The Validity of Testing Aphasic Patients with The Luria-Nebraska Battery: Preliminary Findings

Reg L. Warren
Braintree Hospital, Braintree, Massachusetts

Bruce A. Crosson
Veterans Administration Medical Center, Leavenworth, Kansas

Susan Edwards
Veterans Administration Hospital, Columbia, Missouri

INTRODUCTION

Recently, a standardized version of Luria's neuropsychological techniques was published as the Luria-Nebraska Neuropsychological Battery (Golden, Hammeke and Purisch, 1980). Claiming to incorporate many aspects of Luria's "functional systems approach," the battery is said to be useful for breaking functional deficits of brain injured patients into their component parts. While the Luria-Nebraska has gained wide popularity among psychologists, it has been recently criticized by aphasiologists (Holland, 1982) and a significant portion of the neuropsychological community (Lezak, 1983; Delis and Kaplan, 1982; and Spiers, 1981). The problem is that many of the items of the Luria-Nebraska do not measure what they were designed to measure when given to aphasic patients. Previously we have pointed out that the Luria-Nebraska consistently requires the use of language to measure non-language functions (Crosson and Warren, 1982).

For example, tactile identification of common objects is assessed by blindfolding a patient, placing an object in his hand, and asking the patient to name the item. Under these conditions it is impossible to determine whether an incorrect response from an aphasic patient is due to tactile misperception, anomia, or both. Similar problems exist for tasks which assess visual-perception--for many items there is no way to separate aphasic difficulties from visual agnosia. Of the 34 items which assess intellectual functions, 30 require processing of verbally formulated material or verbal output or both. A Reading scale does not measure reading comprehension; it primarily measures oral reading ability. A patient with apraxia of speech could comprehend everything he/she was reading but fail the entire scale because of an inability to articulate. Of the 33 items measuring comprehension of speech, over half require either written or verbal expression.

The confounding influences of the Luria-Nebraska format recently have been illustrated in a case presentation by Delis and Kaplan (1982), who described a patient with a CT-confirmed, left posterior lesion, presenting no hemiparesis and a mild, fluent aphasia as determined on the Boston Diagnostic Aphasia Examination. The Luria-Nebraska clearly profiled the patient as having a right hemisphere lesion anterior to the Rolandic fissure with expressive aphasia. Golden et al., (1981) claim that the Luria-Nebraska is accurate in localizing unilateral involvement 78% of the time. Certainly, the interpretive skills of most clinicians exceed the actuarial process of adding up test scores and literally interpreting test profiles. However, it is essential that the format of a neuropsychological test allow the clinician to make reasonably direct inferences about the status and location of affected cognitive processes through observation of a patient's behavior.

PURPOSE OF THE STUDY

The Luria-Nebraska contains 14 separate scales which attempt to measure various sensory-perceptual, motor, language, memory and intellectual functions of patients with brain impairments. These are shown in Table 1.

Table 1. Scales of the Luria-Nebraska Neuropsychological Battery.

Motor	
Rhythm	
Tactile	
Visual	Pathognomic
Receptive Speech	Right Hemisphere
Expressive Speech	Left Hemisphere
Writing	R* Scale
Reading	L* Scale
Arithmetic Skill	
Memory	
Intellectual	

The first 11 scales have no item overlap and their names adequately describe the functions the scales attempt to measure. The Pathognomic scale consists of items from the first 11 scales that are most sensitive to brain dysfunction. The Right Hemisphere and Left Hemisphere scales consist of items from the Motor and Tactile scales that use the right and left hands respectively, and thus are sensitive to dysfunction in the sensory motor area of the two hemispheres. Certain test items have been marked to facilitate scoring of two empirically developed scales that measure right hemisphere dysfunction (R* scale) and left hemisphere dysfunction (L* scale). Eight localization and 30 factor-analytic scales have also been developed. We decided to investigate whether certain items from the Tactile and Visual scales validly assess tactile and visual abilities of aphasic patients.

METHODOLOGY

Stimuli. The Tactile scale is designed to test the ability to localize tactile stimuli, to identify various types of tactile stimuli such as common objects, and to discriminate between intensities of tactile stimulation. Two-point discrimination, direction of touch, and identification of geometric shapes, numbers, and letters are assessed. The patient is blind-folded. Right and left limbs are tested separately. Fourteen of the 22 items in this scale require a verbal response. For each of these 14 items two sets of response forms were compared--the standard form of the Luria-Nebraska and an alternate form in which the need for a verbal response was eliminated. For example, in the standard form, discrimination of "hard" versus "soft" touch is tested with the head of a pin applied to the back of the hand and asking the patient to say "hard" or "soft." The alternate form (Figure 1, Appendix) allowed the patient to point to one of two drawings, the top depicting hard, the bottom depicting soft.

