CHAPTER

14

The Limb Apraxia Test: Development of a Short Form

Joseph R. Duffy
Robert J. Duffy
Deborah Uryase

Chapter 13 described the Limb Apraxia Test (LAT) (Duffy, 1974) and its validity, replicable sensitivity to deficits, preservation of information about limb movements, and capacity to provide information about factors that may contribute to limb movement deficits. The LAT has been used in studies of nonverbal communication in aphasia (Duffy and Duffy, 1981) and has potential as a research tool to objectify the presence and severity of limb apraxia and aid our understanding of its behavioral characteristics and nature.

In contrast to its usefulness as a research tool, the LAT has at least three limitations as a clinical measure:

1. It contains eight subtests, with 10 items in each subtest. There are a total of 252 movement components in the test, each of which must be scored and averaged to obtain item, subtest, and overall scores.
2. The scoring system is complex, containing 21 categories that require judgments about five different dimensions of response adequacy.
3. Responses occur relatively quickly, and frequently in sequences of up to six movement components, each of which must be scored; this makes it difficult to assign scores during administration and usually necessitates videotaping and scoring at a later time. Thus, while test administration usually takes only 20 to 30 minutes, scoring of the videotape takes at least that long.

These factors combine to make administration and scoring a tedious process, one unlikely to be adopted for routine clinical use in which the goal of assessment is to identify the presence and severity of the disorder.

The purpose of this study was to develop a short form of the LAT. Two basic approaches to shortening the test seemed reasonable: first, reduction of the number of items; second, reduction in the complexity of the scoring system.

METHOD

SUBJECTS

Because the LAT is designed primarily for identification and description of limb apraxia in people with left-hemisphere damage (LHD), the identification and cross-validation of short forms was based on performance of patients with single unilateral left-hemisphere lesions.

Two groups of subjects will be referred to:

LHD1 (N = 20). These patients represent the LHD sample on whom the LAT was originally developed (Duffy, 1974). Item, subtest, and scoring category analyses of this group's performance were used to identify potential short forms.

LHD2 (N = 21). This group of patients was given the eight-subtest, 80-item original long form of the LAT as well as the short forms. This group served to cross-validate the short forms. Although this cross-validation group was not selected to match the LHD1 group relative to LAT performance, their overall LAT scores were not statistically significantly different ($p > .05$) from those of the LHD1 group or another larger group of 36 LHD subjects who were tested as part of another study (Duffy and Duffy, 1981). In addition, as was true for the LHD1 group, they were statistically significantly inferior to groups of control and right-hemisphere-damaged subjects (Duffy and Duffy, 1988). Therefore, the cross-validation group was as similar in its representativeness of patients with unilateral LHD as was the LHD1 group.

IDENTIFICATION OF SHORT FORMS

Examination of the LHD1 group data indicated that all LAT subtests were statistically significantly ($p < .01$) and highly correlated with one another and the overall LAT score; all subtest correlations with the overall LAT score exceeded .86, and the intersubtest correlations ranged from .67 to .96, with most exceeding .80. This suggested that the overall LAT score might accurately be predicted by a small number of subtests or by a small number of items from among the eight subtests.

Two potential short forms that would reduce the number of test items were identified. First, a stepwise multiple regression with the overall LAT score as the dependent variable and the eight LAT subtest scores as independent variables indicated that subtests VII and II predicted 98 percent of the variance in the overall LAT score (see Table 14-1). This high predicted variance and the fact that these two subtests cover a range of binary contrast features assessed by the LAT made it a reasonable short-form candidate.

The second short form contained one item from each of the eight subtests. The items selected from each subtest were identified in two steps. In the first, a stepwise multiple regression was run for each of the eight subtests, with the overall LAT score as the dependent variable and the subtest's items as independent variables. This identified the best item in each

TABLE 14-1. THE TWO-SUBTEST SHORT FORM VERSION OF THE LAT, FROM STEPWISE MULTIPLE REGRESSION ANALYSIS OF PEFORMANCE OF THE LHD1 GROUP

Subtest	*Features*	*Multiple r*	*r square*
VII	Segmented/Complex/Object (e.g., spooning sugar into cup)	.975	.951
II	Sequenced/Simple/Object (e.g., turn cup over to remove block, put cup down, place block on cup)	.988	.976

subtest at predicting the overall LAT score. Then these best eight items were looked at to be sure there were not duplications of items that differed only on their Segmented/Sequenced characteristics or Object/No-Object characteristics. This duplication occurred for two items. In these cases, the next best item in the Segmented subtest on which the duplication occurred was selected.

