CHAPTER

11

The Differentiation of Normal from Aphasic Performance Using PICA Discriminant Function Scores

Deborah Brauer
Malcolm R. McNeil
Joseph R. Duffy
Robert L. Keith
Michael J. Collins

Most aphasia tests are designed to describe the degree and perhaps the type of language deficit that exists in an individual. In most cases, the person being tested is assumed to have a deficit, which is then described by the test administered. There is, however, a need for aphasia tests that will enable a valid and reliable differentiation between performance that is within the normal versus an impaired range. That is, a test is needed for the detection of aphasia. A number of aphasia tests provide information on normal as well as impaired performance, but none provides specific methodological guidelines or their degree of success in the differentiation of the two.

One means for assigning individuals to groups that is based on test performance is through discriminant analysis. Porch, Friden, and Porec (1976) used this statistical procedure to differentiate aphasic individuals from those who have a nonorganic component to their symptoms (malingerers). Using the Porch Index of Communicative Ability (PICA), Porch and colleagues reported considerable success in differentiating normal subjects from malingerers.

Porch and colleagues' (1976) discriminant function has, however, been extended to other populations without systematic examination of the appropriateness of this extension. Deal, Deal, Wertz, Kitselman, and Dwyer (1979) used the PICA discriminant scores generated by Porch and colleagues to demonstrate the presence of aphasia in right-hemisphere lesioned adults. McNeil, Wertz, Deal, Burkhard, and Collins (1983) used the same discriminant weights to examine the ability of normal adults with varying levels of understanding of aphasia to simulate the disorder. Those subjects whose performance produced a discriminant score in the aphasic range in the Deal and colleagues' study were considered to have an aphasic language deficit, and those in the McNeil and colleagues' study were considered to have the ability to simulate aphasia. However, if normal subjects were not differentiable from asphasics using this technique, then the interpretation of either study could be questioned. That is, in the Deal and colleagues' study, the people who were not identified as belonging to the right-hemisphere group could as easily have belonged to normal group. Likewise, in the McNeil and colleagues' study, those people identified as successfully feigning aphasia could as likely have performed as normal subjects unless normal subjects could be shown to be differentiable from the aphasic group. We will present the results of two studies designed to investigate these issues.

It was the purpose of this first study to examine the accuracy of the aphasic discriminant function cutoff scores generated by Porch and his colleagues for classifying normal subjects.

STUDY I

SUBJECTS

The subjects were 120 normal adults, ranging in age from 20 to 96 years, with a mean of 56 years, ranging in education from 3 to 18 years, with a mean of 12 years (Table 11-1). These normal subjects were derived from the normal PICA reference data published by Duffy and Keith (1980). These subjects were considered normal on the basis of a negative neurological speech and language history, no evidence of a substantive visual acuity or hearing impairment, and a score of not more than two standard deviations below the age mean for part V of the Token Test (Wertz, Keith, and Custer, 1971). The standard PICA was administered to each subject by a trained and reliable test administrator and scorer. The subtest percentiles were calculated for each subject and weighted according to the discriminant function generated by Porch and colleagues (1976). Those weighted scores falling below −.279 were considered nonaphasic and those falling above −.223 were considered aphasic. Those subjects falling between these values were considered unclassifiable. The number of subjects classified as aphasic, nonaphasic, and unclassifiable was then calculated.

RESULTS AND DISCUSSION

Using the discriminant cutoff scores listed above, 115 of the 120 (96%) normal adults were classified as aphasic (Table 11-2). Unlike the 115 subjects classified as aphasic, each of the five subjects not classified as aphasic showed errors on visual subtests with generally good overall performance.

It is concluded that the discriminant function weightings offered by Porch and colleagues (1976) did not differentiate the two groups successfully. Because of this failure to differentiate groups with these weightings, an additional discriminant analysis using aphasic and normal subjects (*not* attempting to feign aphasia) was undertaken.

TABLE 11-1. SUBJECT DESCRIPTIVE INFORMATION

	Normal adult	***Aphasic adult***
Number of subjects	120	78
Mean age (years) [range]	56.31 [20–96]	60.10 [21–89]
Time post-onset (months) [range]	—	12.99 [0.5–144.0]

TABLE 11-2. CLASSIFICATION OF SUBJECTS IN THE CURRENT STUDY USING THE PORCH AND COLLEAGUES' DISCRIMINANT FUNCTION

Group	*Percent classified as aphasic*	*Percent unclassifiable*	*Percent classified as malingering*
Normal	96	1	3
Left	94	5	0

STUDY II

The second study used the normal subjects from the first study along with a second group of 78 aphasic adults. The aphasic subjects were selected from patient files at the Veterans Administration Medical Center Hospital in Madison, Wisconsin. All subjects in this group had undergone a full assessment by a speech-language pathologist and were diagnosed as aphasic on the basis of a battery of tests. The PICA was administered to each patient at some point during the course of treatment.