The Visual scale is designed to test identification of objects, photographs and drawings. Recognition of time and direction is assessed. Analysis of "Raven-Matrices" type designs, drawings of stacked blocks, and rotated figures is tested. An alternate form was constructed for eight of the fourteen items in the Visual scale which required a verbal response. For example, in the standard form a visual identification task requires the patient to name a pencil, eraser, rubber band, and quarter. The alternate form (Figure 2, Appendix) allowed the patient to identify each item by means of association to one of six photographs, so that pointing to the coin machine correctly identified quarter, pointing to the rolled papers identified rubber band, and so on; two pictures are foils.

Three criteria, in order of priority, were used in developing the alternate response forms. First, expressive language should not be required in the response component. Second, the relative level of complexity of the task should be maintained--that is, an identification level task in the standard form should continue to require identification level processing in the alternate form. Third, if possible, only the modality under test should be involved.

Subjects. Eighteen left hemisphere aphasic subjects (mean age = 57.6 yrs), eight right hemisphere nonaphasic subjects (mean age = 60.13 yrs), and 26 normal controls (mean age = 58.62 yrs), participated. All brain damaged subjects had suffered a single, unilateral lesion, and were beyond one month post onset. Aphasic subjects were identified by standardized testing and were given the Boston Naming Test (Goodglass and Kaplan, 1976). The average severity of the aphasic group (derived from overall PICA (Porch, 1967) scores) was 46.5%, the range being 18 to 84%. Right hemisphere patients were screened to rule out significant language difficulties.

Procedures. Across two testing sessions, each subject received counterbalanced presentations of the standard and alternate forms of the 14 Tactile scale items and the 8 Visual scale items. Sessions were conducted on consecutive days. Standard Luria-Nebraska scoring (plus-minus) was used.

RESULTS

Because of the heterogeneity within the Visual and Tactile scales (each measures a number of different functions within a modality) the data analysis focused upon each item as a statistical entity. A t-test for related samples compared the number of errors made on each item for the standard versus the alternate forms. Separate comparisons were constructed for each subject group on each item. Results will be reported as the difference in mean percent error.

Table 2 shows these results for each item of the Tactile scale listed separately for the aphasic, right hemisphere and control groups. The 14 items actually represent seven different tasks, each presented twice, once to the right and once to the left limb. A positive difference in percent error indicates that less errors were made on the alternate, nonverbal form. A negative value means less errors were made on the standard, verbal form. The larger the number, the greater the difference between the two forms.

Clearly, the alternate form had its greatest impact on the aphasic group. Compared to the other two groups, difference scores of the aphasic group are larger, and they are all positive. For seven of the 14 Tactile scale items, the aphasic group made significantly less errors on the alternate form, while performance of the right hemisphere and control groups was not significantly changed. For the seven remaining items, six produced no significant changes in error rates, while one significantly decreased error rates for the aphasic and control groups.

Table 2. Difference in mean percent error for tactile scale items.

Item	Aphasia (18)	Right Hemisphere (8)	Controls (26)
66	14.4**	- 7.5	- 3.8
67	24.4	- 7.5	3.8
68	19.4*	- 3.0	- 2.9
69	33.3**	9.4	- 3.8
70	8.3	-12.5	1.9
71	31.3	0	0
72	11.1	-16.0	0
73	13.9	-12.6	0
76	16.6	0	15.4
77	33.3**	25.0	3.9
78	38.9**	25.0	34.7**
79	33.3**	0	23.1
82	13.9*	12.5	2.9
84	43.1***	9.4	- 1.9

Two-tailed probabilities for t
* $p < .05$
** $p < .01$
*** $p < .001$

Interestingly, the effectiveness of an alternate form did not depend on whether the right limb (usually hemiplegic) or the left (nonhemiplegic) limb was tested. That is, if an alternate form allowed the aphasic group to make significantly less errors in a tactile processing task, it usually did so for the right as well as for the left arm and hand. The success of some alternate forms may have been related to the complexity of the task and the amount of language required in the standard form. For example, the alternate form in Figure 3, Appendix, eliminated the need to verbalize that "one" or "two" points were being presented. Another alternate form (Figure 1, Appendix) eliminated the need to verbalize "hard" or "soft." While aphasic subjects made less errors on the alternate forms of both these tasks, only the latter achieved significance. Two out of three identification level tasks achieved significance in the alternate form. For example, the array in Figure 4, Appendix, allowed the subject to identify the letter "S" which had been traced on the back of his wrist.