In the second step, these eight items served as independent variables in a stepwise multiple regression with the overall LAT score as the dependent variable. The results are summarized in Table 14-2, and indicate that the eight items together predicted 98 percent of the variance in the overall LAT score. Actually, the amount of variance explained was not significantly added to after the fourth item, but we decided to include all eight items for face validity purposes (i.e., to sample behaviors measured by each of the eight LAT subtests).

We were also interested in simplifying the scoring of the LAT. The most direct way of doing this was to use a plus-minus system of scoring. This required a decision about the plus-minus cutoff point on the 21-point LAT scoring scale. The LAT scores for subjects in the LHD1 group were recoded by scoring the 252 movement components in the test as plus or minus (0 or 1) using various cutoff points. Although many cutoff points yielded high correlations between scores computed using multidimensional scoring and plus-minus scores, it was decided to score the top five categories in the multidimensional scale as plus and the bottom sixteen categories as minus (see Chap. 13). Thus, any component scored incomplete-distorted or worse was scored 0 and those better than that (or 17 or above) as 1. There were two reasons for selecting this cutoff point. First, the correlation between overall LAT scores using multidimensional scoring and this method of plus-minus scoring was .933. Second, examination of the distribution of scores in each scoring category for 20 normal sub-

TABLE 14-2. THE ONE-ITEM-FROM-EACH-SUBTEST VERSON OF THE LAT, DERIVED FROM STEPWISE MULTIPLE REGRESSION ANALYSIS OF PERFORMANCE OF THE LHD1 GROUP

Subtest (item #)	*Multiple r*	*r square*
VII (4)	.947	.896
IV (6)	.978	.956
V (7)	.982	.964
VI (2)	.986	.972
III (1)	.988	.975
II (3)	.988	.976
I (3)	.989	.977
VIII (3)	.989	.978

jects indicated that 96 percent of their movement components were scored 17 or above. Similar analysis of 20 right-hemisphere patients indicated that 92 percent of their movement components were scored 17 or above. In contrast, only 78 percent of the LHD1 subjects' components were scored 17 or above. Thus, this cutoff point appeared to capture, on the plus side, adequate features of response by non-left-brain-injured individuals and, on the minus side, inadequate response features from left-brain-injured patients.

A second plus-minus approach was also identified. In this, if *any component* of an item received a multidimensional score of less than 17, the entire item was scored as minus. The correlation between this item approach to scoring and multidimensional scoring was .911.

CROSS-VALIDATION STUDY

Three versions of the LAT were administered to the 21 subjects in the cross-validation group (LHD2): the eight subtest, 80-item LAT; the two-subtest version; and the one-item-from-each-subtest (eight-item) version. Order of administration of the three versions was counterbalanced across subjects. Each LAT version was then scored in three ways: using the 21-point multidimensional scale, using plus-minus scoring for each compo-

nent of the test (component scoring), and using plus-minus scoring for each item of the test (item scoring).

RESULTS

It is best to examine the results for the eight-item short-form version first, as the results were disappointing. The correlation between the eight-item version and the original multidimensionally scored long form was only .185 and was statistically nonsignificant ($p = .21$). Correlations between the eight-item version and all other scoring methods for the eight-subtest and two-subtest versions were similarly low and statistically nonsignificant. Thus, the eight-item version was not adequately cross-validated and was discarded as a viable short form.

Cross-validation results for the remaining short-form versions are summarized versions are summarized in Table 14-3. For the eight-subtest version, plus-minus component scoring was correlated .96 with the criterion measure, or long-form multidimensionally scored LAT. The plus-minus item scoring of the eight-subtest version was correlated .89 with the long form.

For the two-subtest version, multidimensional scoring was better than plus-minus scoring and correlated .87 with the long form. The plus-minus component scoring was nearly as good with a correlation of .85 with the long form. Finally, the plus-minus item scoring of the two-subest version was least adequate, correlating .78 with the long form and explaining 61 percent of its variance.

TABLE 14-3. CROSS-VALIDATION RESULTS FOR THE EIGHT-SUBTEST AND TWO-SUBTEST SHORT FORMS SCORED WITH COMPONENT, ITEM, AND MULTIDIMENSIONAL SCORING METHODS

	Original LAT (multidimensional scoring)	
Short form	***r****	***r square***
Eight subtest		
+/− component	.955	.912
+/− item	.891	.794
Two-subtest		
Multidimensional	.874	.765
+/− component	.851	.725
+/− item	.780	.609

* All r significant ($p < .001$).