A new discriminant analysis was performed using the data from these two subject groups. The classification of each individual subject in this analysis was determined at three different levels of significance: .05, .10, and .20. Following this analysis, a cross-validation study was conducted. This cross-validation was done by assigning 29 (15%) of the cases, randomly selected from the total of the original two subject groups, to "new" groups. These subjects were not used in generating the discriminant analysis. The classifiability of these new groups was then used to check the ability of the discriminant function to classify new cases.

RESULTS AND DISCUSSION

This discriminant analysis using the normal and aphasic subjects resulted in assignment of the vast majority of both normal and aphasic subjects to the appropriate group. The percentage of correctly classified subjects ranged across alpha levels from 82 to 88 percent for the aphasic group and from 97 to 99 percent for the normal group. The overall rate of correct classification ranged from 91 to 95 percent. These percentages, along with the cutoff scores for group assignment, are included in Tables 11-3, 11-4, and 11-5. It can be seen from these tables that no normal subjects and a very small percentage of the aphasic subjects were misclassified. In all cases, considerably more subjects were unclassifiable than were misclassified.

TABLE 11-3. CORRECT CLASSIFICATION OF APHASIC AND NORMAL ADULTS AT THE .05 LEVEL

Group	*Range*	*Percent correctly classified*	*Percent unclassifiable*	*Percent misclassified*
Normal	2.67 to 0.41	97	3	0
Unclassified	0.40 to −1.26			
Left	−1.27 to −5.23	82	14	4
Total classified		91	7	2

TABLE 11-4. CORRECT CLASSIFICATION OF APHASIC AND NORMAL ADULTS AT THE .10 LEVEL

Group	*Range*	*Percent correctly classified*	*Percent unclassifiable*	*Percent misclassified*
Normal	2.67 to 0.41	99	1	0
Unclassified	0.13 to −1.10			
Left	−1.11 to −5.23	83	12	5
Total classified		93	6	2

TABLE 11-5. CORRECT CLASSIFICATION OF APHASIC AND NORMAL ADULTS AT THE .20 LEVEL

Group	*Range*	*Percent correctly classified*	*Percent unclassifiable*	*Percent misclassified*
Normal	2.67 to 0.41	99	1	0
Unclassified	0.13 to −0.82			
Left	−0.83 to −5.23	88	6	5
Total classified		95	3	2

The subjects for the cross-validation study were a subset of each group. The rate of correct classification with this cross-validation was very similar to the results of the original analysis (Table 11-6). The new aphasic group had a higher percentage misclassified and a slightly lower percentage correctly classified, although this may be an artifact of the small number of subjects in this group, which was seven.

In the current study comparing the normal and aphasic subjects, seven variables were selected as providing the best discrimination between the two groups. These were the total subtest score for subtests I, IV, VI, IX, A, and B and time (the total time required for test administration). The weights for each variable are shown in Table 11-7.

The discriminant analysis reflects the statistical function that best characterizes the difference between the groups being analyzed. The Porch and colleagues discriminant cutoff scores are based on a function generated from aphasic versus malingering performance. In the first study, normal performance was not differentiable from aphasic performance using the Porch and colleagues discriminant scores. In the second study, a discriminant function was generated that was quite successful in differentiating aphasic from normal performance. The difference between these two discriminant functions suggests that the features that differentiate malingering from aphasic performance on this test are not the same features that will differentiate normal from aphasic performance.

The different alpha levels do not cause great changes in the numbers classified. As the confidence level increases, the number of unclassified

TABLE 11-6. WEIGHTS USED IN DISCRIMINANT ANALYSIS

Variable	*Weight*
I	0.72231
IV	−0.53449
VI	−0.34390
IX	−0.04337
A	0.57231
B	0.49651
Time	−0.54599
Constant	−2.48663

