functions would be questionable.

Another approach to construct validation is that of factor analytic evaluation to identify the existence of differentiable psychological traits. In the current study, factor analysis of the LNNB identified the presence of an extremely high loading on a general level of performance or impairment level factor, and confirmed that there was a substantial level of overlap between the clinical scales. Three group factors were also identified that accounted for approximately 77 percent of the total variance, and these were interpreted as reflecting Verbal Intelligence, Academic Attainment and Perceptual/Motor factors. The limited number of factors identified from the eleven clinical scales suggests that the disparate titles provided for the separate clinical scales may be somewhat misleading, as there is a lack of evidence that they are actually measuring significantly different functions, and the breadth of dimensional coverage of the LNNB would appear to be considerably lower than suggested by the division into eleven distinct scales.

The results of the current evaluation would suggest that with respect to the construct validity of the LNNB, there is a lack of consistency between the item to scale correlations for a significant number of items, and that the level of intercorrelation of the clinical scale variables and the factor structure of the battery does not support the validity of the separate clinical scales as they are currently published. An exploration of the item to scale correlations allowed for the identification of an abbreviated battery that had a substantially reduced number of items (by elimination of most items that did not correlate the highest with their assigned scale) and which still retained internal consistency reliability that was equal to the full battery, and which correlated with the full battery at a very high level (average correlation between LNNB and LNNB-A = 0.982). This abbreviated battery was found to display an improved level of item to scale correlation consistency, but on factor analysis an essentially identical factor structure was obtained. The consistency between the full LNNB and the abbreviated LNNB-A would support the interpretation that the dimensional structure of the LNNB is restricted to three principal factors, and that even elimination of nonspecific items that may have acted to inflate the loading on the general factor did not lead to identification of more than three significant group factors. These results would serve to challenge the description of the LNNB as a "multidimensional battery designed to assess a broad range of neuropsychological functions" as stated by the test's authors.

The factoral structure and breadth of coverage of the LNNB can also be contrasted to the dimensional structure of other neuropsychological batteries for comparison purposes. The other widely used test battery in this field is the Halstead - Reitan

Neuropsychological Battery (HRNB) and it has also been subjected to factor analysis in a number of studies. The dimensional complexity or coverage of the two batteries can be contrasted by consideration of the number of significant factors obtained using a standard factor extraction termination process, such as Kaiser's rule (Kaiser, 1960), which accepts for interpretation only a number of factors equal to the number of eigenvalues greater than unity. Using this criteria, only a single factor would be accepted for the current study's LNNB data, and only two for the LNNB-A data. In contrast, there have been studies of the HRNB that have applied the same criteria to principal components factor analysis, and which have identified factor structures reported as displaying between four (Goldstein & Shelly, 1972) to five (Goldstein & Shelly, 1971; Strom, Fischer, & Dean, 1987) principal components factors with eigenvalues in excess of unity. With respect to the construct validity of the LNNB, the current study's results would suggest that there is insufficient coverage of the broad range of functional dimensions that have been identified by a variety of researchers and theorists in the area of neuropsychology to warrant describing this battery as a reasonably comprehensive test of neuropsychological abilities.

Criterion-related validity has been defined as: "the effectiveness of a test in predicting an individual's behavior in specified situations. For this purpose, performance on a test is checked against a criterion, i.e., a direct and independent measure of that which the test is designed to predict." (Anastasi, 1968, p. 105). The criterion-related validity of the LNNB was investigated in this study by exploring the concurrent validity of the battery with comparison to widely accepted instruments of intellectual functioning (WISC-R) and academic attainment (WRAT). As the criterion variables of intelligence and academic attainment were not designed or created to represent measures of neuropsychological impairment (as was the LNNB), the critical focus in these comparisons was held to be the degree of validity for the specific LNNB scales that are identified as measures of intellectual and academic performance.

In the examination of the concurrent validity of the Intellectual Processes scale, the correlations of the LNNB with the three I.Q. scores of the WISC-R were examined. Consistent with the prediction of Hypothesis 10, the Intellectual Processes scale was observed to correlate significantly with the three WISC-R I.Q. variables, but significant correlations were also observed between all the other clinical scales and the WISC-R I.Q. variables. The further prediction that the Intellectual Processes scale would correlate with the WISC-R variables higher than any other clinical scale was upheld only for the correlations with the Verbal I.Q., while for both the Performance and Full Scale I.Q. variables there were other LNNB clinical scales that correlated higher than the Intellectual Processes scale. This would indicate that the Intellectual Processes scale, as it exists currently on the LNNB, is predominantly a measure of verbal intellectual ability, and that it has insufficient coverage of nonverbal intellectual factors to be equated with general intelligence. As a measure of verbal intellect, however, it is relatively robust (correlation = 0.73 with the WISC-R V.I.Q.), and if interpreted as a specific verbal intellectual measure, it would be considered to have adequate concurrent validity.

The concurrent validity of the LNNB academic scales of Writing, Reading and Arithmetic were also examined with regard to their correlations with the WRAT variables of Spelling, Reading and Arithmetic. Consistent with prediction, the three academic scales all correlated with a high degree of significance with their equivalent variables from the WRAT, and in each case the correlation between the specific LNNB academic scale and the WRAT variable was higher than between that WRAT variable and any other LNNB clinical scale. This is interpreted as being a significant confirmation of the concurrent validity of these academic attainment scales on the LNNB.

The concurrent validity of the LNNB as a measure of intellectual and educational abilities was further investigated by means of multiple regression analysis, using the clinical scale scores to predict WISC-R I.Q.'s and WRAT grade scores. The findings of this evaluation indicated that weighted combinations of the LNNB clinical scales would result in highly significant multiple correlations with each predictor variable. It was also observed that, with the single exception of the regression formula to predict Performance I.Q., the first variable to enter the regression equation was consistent with expectation, such that for both the Verbal and Full Scale I.Q. equations the Intellectual Processes scale was the first variable entered, and the LNNB Reading, Writing and Arithmetic scales were the first variables entered for the equations to predict WRAT Reading, Spelling and Arithmetic, respectively.

Multiple correlation coefficients obtained also indicated that both Verbal and Full Scale I.Q. scores could be predicted with a very high level of significance, with coefficients in excess of 0.70, while the Performance I.Q. multiple correlation was significantly weaker, falling at only a 0.632 level. This would confirm the relatively limited coverage of nonverbal intellectual skills in this neuropsychological battery, in contrast to a more extensive representation of verbal intellectual tasks. Multiple correlation coefficients between the LNNB and WRAT were observed to be extremely robust, with all coefficients exceeding the 0.70 level. Language arts (reading and spelling) were, however, more adequately predicted than the arithmetic variable, which is consistent with the observation that items requiring functional reading skills are located in several scales of this test, including the Expressive Speech, Writing, Memory, and Intellectual Processes scales in addition to the Reading scale.