There are many other ways in which the adequacy of these short forms in predicting long-term performance can be assessed, but one further set of data sheds some light on their relative adequacy. Because the LAT appears useful as an objective index of the *presence* of limb apraxia, it is reasonable to ask if some of the short forms identify left-hemisphere-damaged patients with and without limb apraxia with sensitivity comparable to that of the long form. In the cross-validation sample (LHD2), 15 of the 21 subjects had limb apraxia based on long-form performance below the poorest score obtained by a control subject in our first study (see Chap. 13). To compare short-form to long-form sensitivity to the presence of limb apraxia, the cross-validation subject who scored in the normal range but was closest to the cutoff score for normal based on control subject performance on the long form was identified. This subject's scores on each short form were then used as the cutoff scores for normal versus apraxic. Using the long-form criterion for identifying subjects as apraxic versus nonapraxic, false-positive and false-negative rates for each short form were derived. As shown in Table 14-4, the false-positive and false-negative rates were small for the plus-minus component and item scoring for the eight-subtest version and were considered quite acceptable. The percentage of false positives and false negatives for the two-subtest version were not quite as good, with good false-positive rates but a tendency toward higher rates of false-negative identification or to calling a patient normal when they were apraxic (based on long-form performance). Generally, therefore, the plus-minus scoring method for the eight-subtest version is quite good. The two-subtest version does not fare as well, particularly in plus-minus scoring of items. It is important to note, however, that false positives and false negatives would be reduced for all short forms if a "gray area" of indeterminate diagnosis were established. Such a gray area score on a short form could be used to establish a need for more complete assessment (i.e., administration of the long form).

TABLE 14-4. FALSE-POSITIVE AND FALSE-NEGATIVE IDENTIFICATION OF LIMB APRAXIA FOR SUBJECTS IN THE LHD2 GROUP (N = 21)

Version	*False positive*	*False negative*
Eight-subtest		
+/− component	1 (5%)	1 (5%)
+/− item	2 (10%)	1 (5%)
Two-subtest		
Multidimensional	1 (5%)	3 (14%)
+/− component	1 (5%)	4 (19%)
+/− item	2 (10%)	4 (19%)

DISCUSSION

The results of this study suggest that there is considerable redundancy in the eight-subtest, 80-item version of the LAT that is scored using a multidimensional scoring table. They suggest that it is possible to shorten the LAT by reducing the number of test items and/or by simplifying the scoring system.

Cross-validation failed to confirm what looked like one promising short form that used only one item from each subtest. We did, however, identify several short forms that on cross-validation have validity as overall indices of limb apraxia. Especially promising are versions that retain all test items but modify the scoring system from multidimensional to plus-minus. A two-subtest version is also promising, especially if a gray area of performance on short forms is adopted to establish a need for further testing or, at least, a diagnosis of indeterminate. It appears that these short-form versions can decrease scoring complexity and scoring time and/or decrease administration time. These characteristics are desirable for many clinical and research purposes. For example, one or more of the short forms may be useful in objectively identifying the presence of limb apraxia in brain-damaged subjects. Similarly, they may prove adequate in documenting change in limb apraxia or as an index of severity of impairment. We do not believe, however, that these short forms should be used to examine issues relative to typologies of limb apraxia or those factors that influence limb apraxia, at least at this time. Such questions right now are best addressed with the long form.

REFERENCES

Duffy, J. R. (1974). *Comparison of brain injured and nonbrain injured subjects on an objective test of manual apraxia.* Unpublished doctoral dissertation, University of Connecticut.

Duffy, R. J., and Duffy, J. R. (1981). Three studies of deficits in pantomime expression and pantomime recognition in aphasia. *Journal of Speech and Hearing Research, 46,* 70–84.

DISCUSSION

Q = question; A = answer; C = comment.

Q. In your previous presentation you mentioned that the presence of limb apraxia in your right-hemisphere group, or in some members of

that group, made you wonder about the possibility of bilateral involvement. How was localization for your patients across those various studies determined?

A. In a variety of ways. CTs were available for some patients, particularly in the latter two left-hemisphere groups. For the remainder of the patients, "localization" was based on a history of only one event and clinical signs that localized the lesion to the middle cerebral artery distribution in the right or left hemisphere.

Q. Was your comment about bilaterality essentially a throw-away or do you think, for example, that in some of those patients who have clinical evidence of the type you've described confirming right-hemisphere lesion that an MRI, for example, might indeed show bilateral involvement or that positron emission tomography might show bilateral involvement?

A. It wasn't a throw-away, and I don't know the answer to the second part of the question. I can say that these patients did not have any clinical evidence of bilateral brain injury, and I simply don't know what an MRI or other study might show.

Q. Do you feel that praxis and apraxia do have some specificity? Do they have some localizing significance? Is there a neuro substrate that, while it may not be unique to the serving of limb movements, is certainly dedicated to limb movements? Will we find localizing value?