TABLE 11-7. CLASSIFICATION OF APHASIC AND NORMAL ADULTS — CROSS-VALIDATION AT THE .05 LEVEL

Group	*Range*	*Percent correctly classified*	*Percent unclassifiable*	*Percent misclassified*
Normal	2.88 to 0.49	95	5	0
New normal	2.53 to 0.42	95	5	0
Unclassified	0.41 to −0.99			
Left	−1.11 to −4.91	85	11	3
New left	−1.00 to −4.03	72	14	14
Total classified		91	7	2

subjects increases. Conversely, the number of correct classifications decreases. The percentage of misclassified subjects does not change. In other words, at the .05 level, there are more unclassifiable subjects and fewer classified either correctly or incorrectly than at the .20 level. The reason for analyzing the data at several probability levels needs explanation. There are times when taking a chance on being wrong on an initial diagnosis 4 out of 20 times would be a risk worth taking. Perhaps this risk level would be appropriate for the assignment of a person to a particular behavioral treatment regimen when the adverse effects (e.g., financial, psychological) would be minimal. Perhaps a 1 in 20 error rate would be more acceptable if one were assigning persons to treatment efficacy studies or to pharmaceutical protocols based on the correct initial classification. Since correct, unclassified, and misclassified rates vary with the alpha level, the clinician or researcher can choose those classification trade-offs that best meet the patient's needs.

The results of this study suggest that caution must be used in extending the applicaton of any discriminant analysis. It appears that this type of analysis is not necessarily applicable to groups other than those for whom it was designed and for whom the function was generated. The characteristics that best distinguish any two groups, and hence those characteristics that are used for deriving discriminant weightings, may differ from those that best differentiate among any other groups. In the McNeil and colleagues' (1983) and Deal and colleagues' (1978) studies, discriminant scores were assumed to reflect an aphasic language disturbance or a simulation thereof, respectively. In those studies, it is equally likely that a discriminant score interpreted as reflecting aphasic performance was actually within the normal range or, perhaps, even within the range of another, as yet undescribed

pathological group. In the Deal and colleagues' study, Porch and colleagues' discriminant score was used to determine whether PICA performance following a right-hemisphere lesion resembled aphasia. The results of the present study support the suggestion that this was an improper application of the function. In that study, 62 percent of the subjects were classified as aphasic. That is, Deal and his colleagues (1979) concluded that the right-hemisphere damaged subjects had subtle and previously undetected aphasia. For those subjects not classified as aphasic, their interpretation was that the PICA was either incapable of detecting the aphasia that was present in this population or that 38 percent of the subjects were not in fact aphasic. A third possible interpretation now exists. It may be that the Porch and colleagues discriminant function was not the appropriate analysis to detect aphasia in this population. It is also possible that of the 62 percent who were classified as aphasic, some or all may have been performing as normals, since normal adults will also be classified as aphasic using those discriminant weights.

In the McNeil and colleagues' study, Porch's discriminant function analysis was applied to one of the groups for whom it was originally intended. However, a portion of the results may require additional interpretation. In this study, a number of people with varying degrees of familiarity with aphasia were asked to simulate aphasia. Those who received discriminant scores within the aphasic range were judged to be able to simulate aphasia accurately. However, as with the Deal and colleagues' study, it is as likely that some or all of those classified as aphasic actually performed as they would have if not attempting to feign aphasia.

The second discriminant function analysis in this investigation, with the cross-validation, offers some additional confidence that the discriminant function generated is capable of correctly classifying aphasic from normal performance on the PICA. The newly derived PICA discriminant scores appear to have immediate applicability for both clinical and research purposes. The vast majority of subjects in each group were classified correctly. No normal subjects were misclassified as aphasic at any alpha level, and only 4 to 5 percent of the aphasic subjects were misclassified as normal. Therefore, any person who is classified as aphasic usng these discriminant weights is extremely unlikely actually to be a normal, although a subject classified as normal has a very small chance of being aphasic. The unclassifiable range reduces the chance of a subject's being classified either correctly or incorrectly. The discriminant weights generated in this study offer a high level of correct classification with minimal misclassification of subjects. Since these findings have been replicated with a smaller set of data, there appears to be sufficient evidence available for its use. However, given that these aphasic and normal subjects may not represent the universe of subjects to which one might like to generalize these dis-

criminant weights, there is of course a need for replication of these findings. Until they have been replicated, we recommend cautious use of these data for the differentiation of normal from aphasic persons with the PICA.

ACKNOWLEDGMENT

This research was supported by NICHD Grant No. 5 P30 HD03352.

REFERENCES

Deal, J. L., Deal, L., Wertz, R. T., Kitselman, K., and Dwyer, C. (1979). Right hemisphere PICA percentiles: Some speculations about aphasia. In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 9). Minneapolis: BRK.

Duffy, J. R., and Keith, R. L. (1980). Performance of non-brain injured adults on the PICA: Descriptive data and a comparison to patients with aphasia. *Aphasia-Apraxia-Agnosia, 2*(2), 1–30.