Table 3 shows the difference in mean percent error for each item of the Visual scale. Again, a positive value indicates that fewer errors were committed on the alternate form, and a negative value means that fewer errors were made on the standard form. For items 88, 89, and 94, aphasic subjects made significantly fewer errors on the alternate form while performance of the other two groups was unchanged. For example, on item 94 the examiner asks "tell me what time these clocks tell on this card," (stimulus card is shown depicting four horizontally-arranged clocks without numerals). For each clock the alternate form (Figure 5, Appendix) allowed the patient to select the time from a list. For items 86, 90, and 91, effects of the alternate form varied. Error rates of the aphasic group were unchanged for items 86 and 90, but decreased for item 91, while errors on the alternate form increased for the right hemisphere and control groups, shown by the significant, negative difference scores.

Table 3. Difference in mean percent of error for visual scale items.

Item	Aphasia (18)	Right Hemisphere (8)	Controls (26)
86	12.5	-21.8*	- 6.7*
87	37.8**	15.1*	13.1***
88	28.9**	5.3	- 1.5
89	27.8**	43.9	3.9
90	18.9	-38.0	-14.6***
91	25.6*	-17.1**	- 8.5**
94	37.5**	9.3	4.8
97	4.2	9.3	1.9

Two-tailed probabilities for t
* $p < .05$
** $p < .01$
*** $p < .001$

In such tasks the subject was asked to name a group of items (tea kettle, bottle, bowl, glass, and fork) drawn in an overlaid fashion. In the alternate form the subject selected five of ten photographs in which the target item was not present, but which depicted a context closely associated with the target item. For instance, a fork is associated with a photograph of a place setting minus the fork, the glass with a photograph of Alka-Seltzer and a pitcher of water, etc. However, in our attempt to maintain an identification level of difficulty we apparently made such alternate forms visually and/or conceptually more complex, adversely affecting the performance of the right hemisphere and control groups, as shown by the significant, negative difference scores in Table 2. Finally, for item 87 the alternate form decreased percent error for all three groups. We attribute this to the fact that even normals have difficulty naming some of the stimuli on the standard form, such as vial and nutcracker. Our alternate array of contextual pictures may have simplified identification of these unusual stimuli.

Because many traditional methods for treating aphasia require intact visual-spatial skills (e.g., Helm-Estabrooks, et al., 1982) it was decided to explore further the performance of the aphasic subjects on the Visual scale items. Table 4 shows error rates from both the standard and the alternate forms of the Visual scale items correlated with the number of correct responses on the Boston Naming Test. The significant negative correlations indicate that the more errors subjects with aphasia make on standard Visual scale items, the smaller is the number of their correct responses in naming, the average correlation being -.798. However, when the verbal component is taken out of the task with response alternatives, the resulting correlations are low, the mean correlation reduced to -.249. In other words, while the Boston Naming Test accounts for over 63% of the variance in these Visual scale items ($r^2 = .636$), the visual component of these two tasks accounts for only 6% of the variance ($r^2 = .062$). Thus, naming may account for 57% of the variance in the Luria-Nebraska Visual Scale items in question.

Table 4. Correlations of standard and alternative visual scale items with Boston Naming Test for aphasia group (N=15).

Item	Standard	Alternative
86	-0.901**	-0.170
87	-0.910**	0.034
88	-0.835**	-0.204
89	-0.374	0.090
90	-0.901**	-0.486
91	-0.938**	-0.346
94	-0.758**	-0.198
97	-0.767**	-0.716**
$\bar{x}$ =	-0.798	-0.249
r^2 =	.636 (63%)	.062 (6%)

CONCLUSIONS

The data reported here question the validity of items of the Luria-Nebraska for assessing visual and tactile abilities of aphasic patients. Specifically, our concern is one of construct validity--to a large extent items designed to measure tactile and visual perception are measuring naming. Certainly, naming is an anticipated result of normal tactile and visual perception. However, tests of these processes should not depend on naming to determine whether tactile and visual perception is impaired.

Golden (1982) has indicated that such problems are not a weakness of the Luria-Nebrasks, but are an indictment of persons who misuse the test. Golden's inability to fully comprehend our criticisms is not surprising when one considers his recent conclusion (Golden et al., 1980) that because most of the Luria-Nebraska scales are more sensitive to left hemispehre dysfunction, such a finding validates Luria's concept that the left hemisphere coordinates activity between the two hemispheres. We suggest that a confounded test battery is not the best method for testing such a hypothesis.