The correlational comparisons as well as the multiple regression analyses both indicated that the LNNB can be viewed as a valid measure of both verbal intelligence as well as academic achievement for the adolescent age group tested in this study. Nonverbal intellectual ability can also be predicted with a statistically significant level of validity, but it is clearly less thoroughly represented on the battery.

The concurrent validity of the LNNB-A was also investigated using the same procedures of correlation with the WISC-R and WRAT, as well as through using the LNNB-A clinical scale scores to predict WISC-R and WRAT variables with multiple regression analyses. The results of these investigations indicated that the Intellectual Processes scale from the abbreviated battery remained a strong measure of verbal intelligence and a moderately strong measure of general intellectual functioning, but a relatively weak measure of nonverbal intelligence. The LNNB-A academic scales were also observed to correlate very highly with their equivalent variables from the WRAT, and for both the WRAT Reading and Arithmetic scores, the corresponding LNNB-A clinical scale displayed a higher correlation than any other clinical scale. The WRAT Spelling grade score correlated with equal strength to both the LNNB-A Writing and Reading scales, and examination of the intercorrelations between the two language arts scales of the LNNB-A as well as between the Reading and Spelling scales of the WRAT would suggest that these academic skills are very highly intercorrelated and probably reflect a common factor or skill. Multiple regression analyses of the

LNNB-A to predict WISC-R and WRAT variables were observed to be equivalent to those obtained from similar procedures with the LNNB data, with only the multiple correlation with Performance I.Q. displaying a reduction in coefficient value. This would support the interpretation that the abbreviation procedure has not significantly reduced the validity of the battery for predicting verbal intellectual or academic skills, but the weakness of the full LNNB battery in assessing nonverbal intellectual abilities continues to be present and has been mildly increased by the abbreviation process.

An additional approach to the concurrent validity evaluation was to explore the ability of the LNNB to accurately identify subgroups of students that had been classified a priori according to patterns of academic achievement on the WRAT. The linear discriminant analysis of the LNNB data resulted in a highly significant overall classification hit rate of 76.3 percent. This result, while at present in need of cross validation, would suggest that the data contained in the LNNB clinical scale profile can be related to differing patterns of academic performance, and that while there may be some question as to the breadth of the functional dimensions covered by the LNNB, it has sufficient information to discriminate between cases with differing patterns of educational achievement. As recent trends in educational theory regarding learning disabilities emphasize the role of neuropsychological factors, this finding would support the validity of the LNNB as a suitable measure of such educationally related factors. A similar linear discriminant analysis of the LNNB-A data resulted in an identical overall hit rate of 76.3 percent, and suggested that the abbreviated battery was equally valid for educational classification purposes.

A further exploration of the concurrent validity of the LNNB was conducted by means of cluster analysis of the data from this adolescent sample, to determine if the LNNB clinical scale data could be subgrouped into clusters that had educational significance. A number of cluster methods were investigated and evaluated to determine the most stable and suitable cluster solution. Use of an average linkage between groups method and a squared Euclidean distance similarity metric was adopted as the most suitable procedure for describing the naturally occurring subgroups within this data set, and a three cluster solution was determined to be optimal. The obtained cluster groups were observed to display highly significant differences between with regard to the LNNB clinical scale profiles, with the three groups being identified as representing a group with normal neuropsychological functioning, a group with mild neuropsychological dysfunction and a third group with marked neuropsychological impairment. The profiles were distinct with respect to profile elevation, and there was also a tendency for increased profile scatter or variability to occur with increased elevation. The general profile configurations or patterns of relative strength and weakness were, however, essentially similar between the three groups and the effect of elevation was judged to be the most significant discriminating feature between these three group profiles.

The external validation of the obtained three cluster solution was observed to display highly significant differences between the three groups on both the WISC-R and the WRAT variables, despite their independence from the clustering process. There was a clear seperation between the three groups with respect to both intellectual ability and educational achievement levels, a finding that confirmed that the groups were significantly distinct with respect to these educationally relevant factors. The profile configurations on both the WISC-R as well as the WRAT scores were essentially similar, however, and the seperation was observed to reflect cluster group differences based on overall level of functioning, rather than specific subtype of learning disorder.

These findings from the cluster analysis with the LNNB data would support an interpretation that there were real subgroups within the total sample, but the failure to achieve a significant correspondence between the educational classifications based on WRAT academic performance pattern and the cluster memberships supported the interpretation that the subgroups identified were primarily defined on the basis of similar levels of overall impairment or functioning, rather than on similarities in profile shape. An attempt to cluster the sample using a correlational similarity metric (to obviate the effect of profile elevation) failed to achieve a satisfactory cluster solution; with a chaining pattern of cluster assignments arising rather than identification of discrete subgroups. It is concluded that the high level of intercorrelation between the clinical scale is likely a major factor in these results, such that the tendency for the clinical scales to covary is so strong as to result in differences between individuals being expressed more clearly in the overall LNNB profile elevation than in variations in the profile configuration or pattern.

A similar cluster analysis of the LNNB-A clinical scale data also succeeded in identifying three subgroups within the sample, and on external validation, these subgroups were also observed to display highly significant differences on both intellectual and academic achievement test variables. It is therefore concluded that the LNNB-A clinical scales are also significantly related to educationally relevant factors of intelligence and academic achievement, and that the LNNB-A could also be used to identify adolescent students with deficits in intellectual and/or academic functioning.

The overall conclusions derived from these series of analyses on the LNNB and the derived abbreviated form is that this neuropsychological test battery can be shown to have an adequate level of internal consistency to be considered as a reliable instrument for the adolescent age range. The content and construct validities of the test, as a multidimensional measure of the broad range of cognitive abilities that can be affected by brain damage would, however, be challenged by the findings of this study, as there are a relatively high number of items that display correlations across the scales that exceed the correlations with their assigned scales. The high level of intercorrelation between the scales and the limited number of factors with eigenvalues in excess of unity would also be viewed as indicating that the separation of the items into eleven discrete scales with disparate functional labels is also questionable, and likely represents an artifactual rather than legitimate division. The factor analyses of both forms of the test battery did, however, suggest that there are three interpretable group factors present, with these reflecting a combination of verbal intellectual, academic achievement and perceptual/motor dimensions, which would represent three functional areas of neuropsychological significance.

