Measures Of PICA Subtest Variance: A Preliminary Assessment Of Their Value As Predictors Of Language Recovery In Aphasic Patients

James L. Aten
Jon G. Lyon

Veterans Administration Hospital
Long Beach, California

Reliable measures of predicting language recovery are needed if aphasiologists are to be accountable in their language rehabilitative roles. Attaining reliability in predicting recovery has not been easy given the inherent complexity of the task. Initial efforts to accomplish this task have focused primarily on patient parameters (Eisenson, 1949; Wepman, 1951). Darley (1972) listed 10 patient parameters that appear to influence language recovery. Some of these parameters, such as premorbid language level, social status, and even extent of brain damage are difficult to quantify accurately. Post-morbid residual language, however, can be measured objectively.

Porch (1967, 1971) has attempted to utilize patient test scores from the PICA to predict language recovery. He has plotted overall percentile change in aphasic patients between the first and sixth month post onset (MPO), thus deriving a set of predictive curves called HOAP slopes. Also Porch has used a percentile difference between a mean of the nine highest and nine lowest subtest scores, a "high-low gap," to predict change. In 1974, Porch, Collins, and Wertz used an analysis of covariance to assign predictive "weights" to different groupings of PICA scores at 1, 3, and 6 MPO. The adjusted predictive values were then correlated with attained values. Their results yielded high correlations for all test intervals, with accuracy of prediction being directly correlated with length of time post onset. Porch (1978) has introduced still another measure, peak-mean difference (PMD), an accumulating score over subtests using the difference between the highest score and mean for each subtest. This measure is based on the premise that the response receiving the highest score within a subtest is the best indicator of potential change. It remains unconfirmed whether PMD is sensitive to later changes in subtests scores, and whether it is superior to the more classical measures of variance, such as standard deviation or range.

The purpose of this investigation was to determine how five measures of variability on PICA subtests correlate with subsequent language gains in a group of aphasic patients undergoing language therapy during and following the classical spontaneous recovery period.

METHODS

Subjects. The records of 72 aphasic patients at the Long Beach Veterans Administration Hospital were examined to determine the regularity with which PICA tests were administered over the past four years. Patients who had been tested on the PICA between the first and fourth MPO, and between the sixth and ninth MPO, were included in the test sample.

Twenty-four male aphasic patients with a mean age of 53 years and a range of 27 to 68 years satisfied the above criteria. The mean elapsed time for the initial testing of patients was 2.1 months while their overall PICA percentile averaged 42.5. For the second testing, the mean MPO was

7.8 months and the overall percentile mean was 63. Of the 24 patients included in the test data for the first two testings, 13 had been given PICAs between the 10th and 13th MPO as well, the mean being 11.6 and the mean overall percentile score being 67.5.

All patients exhibited unilateral, left hemisphere brain damage; 22 attributable to CVA, 1 due to trauma and 1 due to infectious disease. Of those patients suffering a CVA, 17 were thromboembolic and 5 were hemorrhagic. Eighteen of the 24 were classified independently by a panel of three judges as being predominantly nonfluent, the others as fluent. Throughout the testing periods, all were seen 3 to 4 times per week in individual therapy and approximately half were seen biweekly in group therapy.

Measures of Variance. The five measures of variance examined within PICA subtests were: standard deviation, range, high-low gap[1], peak-mean difference, and a modified peak-mean difference[2].

Procedures. PICA tests were examined over the three post-onset periods. Each of the five measures of variance was determined for each subtest. In addition to subtest variances, modality and overall variances were determined by averaging appropriate subtest variances together. Language gains were calculated by subtracting subtest scores achieved on the initial test from those on the second and third testings. Again, means of these subtest score changes were used to determine modality and overall changes.

Analysis of the Data. Multiple Pearson Product-Moment Correlations were performed between all measures of variance at approximately two months post-onset and subsequent changes in scores at 6-9 MPO and 10-13 MPO.

RESULTS

There were minimal differences between the five measures of variance. Of these variance measures, though, PMD produced the highest positive correlations.