The utility of the Luria-Nebraska for delineating functional deficits in aphasia is obviously limited. At best, use of the battery in its current form requires administration and interpretation by a clinician who is highly skilled in the diagnosis of aphasia and who possesses supplementary evaluative techniques. At worst, it can lead to misinterpretation of deficits in aphasic patients. Incorrect assessment of sensory functions in the aphasic patient can lead to an underestimation of his potential for rehabilitation. Conversely, those aphasic patients who do have sensory deficits need to be identified.

REFERENCES

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

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

Goodglass, H. and Kaplan, E. The Assessment of Aphasia and Related Disorders. Philadelphia: Lea and Febiger, 1976.

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

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

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

Helm-Estabrooks, N., Fitzpatrick, P. and Barresi, B. Visual action therapy for global aphasia. Journal of Speech and Hearing Disorders, 47, 385-389, 1982.

Holland, A. An analysis of the utility of the Luria-Nebraska Neuropsychological Battery in assessing patients with language disturbances. A workshop presented at the Northport V.A. Regional Educational Medical Center, Northport, New York, 1982.

Lezak, M.D. Neuropsychological Assessment. New York: Oxford University Press, 1983.

Porch, B.E. Porch Index of Communicative Ability. Palo Alto, California: Consulting Psychologists Press, 1967.

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

APPENDIX

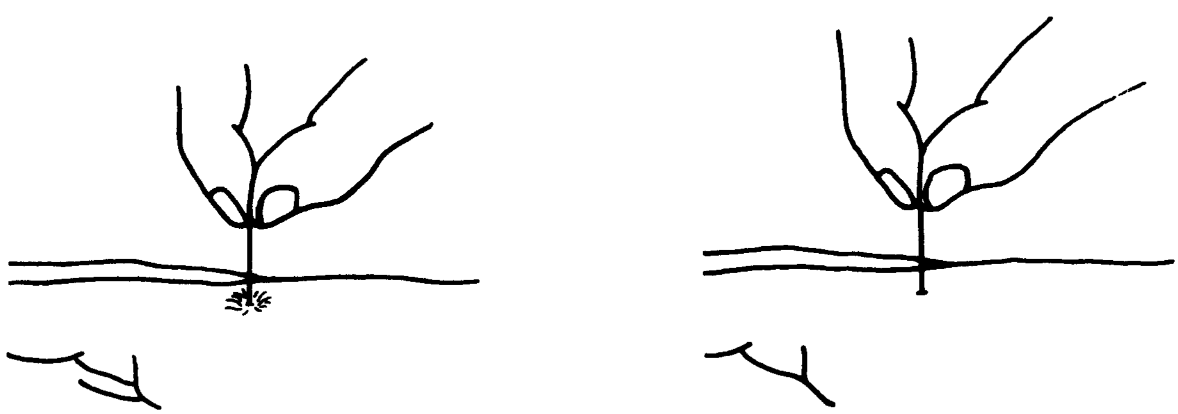

Figure 1. Alternate response form for testing tactile perception of "hard" vs "soft".