While the evaluations of the content and construct validity of the LNNB suggest a need for refinement and redefinition of what the test is measuring, the evaluations of concurrent validity with respect to intellectual and educational variables were considerably more positive. With the notable exception of nonverbal intelligence, the appropriate clinical scales of the LNNB were observed to be highly correlated with independent measures of verbal and general intelligence and the academic skills of reading, spelling and arithmetic. The battery as a whole was also demonstrated to be capable of predicting (through multiple correlations) WISC-R and WRAT test variables with a high level of significance. Performance I.Q. remained the only area of relative weakness in this regard for both the LNNB and LNNB-A. Educational classification was also observed to be possible with the LNNB and LNNB-A through the use of linear discriminant analysis to identify cases classified on the basis of academic performance pattern, and to subgroup cases on the basis of the level of general functioning, both neuropsychological as well as educational.

The results of this study would therefore be interpreted by this investigator as raising some serious questions regarding the content and construct validity of the LNNB as a comprehensive battery of neuropsychological functioning, with the observed dimensional structure of the test strongly suggesting that it is too limited in its coverage to represent an adequate measure of all relevant neuropsychological functions. This is especially true with respect to the coverage of the nonverbal functions that are typically mediated by the right cerebral hemisphere. However, despite this short-coming, it is also apparent that the battery, in either its full or the abbreviated form, was highly successful in displaying adequate concurrent validity with respect to verbal intellectual and academic variables, and to be able to discriminate between educationally defined subgroups of cases and to identify naturally occurring subgroups within the sample that had educationally significant differences in both intelligence and academic achievement. Thus, while the breadth of coverage of neuropsychological factors may be criticized and the battery held up to question as to its adequacy to reflect the broad range of cognitive functions that are susceptible to brain damage, it would appear that the test has sufficient coverage of those functions that are particularly relevant to academic functioning to allow for meaningful educational classifications. Whether the LNNB can successfully discriminate between specific subtypes of learning disabilities was not adequately determined in this study, however, due to an insufficient range of sampling among the clinical referral base, such that there was an inadequate number of cases available that had a distinctly different pattern of learning disability subtype (i.e., specific language arts deficit with normal arithmetic ability) to test this capability.

Limitations of the Study

While this study has been successful in demonstrating that the LNNB has adequate reliability and concurrent validity with respect to educationally related variables for the adolescent age range, there are a number of significant limitations that are apparent that would act to restrict the generalizability of these results. A primary limitation is related to the composition of the sample used in this research. Cases studied in the course of this investigation were representative of clinical referrals for neuropsychological evaluation conducted by the writer through a community mental health clinic and through a private practice setting. As such, the cases are clearly not representative of the typical caseload of a practicing school psychologist, and were conversely representative of youths with multi-problem situations. In the vast majority of cases the youths had been referred for neuropsychological evaluation on the basis of a combination of presenting complaints, most typically including aggressiveness, delinquency, poor impulse control, or emotional disturbances in addition to a history of inadequate academic performance. As a result, the sampling of cases was significantly skewed toward displaying a considerably lower than average level of functioning on many variables, as reflected in the mean intellectual and educational attainment levels for this sample.

The composition of the sample was also observed to be relatively unusual with respect to the proportion of different types of educational achievement patterns on the WRAT, as a high percentage of the cases were found to display selective deficiencies in arithmetic achievement with no impairment in language arts achievement or to display a pattern of global deficit in both language arts and arithmetic attainment. Out of the initial sample of 120 subjects, only two cases were found of selective deficiency in language arts with normal arithmetic functioning. One consequence of this abnormal sample composition that there was an inadequate number of cases of specific language arts disability available to include these cases in the linear discriminant analysis, such that the capability of the LNNB to discriminate between such language arts deficits and selective deficits in arithmetic or global learning deficits could not be tested.

Another limitation of the current study is that the multiple regression equations and linear discriminant functions have not been subjected to cross validation. As both of these procedures operate to seek the maximal solution based on the data at hand, it would be expected that considerable shrinkage will occur upon cross validation in both the multiple correlations as well as hit rates in discriminant classifications. Restrictions in sample size as well as the recognition that the current sample was not representative of the general adolescent population mitigated against attempts to cross validate these solutions by division of the sample for independent analyses. It is felt that the current results are best viewed as representative of an initial exploration of the use of the LNNB in the adolescent age range, and efforts to prepare a more comprehensive evaluation of this battery (including cross validational procedures) should be directed toward a sample of cases compiled to more accurately represent the populations of normal and learning disabled youths most likely to be seen by a school psychologist.

A further limitation of the current research as a validational investigation of the LNNB in the adolescent age range is the focus upon cases that were included primarily on the basis of combined behavioural/emotional and educational problems, and a lack of sampling of cases of specific neurological disorders. As the LNNB was designed and initially validated as a measure of neuropsychological impairment for brain damaged patients, the limited dimensional structure displayed in this study may represent a lack of inclusion of cases with the types of deficits these tests were designed to detect. As such, the obtained results of this study should only be interpreted as raising questions about the content and construct validity of the LNNB, rather than clearly refuting them. Further explorations using carefully selected groups of adolescents with definitive neurological diagnoses would be required to determine the overall effectiveness of the battery for neuropsychological assessment in such clinical cases.

A further area of limitation in the in the current study would be with regards to the adequacy of the academic achievement test selected as an external criteria for the concurrent validity investigation. While the WRAT is a well recognized and highly utilized measure of educational achievement, it is also clearly limited in the breadth and comprehensiveness of its assessment. It represents a very brief screening assessment of a limited selection of academic skills, and cannot be considered as a thorough evaluation of educational performance. For example, the reading subtest of the WRAT represents a very isolated subset of reading skills (oral reading of words in isolation) and no assessment is made of the individual's ability to comprehend what is read or to deal with words in context. The spelling test is limited to spelling of words from dictation, and there is no distinction made between words that display phonetically consistent and phonetically irregular spellings, nor of the abilities to use phonic attack versus rote visual memory strategies to spell words. The arithmetic subtest is composed of a series of written calculation problems ranging from simple addition and subtraction items to complex items of algebra and logarithms, but it is time limited and dependent on speed as well as accuracy. With these restrictions in mind, it is apparent that the ability to make specific predictions regarding educational classifications would also be necessarily limited by the restricted nature of the educational criterion available. A more detailed and thorough educational functioning assessment would be desirable in such investigations, to determine more accurately the capabilities of the neuropsychological battery to identify meaningful patterns or syndromes of educational functioning.

Implications for Future Research

Based upon the findings of this study, and a consideration of the limitations of the current investigations, it is recommended that future research be directed toward obtaining a more representative sample of adolescent cases upon which a more definitive evaluation of the content and construct validities can be conducted, and in which the calculation of multiple regression equations and linear discriminant functions can be subjected to cross validation.