A. The literature certainly suggests there is. It strongly suggests that praxis is a dominant hemisphere function and strongly suggests that there is considerable overlap of the cortical and subcortical regions that are responsible for language and praxis. To that extent, this is a localizable function. How finely localized is a matter of further investigation.

Q. Why would you not expect to see some limb motor programming impairment in the right hemisphere? I mean, I know we don't see it very often, but what's your guess as to why?

A. The assumption, on the basis of available data, is that praxis is a dominant hemisphere function for control of limb movements as well as speech, so the prediction would be that this problem should not occur in lesions of the nondominant hemisphere.

One interesting thing about the performance of some of the right-hemisphere patients on the LAT is that there was a tendency for the group and certainly for some individual subjects in the group to have more trouble on some of the objects tasks than no-object tasks. This is interesting relative to the presumed constructional and visual-

spatial functions of the right hemisphere. I don't believe that factor explains all of the reduction in their scores, though.

C. A tiny point left over from the first study. I'm a little worried about the representativeness of your aphasic sample. I seem to remember that you said 70 percent of them or maybe 80 percent of them had hemiplegias.

A. Most were hemiplegic or hemiparetic. The figure is in the neighborhood of two-thirds to 75 percent.

Q. Doesn't it bother you a little bit in terms of people with posterior lesions? I mean, it seems that would be a group that we can logically expect to show some kinds of limb praxis, and aren't you a little worried that they were under-represented in your sample?

A. There is no doubt that they were not equally represented in the sample, although the sample is probably quite representative of the percentage of patients with hemiplegia and hemiparesis who come through our clinics and rehabilitation centers. We have looked at the performance of some fluent patients who had no hemiparesis and were tested using both hands. The small number whom we have data on have at least as high an incidence of the problem and had it in both limbs.

Q. I'm interested in your establishing that gray area in the short version of the test that will allow you better predictability. How will you go about establishing that gray area?

A. I think that's somewhat arbitrary or dependent on purpose. I can tell you that if on the two-subtest version you take the cutoff point, at about 19.3, and add and subtract a point and go from 18.3 to 20.3 and use that as your gray area, then you eliminate all false positives and negatives for the LHD2 sample. And you also would not have too many patients in the sample in the gray area.

C. It seems like your longer or complete plus-minus version, which may save you a lot of time, is a pretty darn good way to go, but maybe even that takes too long.

A. Well, that goes back to the question about purpose. Why do you want to give this test? Do you want to be able to say the person has the problem or doesn't have it, and if so, why do you want to do that? Are you using the test as a criterion for entry into a study? That is, you don't want any people who have the problem, or you want only people who have the problem? It's really a matter of purpose.

Q. Are you having them imitate you from across the table or side by side?

A. Across the table.

Q. So, they've got to do a switch. What would happen if you did it side by side or in unison?

A. We used the limb contralateral to the one used by the patient so they didn't have to reverse direction. Whatever they had to imitate went in the same direction as our movements from across the table.

Q. We know that aphasic patients can produce the same impaired performance for a number of different reasons. How comfortable do you feel in saying that 68 percent of left-hemisphere-damaged patients and some of your right-hemisphere-damaged patients produce the same impaired gesture for basically the same reason? You mentioned that there might be different types of apraxia. Could it be completely different impairments? Could the right-hemisphere-damaged patients, for instance, have impaired spatial representation? Could there be a lot of different underlying causes?

A. That is possible. Certainly the classic typologies of apraxia suggest differences. The patients with limb kinetic apraxia, according to much of the literature, are primarily awkward in their movements and do not necessarily make frank errors; some people argue that they shouldn't be called apraxic at all. People certainly talk about a difference between transitive and intransitive movements. The visual-spatial issue is an important one and may play a role in patients with right-hemisphere lesions.

One of the nice things about our multidimensional scoring system is that it does retain some characteristics of the behaviors that contribute to a score, and I think that may be very useful in identifying different forms of disturbance, if they exist. We've looked at profiles of patients, though, to see how many of them might correspond to the classic types; for example, transitive versus intransitive, or object versus no-object, and there just aren't many subjects who fall neatly into those categories. If you do poorly on one of the LAT subtests, you tend to do poorly on all. You may have relatively more difficulty on one subtest than another, but they tend to predict one another's performance. But at a basic level I think you're right. You can take a language test and fail for reasons that go well beyond what we call aphasia, and I'm sure that can happen on the LAT as well. The advantage of the test is its standardized objectivity, but the clinical diagnosis of limb apraxia, like the diagnosis of aphasia, also requires clinical judgment.