McNeil, M. R., Wertz, R. L., Deal, J. L., Burkhard, B., and Collins, M. J. (1983). Aphasia: Can it be taught? Can it be caught? In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 13). Minneapolis: BRK.

Porch, B. E. (1971). *Porch Index of Communicative Ability.* Palo Alto, CA: Consulting Psychologists Press.

Porch, B. E., Friden, T., and Porec, J. (1976). *Objective differentiation of aphasic versus nonorganic patients.* Paper presented to the meeting of the International Neuropsychology Association, Santa Fe, NM.

Wertz, R. T., Dronkers, N. F., and Shubitowski, Y. (1986). Discriminant function analysis of performance by normals and left hemisphere, right hemisphere, and bilaterally brain damaged patients on a word fluency measure. In R. H. Brookshire (Ed.), *Clinical aphasiology* (Vol. 16). Minneapolis: BRK.

Wertz, R. T., Keith, R. L., and Custer, D. D. (1971). *Normal and aphasic behavior on a measure of auditory input and a measure of verbal output.* Paper presented at the Annual Convention of the American Speech and Hearing Association, Chicago.

DISCUSSION

Q = question; A = answer; C = comment

Q. Can you tell us a little bit about the aphasic patients who were misclassified? Were they mild or severe?

A. There was a tendency for them to be the more mild patients. However, they were not only the mild patients. It wasn't anything as clear-cut as severity that resulted in the misclassifications.

Q. Also, I'm familiar with Porch's 1976 study. He didn't use time. Is this the first time you've used time as one of the variables?

A. Yes. In the Duffy and Keith study they used time in describing the normal performance, and there did seem to be a clear range, so we thought it was worth including.

C. I'm impressed. Porch has always said it's an important factor and yet, until I heard you mention it, I didn't think of using it in that way.

Q. Would you pass your weights out now and say this is what you need for the differentiation of aphasic from normal performance?

A. While the subject numbers were fairly large, the study does need to be replicated before I would say these are the final weights.

Q. It seems like this kind of questioning is one of specificity. Do we have a sign, a measure, a test that will specify the presence of aphasia? Some diseases and some tests are highly specific. How specific would you need to be? You classified 92 percent of the patients, and that's pretty good. And the whole purpose in this, beyond the theoretical question, is can you use it clinically? And when do we stop generating large-*N* studies and pass out the weights and say this is good? We don't have to get 100 normals and 78 aphasics and do it again. When do we say that your weights are good enough for all of us to use?

A. I think it would be safe to use them clinically with some caution. I wouldn't use this as the only means of determining the presence of aphasia. But, along with other things, it would be worth using them.

Q. How was the correct classification defined? How was the original decision made that the person was aphasic, and then how does that relate to the discriminant function analysis?

A. How did we select our aphasic subjects? Is that the question?

Q. Yes. The discriminant analysis shows the aphasic person is, in fact, aphasic. How did you decide in the first place that he was aphasic?

A. The data were all taken from patient files at a VA hospital, and there was no set criterion for what aphasia was. It was based on the results of a battery of tests, and there was a general definition of aphasia that everybody had in mind when they made the decisions.

Q. The reason I ask the question is because there is a problem in regarding aphasia and nonaphasia as a binary system. You can be a little bit aphasic, or you can be grossly aphasic. It seems to me that one has to predefine, in this sort of a situation, what one is going to regard as aphasic. And my question is, do you have some people who were a

bit aphasic and do you have some people who were grossly aphasic? So the question is, what were your original decisions regarding the groups — aphasic versus nonaphasic — based upon?

A. We had a wide range of severity in the aphasic group. The overall PICA percentiles ranged from the 2nd to the 94th percentile. So we have some subjects who were a little aphasic and some subjects who were severely aphasic.

Q. OK. It's a sort of philosophical mind-type question that just occurred to me, and I thought I'd toss it out. Does anybody have a response?

A. In fact it is a good theoretical and philosophical question. From my point of view, and that's not just to defend the study, but it's also for that reason, I think that the more loosely you set your criteria for the initial group selection, the more valid the selection criteria are, rather than the other way around. If you set up your criteria very strictly, you're very likely to include the people you want. In this case we just used the clinical diagnoses of what turned out to be Rosenbek, Collins, and Wertz, based upon a battery of tests. There were no defined criteria from any of the tests that they used to detect the presence of aphasia. The assessment procedures utilized many tasks and tests. So philosophically, I think it's more ecologically valid to set loose criteria for the initial group selection, and then see how they fall out. It certainly isn't dependent on severity. In fact we did some other analyses in a subsequent study to this one to look at the relationship between classifiability and severity, and it doesn't work very well. However, philosophically I agree with you that it's a continuum. I don't think there are any aphasic and nonaphasic individuals, but there are points at which the decision regarding the presence of aphasia has clinical importance.