Specific efforts to obtaining a sizable sample of normal adolescents without neurological, educational, or emotional/behavioural disturbances would be useful to allow for the development of specific adolescent norms for the LNNB. This would be especially useful in this age range as the clinical scale T-Score transformations are currently based upon the standardization sample of adult cases which had an average age of 42.0 years (standard deviation of 14.8 years) and an average educational level of 12.2 years of education (standard deviation of 2.9 years). While the calculation of the critical level is advocated as a means of compensating for the effects of age and education level on the LNNB, the published procedures for this calculation allows for additional compensation for age only down to age 25, and all cases under that age are (for the purpose of the critical level calculation) assumed to be equal to a 25 years old level. As significant changes in ability and knowledge level would be expected to occur within the age range of 13 to 25 years (particularly on the items sampled on the Intellectual Processes, Writing, Reading and Arithmetic scales), it would appear to be critical to establish some more appropriate adolescent age norms for this test battery if its use in these lower age levels is to be recommended.

A further area of needed research with this battery would be the systematic collection of adolescent cases of well documented and clearly defined neurological damage, and the replication of validational studies of discrimination between such neurological cases and non-neurological cases in this age range, as well as explorations of the ability of the battery to lateralize and localize neurological damage in these lower age ranges.

With respect to further validation of the LNNB for educational neuropsychological evaluations, collection of appropriate samples of learning disabled youths without complicating factors of emotional and behavioural disturbances should be sought, drawing such cases from a school setting, rather than from the inherently biased population of referrals for clinical neuropsychological assessment.

In addition to the selection of a more representative sample of youths for such educational research, further efforts should be directed toward the design of an expanded and more detailed external educational criterion, such that further research with the LNNB (and/or the LNNB-A) would allow for a more detailed description of the specific subtypes of learning disorder present in the sample, as a more thorough and realistic evaluation of the neuropsychological battery's capacity for classification and subgrouping of the sample.

Finally, based upon the observations of a relatively weak representation of nonverbal intellectual/cognitive processes on the LNNB and the reports of other researchers that failed to achieve adequate identification of neurological patient groups with right hemispheric brain damage, a revision of the LNNB should be explored, with the express intent of incorporating into the test battery additional measures that would be designed to be sensitive to such nonverbal cognitive functions of the right hemisphere. In this connexion, the current identification of an abbreviation of the LNNB that has equal reliability, correlates highly with the original battery and which has essentially identical factor structure would suggest that the published LNNB can be shortened without significant loss of coverage of neuropsychological functions. Further research based upon the use of the abbreviated LNNB-A in conjunction with additional nonverbal or "right hemispheric" tasks may therefore result in the improvement of the functional coverage of this approach, while still maintaining the practical advantages of the current LNNB battery with respect to the length of administration time.

References

Adams, K.M. Linear discriminant analysis in clinical neuropsychological research. Journal of Clinical Neuropsychology, 1979, 1(3), 259-272.

Adams, K.M. In search of Luria's battery: A false start. Journal of Consulting and Clinical Psychology, 1980a, 48(4), 511-516.

Adams, K.M. An end of innocence for behavioral neurology? Adams replies. Journal of Consulting and Clinical Psychology, 1980b, 48(4), 522-524.

Ajax, E.T. Acquired dyslexia. Archives of Neurology, 1964, 11, 66-72.

Anastasi, A. Psychological Testing. New York: Macmillan, 1968.

Anastasi, A. Psychological testing (3rd. ed.). New York: Macmillan, 1982.

Bach, P.J., Harowski, K., Kirby, K., Peterson, P., & Schulein, M. The interrater reliability of the Luria - Nebraska Neuropsychological Battery. Clinical Neuropsychology, 1981, III(3), 19-21.

Bakker, D.J. Hemispheric specialization and stages in the learning to read process. Bulletin of the Orton Society, 1973, 23, 15-27.

Bakker, D.J. Perceptual asymmetries and reading strategies: Two dyslexias? Bulletin of the Orton Society, 1979, 29, 84-100.

Bakker, D.J. Hemisphere-specific dyslexia models. In R.N. Malatesha & L.C. Hartlage (eds.), Neuropsychology and cognition (Vol. 1). The Hague: Martinus Nijhoff, 1982.

Bakker, D.J. Hemispheric specialization and specific reading retardation. In M. Rutter (Ed.), Developmental neuropsychiatry. New York: Guilford, 1983.

Bakker, D.J., Licht, R., Kok, A., & Bouma, A. Cortical responses to word reading by right- and left-eared normal and reading-disabled children. Journal of Clinical Neuropsychology, 1980, 2, 1-12.

Bartko, J.J., Strauss, J.S., & Carpenter, W.T. An evaluation of taxometric techniques for psychiatric data. Classification Society Bulletin, 1971, 2, 2-28.

Bateman, B. An educator's view of a diagnostic approach to learning disorders. In J. Hellmuth (Ed.), Learning disorders, Vol. 1. Seattle: Special Child Publications, 1965.

Benson, D.F., Segara, J., & Albert, M.L. Visual agnosia - prosopagnosia. Archives of Neurology, 1974, 30, 307-310.

Benton, A.L. The function of the "Gerstmann syndrome." Journal of Neurology, Neurosurgery and Psychiatry, 1961, 24, 176-181.

Berg, R.A., Bolter, J.F., Ch'ien, L.T., Williams, S.J., Lancaster, W., & Gummins, J. Comparative diagnostic accuracy of the Halstead - Reitan and Luria - Nebraska Neuropsychological Adult and Children's Batteries. International Journal of Clinical Neuropsychology, 1984, VI(3), 200-204.

Berg, R.A., & Golden, C.J. Identification of neuropsychological deficits in epilepsy using the Luria - Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1981, 49(5), 745-747.

Boder, E. A neuropediatric approach to school behavioral and learning disorders: Diagnosis and management. In J. Hellmuth (Ed.), Learning disorders (Vol. 2). Seattle: Special Child Publications, 1966.

Boder, E. Developmental dyslexia: A diagnostic screening procedure based on three characteristic patterns of reading and spelling. In M. Douglas (Ed.), Claremont Reading Conference 32nd Yearbook, Claremont, California: Claremont University Centre, 1968.

Boder, E. Developmental dyslexia: Prevailing diagnostic concepts and a new approach. Bulletin of the Orton Society, 1973, 23, 106-118.

Boder, E. & Jarrico, S. The Boder Test of reading - spelling patterns: A diagnostic screening test for subtypes of reading disability. New York: Grune & Stratton, 1982.

Botez, M.I., & Wertheim, N. Expressive aphasia and amusia following right frontal lesion in a right handed man. Brain, 1959, 82, 186-202.

Bloom, B.S. Human characteristics and school learning. New York: McGraw - Hill Book Co., 1976.

Brentzen, F. Sex ratios in learning and behavior disorders. American Journal of Orthopsychiatry, 1963, 33, 81-92.

Butters, N., & Brody, B.A. The role of the left parietal lobe in the mediation of intra- and cross-modal associations. Cortex, 1969, 4, 328-331.