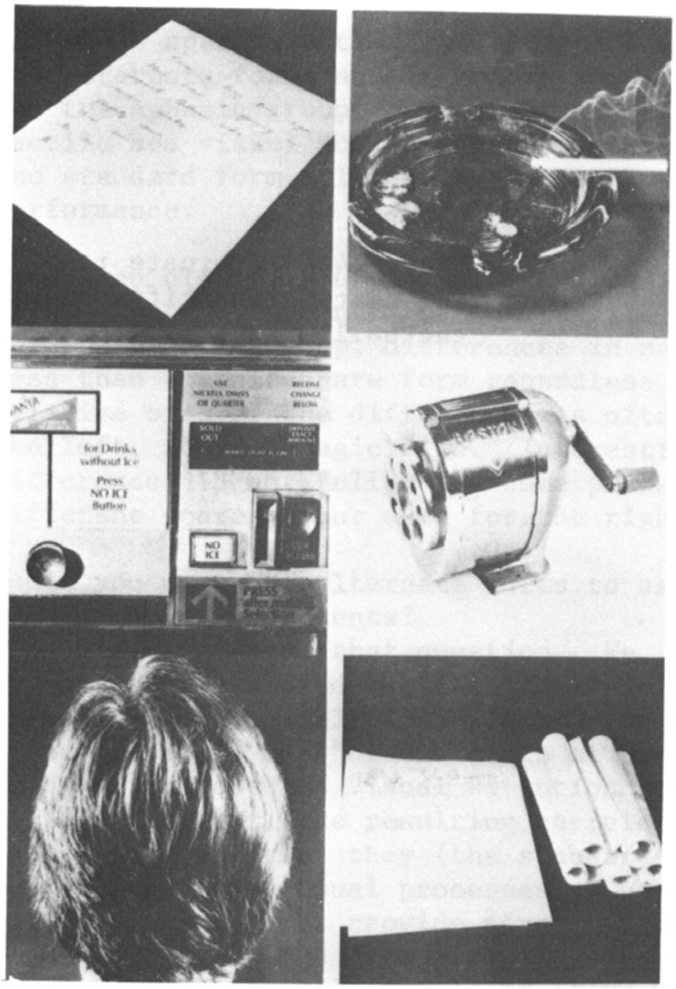

Figure 2. Alternate response form for testing visual identification.

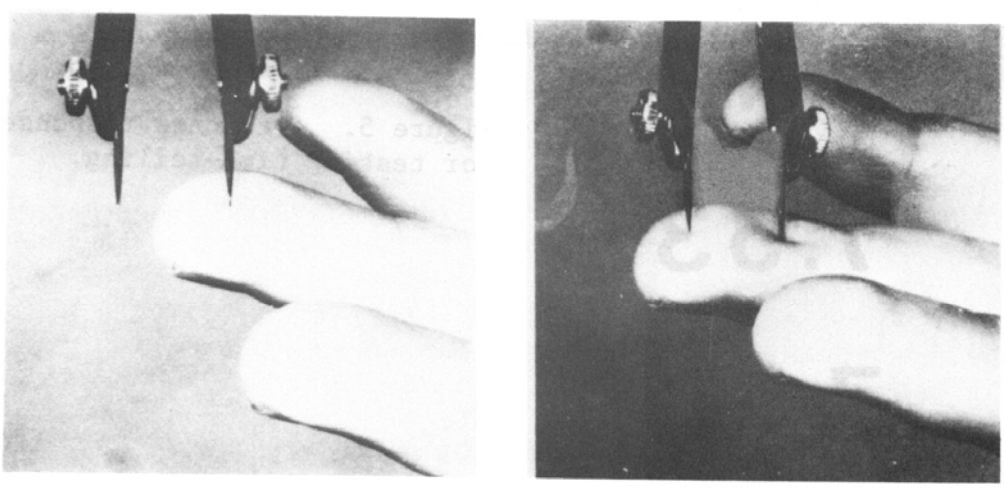

Figure 3. Alternate response form for testing two-point discrimination.

T

S

R

B

Figure 4. Alternate response form for testing identification of tactually traced letters.

7:55

10:40

7:53

7:51

Figure 5. Alternate response form for testing time-telling.

DISCUSSION

Q: Of the 18 aphasic patients, were there any who did not perform better on the alternate form; and if so, to what did you attribute the difference?

A: For the aphasic group, percent error rate in the alternate form for all tactile and visual scale items was always less than percent error rate in the standard form. I cannot adequately comment on individual subject performance.

Q: Would you clarify your point regarding influence of hemiplegia on right versus left performance for the tactile scale items?

A: For the aphasic group, differences in mean percent error were consistently less than the alternate form regardless of the limb tested. However, the relative size of the difference was often greater for items which tested the left, nonhemiplegic limb. Interestingly, significance of mean percent difference did not follow the same pattern. Out of eight significant different scores, four were for the right, usually hemiplegic limb.

Q: Would you use your alternate forms to assess visual and tactile abilities in your aphasic patients?

A: I am glad you asked that question. We, too, have been curious as to what aspects of visual and tactile processing are being tapped by our alternate forms. We did correlations between the performance of the standard and alternate items and "other" tests of visual perception such as the Facial Recognition Test and Visual Retention Test developed by Benton, et al. (1978). However, the resulting correlations were quite low. At present we can only say that they (the standard and alternate forms) apparently do not tap the same visual processes as other batteries. To put it another way, our data do not provide strong evidence of validity for the alternate forms. However, the data strongly suggest that performance on standard Luria-Nebraska tactile and scale items is highly related to naming ability.

Q: Is this test (Luria-Nebraska) widely used?

A: That's really a question for my collaborator who is a neuropsychologist. My guess is that the use and popularity of this test is somewhat regional. Regardless, the battery has received significant attention in the psychological literature.