Q. Part of my concern is, what do you do with this after you're done with it? You use a system and you take someone off the street and give them the measures, and you do the discriminant function, and you say this guy is or is not aphasic? Why do we need it?

A. My opinion is a little different from those reflected by the preceding comments. But I wouldn't use the discriminant function scores clinically. I don't think they mean anything at all, clinically. But I might use it for subject selection in research. That's where I would use it. Right now there's no way to determine if your subjects are in fact normal. We just say they were normal because we said they were normal. I would use it for that purpose but I wouldn't use it clinically. If a guy comes in and he's having trouble, frankly I don't care if you call him aphasic or not. I'll probably try and work on whatever I see

and identify as a problem. But for research purposes I would use it. If I felt comfortable enough with the data, I would use it for that purpose.

C. But if one's going to use it for research, it seems to me you've got to be tremendously careful about what went into your original decision that this group was the aphasic group upon which the discriminant function was based. You've really got to define that carefully.

A. No, I would say loosely. You definitely have to define aphasia, but it's definitely got to be done loosely. I believe that Terry, Jay, and Mike have a conceptual framework for what aphasia is, based probably on Darley's definition of aphasia. They use, to define its presence, the cardinal features such as crosses all modalities, and so on. That's as specific as I would get for this kind of a study. Anything more specific would be too restrictive and reduce the generalizability of the results.

Q. What I thought Bob was going to do, but he's too nice a guy to have done it, is to ask the question, what was your criterion? What were your criteria for saying these folks were aphasic to begin with, and once you told him he'd say, then why did you need to do the discriminant analysis? I think I agree with Mick, in that in so many studies we say we have *X* number of aphasic folks, and it would be nice to say these folks were aphasic as demonstrated by the PICA. That could be applied universally. And as you suggest, through replication using discriminant function analysis we get a group of people we think are aphasic, force them through a discriminant analysis, and we get a high percentage of classification in our shop; you do it in Wisconsin and Bob does it in Minnesota, and we all find that the test classifies people based on whatever our independent assumptions we have about who's aphasic and who's not. Then you can do away with the independent assumptions about who is aphasic and who is not, and you eliminate a lot of quibbling about whether you really have aphasic people in the study or not. It's a question of validity of the diagnosis.

A. I think you can use this clinically as well. An analogy might be the neurologist who sees someone who presents a history and a clinical exam that suggests that they have ALS. And then they order an EMG to confirm the presence of the disease. And the EMG comes back and it says there's evidence of diffuse motor neuron disease, and that confirms what they have said and lends some objectivity to that clinical judgment. Or, perhaps, it comes back and says the findings are equivocal or do not clearly indicate the presence of this disease. And what the neurologist usually does is say, "Well yeah but he's got the disease anyway and the findings are just not prominent enough yet for us to know," or he may say, "Well maybe I'm wrong" and withhold that diagnosis. But it lends some objectivity to what we do. And

I see patients frequently where they're sent over without any definable lesion in which the question is, is there something going on in this person's nervous system? Are they aphasic or aren't they? And I evaluate the patient in whatever way I evaluate them, and I say I think they are aphasic or I think they're not, but that judgment could be objectified with the results of information like this that may lend support to a clinical decision you make, or that may refute it. But I think there are times when this might be a good objective piece of evidence that would really help in a clinical situation.

Q. Procedural question: When you applied the analysis, did you use prior probabilities proportional to your sample size? You didn't have an equal number of aphasic subjects and normal subjects. Did you weight the analysis so that that would be taken into account?

A. We used the BMDP software package, and I'm not sure how it handles that.

C. You can set proportions. If you don't set that kind of variable, it assumes that it's a 50-50 split.

A. We had it assume that it was a 50-50 split.

C. OK. But what I'm saying is that if you let the program know that you are expecting 60 percent of the people to be normal and 40 percent of them to be aphasic, then the results could be different. That option is available on some of these software packages. The other question I had was whether you had any concerns about the cross-validation being done on 15 percent of only 80 aphasic patients. I think with the large numbers that you had you'd be safe to do your original analysis on say 60 percent and cross-validate with 40 percent, and feel more comfortable with that before passing out the weights.

A. A higher subject pool in the cross-validation would have been helpful, but then there would have been fewer subjects used in generating the discriminant function. So there's a trade-off.