Carmon, A. Impaired memory for duration of time intervals in patients with cerebral damages. European Neurology, 1972, 7, 339-348.

Chelune, G.J. A reexamination of the relationship between the Luria - Nebraska and Halstead - Reitan batteries: Overlap with the WAIS. Journal of Consulting and Clinical Psychology, 1982, 50, 578-580.

Christensen, A. Luria's neuropsychological investigation. New York: Spectrum, 1975.

Comrey, A.C. Common methodological problems in factor analytic studies. Journal of Consulting and Clinical Psychology, 1978, 46, 648-659.

Corkin, S. Tactually guided maze learning in man: Effects of unilateral cortical excisions and bilateral hippocampal lesions. Neuropsychologia, 1974, 12, 347-352.

Cormack, R.M. A review of classification. Journal of the Royal Statistical Society (Series A), 1971, 134, 321-367.

Crosson, B., & Warren, R.L. Use of the Luria - Nebraska Neuropsychological Battery in aphasia: A conceptual critique. Journal of Consulting and Clinical Psychology, 1982, 50, 22-31.

Dalby, J.T., & Gibson, D. Functional cerebral lateralization in subtypes of disabled readers. Brain and Language, 1981, 14, 34-48.

Delis, D.C., & Kaplan, E. The assessment of aphasia with the Luria - Nebraska Neuropsychological Battery: A case critique. Journal of Consulting and Clinical Psychology, 1982, 50, 32-39.

Delis, D.C., & Kaplan, E. Hazards of a standardized neuropsychological test with low content validity: Comment on the Luria-Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1983, 51(3), 396-398.

Denckla, M.B. Minimal brain dysfunction and dyslexia: Beyonf diagnosis by exclusion. In M.E. Blaw, J. Rapin, & M. Kinsbourne (Eds.), Child neurology. New York: Spectrum, 1977.

Denckla, M.B. Childhood learning disabilities. In K.M. Heilman & E. Valenstein (Eds.), Clinical neuropsychology. New York: Oxford University Press, 1979.

Denckla, M.B. & Bowen, F.B. Dyslexia after left occipitotemporal lobectomy: A case report. Cortex, 1973, 9, 321-328.

Doehring, D.G., & Hoshko, I.M. Classification of reading problems by the Q technique of factor analysis. Cortex, 1977, 13, 281-294.

Doehring, D.G., Hoshko, I.M., & Bryans, B.N. Statistical classification of children with reading problems. Journal of Clinical Neuropsychology, 1979, 1, 5-16.

Doehring, D.G., Trites, R.L., Patel, P.G., & Fiedorowizc, C.A.M. Reading disabilities: The interaction of reading, language and neurological deficits. New York: Academic Press, 1981.

Dunn, L.M., & Markwardt, F.C. Peabody Individual Achievement Test. Circle Pines, MN.: American Guidance Service, 1970.

Ettinger, G., Warrington, E., & Zangwell, O.L. A further study of visual spatial agnosia. Brain, 1957, 80, 335-361.

Everitt, B.S. Cluster analysis (2nd ed.). London: Heineman Educational Books, 1980.

Fisk, J.L., & Rouke, B.P. Identification of subtypes of learning disabled children at three age levels: A multivariate approach. Journal of Clinical Neuropsychology, 1979, 1, 289-310.

Fletcher, J.M., & Satz, P. Cluster analysis and the search for learning disability subtypes. In B.P. Rourke (Ed.), Neuropsychology of learning disabilities: Essentials of subtype analysis. New York: The Guilford Press, 1985.

Fried, I. Cerebral dominance and subtypes of developmental dyslexia. Bulletin of the Orton Society, 1979, 29, 101-112.

Fried, I., Tanguay, P.E., Boder, E., Doubleday, C., & Greensite, M. Developmental dyslexia: Electrophysiological evidence of clinical subgroups. Brain and Language, 1981, 12, 14-22.

Gaddes, W.H. Prevalence estimates and the need for definition of learning disabilities. In R. Knights and D. Bakker (Eds.), The neuropsychology of learning disorders. Baltimore: University Park Press, 1976.

Gallagher, J. Children with developmental imbalances: A psycho-educational definition. In W. Cruickshank (Ed.), The teacher of brain-injured children: A discussion of the bases of competency. Syracuse, N.Y.: Syracuse University Press, 1966.

Gazzaniga, M.S., & Hillyard, S.A. Language and speech capacity in the right hemisphere. Neuropsychologia, 1971, 9, 273-279.

Gillen, R.W., Ginn, C., Strider, M.A., Kreuch, T.J., & Golden, C.J. The Luria - Nebraska Neuropsychological Battery and the Peabody Individual Achievement Test: A correlational analysis. International Journal of Neuroscience, 1983, 21, 51-62.

Golden, C.J. Validity of the Halstead - Reitan Neuropsychological Battery in a mixed psychiatric and brain-injured population. Journal of Consulting and Clinical Psychology, 1977, 45, 1043-1051.

Golden, C.J. Diagnosis and rehabilitation in clinical neuropsychology. Springfield, IL.: Charles C. Thomas, 1978.

Golden, C.J. In reply to Adam's "In search of Luria's battery: A false start", Journal of Consulting and Clinical Psychology, 1980, 48(4), 517-521.

Golden, C.J. A standardized version of Luria's neuropsychological tests: A quantitative and qualitative approach to neuropsychological evaluation. In S.B. Filskov & T.J. Boll (Eds.), Handbook of clinical neuropsychology, New York: Wiley, 1981.

Golden, C.J., Ariel, R.N., Moses, J.A., Wilkening, G.N., McKay, S.E., & MacInnes, W.D. Analytic techniques in the interpretation of the Luria - Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1982, 50, 40-48.

Golden, C.J., Berg, R.A., & Graber, B. Test-retest reliability of the Luria - Nebraska Neuropsychological Battery in stable, chronically impaired patients. Journal of Consulting and Clinical Psychology, 1982, 50(3), 452-454.

Golden, C.J., Fross, K.H., & Graber, B. Split-half reliability and item-scale consistency of the Luria - Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1981, 49(2), 304-305.

Golden, C.J., Hammeke, T.A., & Purisch, A.D. Diagnostic validity of a standardized neuropsychological battery derived from Luria's neuropsychological tests. Journal of Consulting and Clinical Psychology, 1978, 46, 1258-1265.

Golden, C.J., Hammeke, T.A., & Purisch, A.D. The Luria -- Nebraska Neuropsychological Battery: Manual. Los Angeles: Western Psychological Services, 1980.

Golden, C.J., Kane, R., Sweet, J., Moses, J.A., Cardellino, J.P., Templeton, R., Vincente, P., & Graber, B. Relationship of the Halstead - Reitan Neuropsychological Battery to the Luria - Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1981, 49, 410-417.

Golden, C.J., Moses, J.A., Jr., Fishburne, F.J., Engum, E., Lewis, G.P., Wisniewski, A., Conley, F.K., Berg, R.A., & Grabber, B. Cross validation of the Luria Nebraska Neuropsychological Battery for the presence, lateralization, and localization of brain damage. Journal of Consulting and Clinical Psychology, 1981, 49, 491-507.

Golden, C.J., Purisch, A.D., & Hammeke, T.A. Luria-Nebraska Neuropsychological Battery: Forms I and II Manual. Los Angeles, Ca.: Western Psychological Services, 1985.

Golden, C.J., Purisch, A.D., Sweet, J., Graber, B., Osmon, D., & Hammeke, T.A. Factor analysis of the Luria-Nebraska Neuropsychological Battery II: Visual, Receptive, Expressive and Reading scales. International Journal of Neuroscience, 1980, 11, 227-236.

Goldstein, G. The use of clinical neuropsychological methods in the lateralization of brain lesions. In J. Dimond & J.G. Beaumont, (Eds.), Hemisphere function in the brain. New York: Wiley, 1974.

Goldstein, G., & Shelly, C. Field dependence and cognitive, perceptual and motor skills in alcoholics: A factor analytic study. Quarterly Journal of Studies on Alcohol, 1971, 32, 29-40.

Goldstein, G., & Shelly, C. Statistical and normative studies of the Halstead Neuropsychological Test Battery relevant to a neuropsychiatric hospital setting. Perceptual and Motor Skills, 1972, 34, 603-620.

Goldstein, G., & Shelly, C. Relationships between language skills as assessed by the Halstead - Reitan Battery and the Luria - Nebraska language - related factor scales in a nonaphasic patient population. Journal of Clinical Neuropsychology, 1984, 6(2), 143-156.

Hale, R.L., & Foltz, S.G. Prediction of academic achievement in handicapped adolescents using a modified form of the Luria - Nebraska Pathognomonic Scale and WISC-R Full Scale IQ. Clinical Neuropsychology, 1982, IV(3), 99-103.

Halstead, W.C. Brain and intelligence. Chicago: University of Chicago Press, 1947.

Hammeke, T.A., Golden, C.J., & Purisch, A.D. A standardized, short, and comprehensive neuropsychological test battery based on the Luria Neuropsychological Evaluation. International Journal of Neuroscience, 1978, 8, 135-141.

Hebb, D.O. Intelligence, brain function and the theory of mind. Brain, 1959, 82, 260-275.

Hecean, H., & Angelergues, R. Agnosia for faces (prosopagnosia). Archives of Neurology, 1962, 7, 92-100.

Hill, T.D., Reddon, J.R., & Jackson, D.N. The factor structure of the Wechsler Scales: A brief review. Clinical Psychology Review, 1985, 5, 287-306.

Hutchinson, G.L. The Luria-Nebraska Neuropsychological Battery Controversy: A reply to Spiers. Journal of Consulting and Clinical Psychology, 1984, 52(4), 539-545.

Hynd, G.W., & Obrzut, J.E. Neuropsychological assessment and the school - age child: Issues and procedures. New York: Grune [illegible]

Jastak, J.F., & Jastak, S. The Wide Range Achievement Test: Manual of Instructions. Wilmington, Dela.: Jastak Associates, Inc., 1978.

Johnson, D., & Myklebust, H. Learning disabilities: Educational principles and practices. New York: Grune & Stratton, 1967.

Johnston, C. Learning disability subtypes: A cross-validation. Doctoral dissertation, University of Florida, 1984.

Joschko, M., & Rourke, B.P. Neuropsychological subtypes of learning- disabled children who exhibit the ACID pattern on the WISC. In Rourke, B.P. (Ed.) Neuropsychology of learning disabilities: Essentials of subtype analysis. New York: The Guilford Press, 1985.

Joynt, R.J., & Goldstein, M.N. Minor cerebral hemisphere. Advances in Neurology, 1975, 7, 147-159.

Kaiser, H.F. The application of electronic computers to factor analysis. Educational and Psychological Measurement, 1960, 20, 141-151.

Kane, R.L., Sweet, J.J., Golden, C.J., Parsons, O.A., & Moses, J.A. Comparative diagnostic accuracy of the Halstead - Reitan and standardized Luria - Nebraska Neuropsychological Batteries in a mixed psychiatric and brain-damaged population. Journal of Consulting and Clinical Psychology, 1981, 49, 484-485.

Kear-Colwell, J.J., & Heller, M. A normative study of the Wechsler Memory Scale. Journal of Clinical Psychology, 1978, 34, 437-442.

Kimura, D. Some effects of temporal lobe damage on auditory perception. Canadian Journal of Psychology, 1961, 15, 156-165.

Kimura, D. Right temporal lobe damage. Archives of Neurology, 1963, 8, 264-271.

Kimura, D., & Archibald, Y. Motor functions of the left hemisphere. Brain, 1974, 97, 337-350.

Kirk, S.A., & Bateman, B. Diagnosis and remediation of learning disabilities. Exceptional Children, 1962, 29, 73-78.

Lackner, J.R., & Teuber, H.L. Alterations in auditory fusion thresholds after cerebral injury in man. Neuropsychologia, 1973, 11, 409-413.

Lansdell, H. Relation of extent of temporal removals to closure and visuomotor factors. Perceptual Motor Skills, 1970, 31, 491-498.

Lerner, J.W. Children with learning disabilities: Theories, diagnosis, and teaching strategies. New York: Houghton Mifflin Co., 1971.

Lewis, G.P., Golden, C.J., Moses, J.A., Osmon, D.C., Purisch, A.D., & Hammeke, T.A. Localization of cerebral dysfunction with a standardized version of Luria's neuropsychological battery. Journal of Consulting and Clinical Psychology, 1979, 47, 1003-1019.

Luria, A.R. Higher cortical functions in man. New York: Basic Books, 1966.

Luria, A.R. The working brain. New York: Basic Books, 1973.

Luria, A.R., Pribram, K., & Homokaya, E.D. An experimental analysis of the behavioral disturbance produced by a left frontal arachnoidal endothelioma (meningioma). Neuropsychologia, 1964, 2, 257-264.

Malloy, P.F., & Webster, J.S. Detecting mild brain impairment using the Luria - Nebraska Neuropsychological Battery. Journal of Consulting and Clinical Psychology, 1981, 49(5), 768-770.

Maruish, M.E., Sawicki, R.F., Franzon, M.D., & Golden, C.J. Alpha coefficient reliabilities for the Luria - Nebraska Neuropsychological Battery Summary and Localization scales by diagnostic category. The International Journal of Clinical Neuropsychology, 1985, VII(1), 10-12.

Mathura, C.B. The vulnerability of the male CNS to early trauma: Implications for clinical neuropsychology. Clinical Neuropsychology, 1979, I(4), 34-35.

Mattis, S. Dyslexia syndromes: A working hypothesis that works. In A.L. Benton & D. Pearl (Eds.), Dyslexia: An Appraisal of current knowledge. New York: Oxford University Press, 1978.

Mattis, S., French, J.H., & Rapin, I. Dyslexia in children and young adults: Three independent neuropsychological syndromes. Developmental Medicine and Child Neurology, 1975, 17, 150-163.

McCue, M., Shelly, C., Goldstein, G., & Katz-Garris, L. Neuropsychological aspects of learning disability in young adults. International Journal of Clinical Neuropsychology, 1984, VI(4), 229-233.

McFie, J. Assessment of organic intellectual impairment. New York: Academic Press, 1975.

McKay, S., & Golden, C.J. Empirical derivation of experimental scales for localizing brain lesions using the Luria - Nebraska Neuropsychological Battery. Clinical Neuropsychology, 1979, I(2), 19-23.

McKay, S.E., & Golden, C.J. Re-examination of the factor structure of the receptive scale of the Luria - Nebraska Neuropsychological Battery. International Journal of Neuroscience, 1981, 14, 183-188.

McKay, S.E., Golden, C.J., Moses, J.A., Fishburne, F., & Wisniewski, A. Correlation of the Luria - Nebraska Neuropsychological Battery with the WAIS. Journal of Consulting and Clinical Psychology, 1981, 49(6), 940-946.

McKay, S., & Ramsey, R. Correlation of the Wechsler Memory Scale and the Luria - Nebraska Memory Scale. Clinical Neuropsychology, 1983, V(4), 168-170.

Milner, B. Psychological defects produced by temporal lobe excision. In H.C. Solomon, S. Cobb, & W. Renfield (Eds.), The brain and human behavior. Baltimore: Williams & Wilkins, 1958.

Milner, B. Interhemispheric difference in the localization of psychological processes in man. Brain Medical Bulletin, 1971, 27, 272-279.

Morris, R., Blashfield, R., & Satz, P. Neuropsychology and cluster analysis: Potentials and problems. Journal of Clinical Neuropsychology, 1981, 3(1), 79-99.

Moses, J.A., & Golden, C.J. Cross validation of the discriminative effectiveness of the standardized Luria - Nebraska Neuropsychological Battery. International Journal of Neuroscience, 1979, 9, 149-155.

Moses, J.A., & Golden, C.J. Discrimination between schizophrenic and brain-damaged patients with the Luria - Nebraska Neuropsychological Test Battery. International Journal of Neuroscience, 1980, 10, 121-128.

Moses, J.A., Golden, C.J., Ariel, R., & Gustavson, J.L. Interpretation of the Luria - Nebraska Neuropsychological Battery, Vol. 1. New York: Grune & Stratton, 1983.

Moses, J.A., & Schefft, B.K. Interrater reliability analyses of the Luria - Nebraska Neuropsychological Battery. International Journal of Clinical Neuropsychology, 1985, VII(1), 31-38.

Myklebust, H.R. Toward a science of dyslexiology. In H.R. Myklebust (Ed.), Progress in learning disabilities (Vol. 4). New York: Grune & Stratton, 1978.

Myklebust, H.R., & Boshes, B. Final report: Minimal brain damage in children. Washington, D. C.: Department of Health, Education and Welfare, 1969.

Nie, N., Hull, C., Jenkins, J., Steinbrenner, K., & Bent, D. SPSS: Statistical package for the social sciences. New York: McGraw - Hill, 1975.

Osmon, D.C., Golden, C.J., Purisch, A.D., Hammeke, T.A., & Blume, H.G. The use of a standardized battery of Luria's tests in the diagnosis of lateralized cerebral dysfunction. International Journal of Neuroscience, 1979, 9, 1-9.

Osmon, D.C., Sweet, J., & Golden, C.J. Neuropsychological implications of rhythm and aphasia deficits in unilateral left hemisphere injuries. International Journal of Neuroscience, 1978, 8, 499-524.

Parolini, R. Reading, spelling, and arithmetic disabilities: A neuropsychological investigation using Luria's methods. (Doctoral dissertation, University of Nebraska, Lincoln, 1982). Dissertation Abstracts International, 1983, 43, 1996B.

Petrauskas, R.J., & Rourke, B.P. Identification of subtypes of retarded readers: A neuropsychological, multivariate approach. Journal of Clinical Neuropsychology, 1979, 1, 17-38.

Picker, W.R., & Schlottmann, R.S. An investigation of the Intellectual Processes scale of the Luria - Nebraska Neuropsychological Battery. Clinical Neuropsychology, 1982, IV(3), 120-124.

Piercy, M., & Smyth, V.O.G. Right hemisphere dominance for certain non-verbal intellectual skills. Brain, 1962, 85, 775-790.

Pirozzolo, F.J., & Rayner, K. Cerebral organization and reading disability. Neuropsychologia, 1979, 17, 485-491.

Plaisted, J.R., & Golden, C.J. Test-retest reliability of the Clinical, Factor, and Localization scales of the Luria - Nebraska Neuropsychological Battery. International Journal of Neuroscience, 1982, 17, 163-167.

Press, S.J. Applied Multivariate Analysis: Using Bayesian aand Frequentist methods of inference.. Malabar, Florida: Robert E. Krieger Publishing Company, 1982.

Prifitera, A., & Ryan, J.J. Validity of the Luria - Nebraska Intellectual Process scale as a measure of adult intelligence. Journal of Consulting and Clinical Psychology, 1981, 49(5), 755-756.

Purisch, A.D., Golden, C.J., & Hammeke, T. Discrimination of schizophrenic and brain - injured patients by a standardized version of Luria's neuropsychological tests. Journal of Consulting and Clinical Psychology, 1978, 48, 1266-1273.

Reitan, R.M. Psychological deficits resulting from cerebral lesions in man. In J.M. Warren and K.A. Akert (Eds.), The frontal granular cortex and behavior. New York: McGraw-Hill, 1964.

Reitan, R.M. Neuropsychology: The vulgarization Luria always wanted. Contemporary Psychology, 1976, 21, 737.

Reitan, R.M., & Boll, T.J. Neuropsychological correlates of minimal brain dysfunction. Annals of the New York Academy of Sciences, 1973, 205, 65-89.

Reitan, R.M. & Davison, L.A. Clinical neuropsychology: Current status and applications. New York: Hemisphere, 1974.

Reitan, R.M., & Wolfson, D. The Halstead - Reitan Neuropsycholgical Test Battery: Theory and clinical interpretation. Tucson, Arizona: Neuropsychology Press, 1985.

Rourke, B.P. Neuropsychological research in reading retardation: A review. In A.L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge. New York: Oxford University Press, 1978.(a)

Rourke, B.P. Reading, spelling, arithmetic disabilities: A neuropsychologic perspective. In H.R. Myklebust (Ed.), Progress in learning disabilities (Vol. 4). New York: Grune & Stratton, 1978.(b)

Rourke, B.P. Neuropsychological assessment of children with learning disabilities. In S.B. Filskov & T.J. Boll (Eds.), Handbook of clinical neuropsychology. New York: Wiley - Interscience, 1981.

Rourke, B.P. (Ed.) Neuropsychology of learning disabilities: Essentials of subtype analysis. New York: Guilford Press, 1985.

Rourke, B.P., & Finlayson, M.A.J. Neuropsychological significance of variations in patterns of performance on the Trail Making Test for older children with learning disabilities. Journal of Abnormal Psychology, 1975, 84(4), 412-421.

Rourke, B.P., & Finlayson, M.A.J. Neuropsychological significance of variations in patterns of academic performance: Verbal and visual - spatial abilities. Journal of Abnormal Child Psychology, 1978, 6, 121-133.

Rourke, B.P., & Fisk, J.L. Socio - emotional disturbances of learning disabled children: The role of central processing deficits. Bulletin of the Orton Society, 1981, 31, 77-88.

Rourke, B.P., & Orr, R.R. Prediction of the reading and spelling performances of normal and retarded readers: A 4 - year follow-up. Journal of Abnormal Child Psychology, 1977, 5, 9-20.

Rourke, B.P., & Strang, J.D. Subtypes of reading and arithmetic disabilities: A neuropsychological analysis. In M. Rutter (Ed.), Developmental neuropsychiatry. New York: Guilford Press, 1983.

Rourke, B.P., Yanni, D.W., MacDonald, G.W., & Young, G.E. Neuropsychological significance of lateralized deficits on the Grooved Pegboard Test for older children with learning disabilities. Journal of Consulting and Clinical Psychology, 1973, 41(1), 128-134.

Rutter, M. Prevalence and types of dyslexia. In A.L. Benton & D. Pearl (Eds.), Dyslexia: An appraisal of current knowledge. New York: Oxford University Press, 1978.

Ryan, J.J., & Prifitera, A. Concurrent validity of the Luria - Nebraska Memory scale. Journal of Clinical Psychology, 1982, 38, 378-379.

Sattler, J.M. Assessment of children's intelligence. Philadelphia: W.B. Saunders Co., 1974.

Satz, P., Friel, J., & Rudegair, F. Some predictive antecedants of specific reading disability: A two-, three-, and four-year follow-up. In J.T. Guthrie (Ed.), Aspects of reading acquisition. Baltimore: John Hopkins University Press, 1976.

Satz, P., & Morris, R. Learning disability subtypes: A review. In F.J. Pirozzolo & M.C. Wittrock (Eds.), Neuropsychological and cognitive processes in reading. New York: Academic Press, 1981.

Shelly, C., & Goldstein, G. Intelligence, achievement, and the Luria - Nebraska Battery in a neuropsychiatric population: A factor analytic study. Clinical Neuropsychology, 1982, IV(4), 164-169.

Shelly, C., & Goldstein, G. Discrimination of chronic schizophrenia and brain damage with the Luria - Nebraska Battery: A partially successful replication. Clinical Neuropsychology, 1983, V(2), 82-85.

Spiers, P.A. Have they come to praise Luria or to bury him? The Luria-Nebraska battery controversy. Journal of Consulting and Clinical Psychology, 1981, 49, 331-341.

Spiers, P.A. The Luria - Nebraska Neuropsychological Battery revisited: A theory in practice or just practicing? Journal of Consulting and Clinical Psychology, 1982, 50, 301-306.

Stambrook, M. The Luria - Nebraska Neuropsychological Battery: A promise that may be partly fulfilled. Journal of Clinical Neuropsychology, 1983, 5(3), 247-269.

Stom, D.A., Fischer, W.E., & Dean, R.S. Factor structure of the HRNB by gender with a learning disabled population. Paper presented at the Seventh Annual Meeting of the National Academy of Neuropsychologists, Chicago, Ill., October 28, 1987.

Swiercinsky, D. Manual for the adult neuropsychological evaluation. Springfield, Ill.: C. C. Thomas, 1978.

Taylor, H.G., Satz, P., & Friel, J. Developmental dyslexia in relation to other childhood reading disorders: Significance and utility. Reading Research Quarterly, 1979, 15, 84-101.

Teuber, H.L. Space perception and its disturbance after brain injury in man. Neuropsychologia, 1963, 2, 209-214.

Townes, B.D., Trupin, E.W., Martin, D.C., & Goldstien, D. Neuropsychological correlates of academic success among elementary school children. Journal of Consulting and Clinical Psychology, 1980, 48(6), 675-684.

Tramontana, M.G., Sherrets, S.D., & Wolf, B.A. Comparability of the Luria – Nebraska and Halstead – Reitan Neuropsychological Batteries for older children. Clinical Neuropsychology, 1983, V(4), 186-190.

Trites, R.L., & Fiedorowicz, C. Follow-up study of children with specific (or primary) reading disability. In R.M. Knights & D.J. Bakker (Eds.), The neur[illegible]f learning disorders: Theoretical approa[illegible]more, Md.: University Park Press, 1976.

van der Vlugt, H., & Satz, P. Subgroups and subtypes of learning disabled and normal children: A cross-cultural replication. In B.P. Rourke (Ed.), Neuropsychology of learning disabilities: Essentials of subtype analysis. New York: The Guilford Press, 1985.

Walsh, K.W. Neuropsychology. New York: Churchill Livingston/Longman, 1978.

Warrington, E.K., Logue, V., & Pratt, R.T.C. The anatomical localization of selective impairment of auditory short term memory. Neuropsychologia, 1971, 9, 377-381.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children - Revised. New York: Psychological Corporation, 1974.

Wechsler, D., & Stone, C.P. Wechsler Memory Scale manual. New York: The Psychological Corporation, 1945.

Weisenburg, T.H., & McBride, K.E. Aphasia: A clinical and psychological study. New York: Commonwealth Fund, 1935.

Wheeler, L., & Reitan, R.M. Discriminant functions applied to the problem of predicting cerebral damage from behavioral tests: A cross validation study. Perceptual and Motor Skills, 1963, 16, 681-701.

Wishart, D. CLUSTAN user manual (3rd ed.). London: University of London, 1975.

Wyke, M. The effects of brain lesions on the performance of bilateral arm movements. Neuropsychologia, 1971, 9, 33-40.

Zangwill, O.L. Psychological deficit associated with frontal lobe lesions. International Journal of Neurology, 1966, 5, 395-402.