Figure 1 displays the difference between correlations obtained using PMD and each of the other variance measures. The correlations are between subtest variance and overall change in PICA scores. The top graph illustrates that standard deviation correlations were below PMD correlations across all subtests. Range and high-low gap correlations were below PMD correlations on a majority of subtests as well. Modified PMD correlations differed the least. Correlation differences among variance measures were not significant, with PMD revealing the highest correlations. Consequently, the remaining data analyses were restricted to PMD.

Our initial aim was to determine which PMD measure produced the highest correlation with overall change in subtest scores. Table 1 shows how group PMD measures on the first testing correlated wtih overall change from 1st to 2nd, and from 1st to 3rd testings. Overall PMD failed to correlate with overall change at either interval. An accumulative PMD for the gestural

---

[1]Mean of the five highest scores minus the mean of the five lowest.

[2]When the highest score within a subtest occurred but once, an average of the two highest scores was taken and then the mean was subtracted from it.

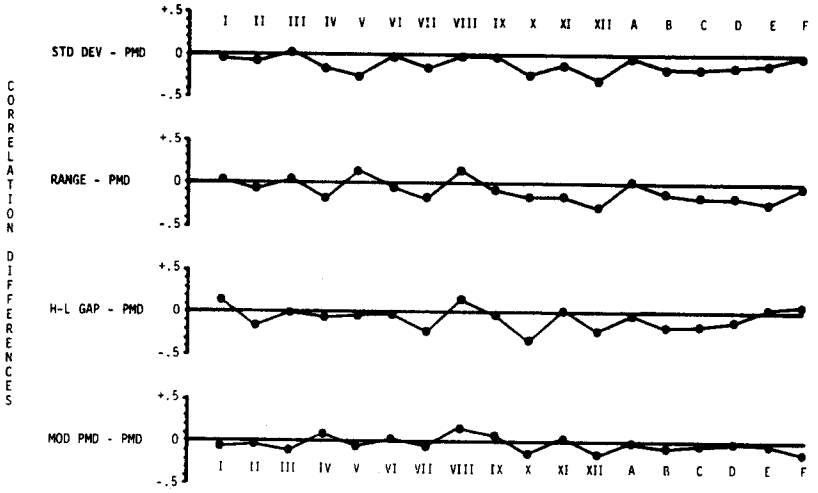

Figure 1. Correlation Differences Between PMD and Other Variance Measures.

modality subtests produced the highest positive correlation of .577 when examining change at 6-9 months, and .701 when examining change at about 1 year. The subgroup of auditory subtests (VI and X) yielded similar positive correlations. Visual, reading and gestural subgroups were positively correlated, although substantially reduced. These gestural modality PMD measures, although positively correlated to overall change, are of minimal predictive value, the square of their correlation reflecting that they are less than 50% accurate. As for the remaining correlations, PMD groupings of verbal and graphic subtests correlated inversely with overall change, achieving a value of -.518 on a combined measure of subtests A and B. This latter finding suggests the erroneous conclusion that minimal variance on verbal and graphic measures is a good sign of later language recovery—a relationship negated by the severly involved aphasic patient.

To explain better the diversity in correlations, we prepared scatter plots for those PMD measures yielding the highest positive and negative correlations. Figure 2 displays PMD on gestural modality subtests

Table 1. Correlations Of PMD And Overall Mean Change On PICA

PMD	Overall Mean Change	
	1st - 2nd Testing (N = 24)	1st - 3rd Testing (N = 13)
Overall	.023	.021
Gestural	.577	.701
Auditory (VI, X)	.566	.662
Visual (VII, XI)	.327	.457
Reading (V, VII)	.265	.159
Gestural (II, III)	.180	.358
Verbal	-.128	-.240
Spontaneous (I, IV)	-.098	-.080
Cued (IX, XII)	-.118	-.316
Graphic	-.320	-.338
Spontaneous (A, B)	-.518	-.580
Copying (E, F)	-.021	-.013

as it relates to overall change between first and second testings. The circled scores represent the thirteen patients who were included in the test sample at approximately 1 year. Similar distributions of circles and triangles throughout the scatter plot suggest that the increase in the correlation from .577 to .701 was legitimate and not merely a reflection of the reduced number of subjects at third testing. Subtest scores are well distributed along a positive slope with a slight predominance toward the lower ranges of each variable.

Figure 3 shows the scatter plot for the inverse relationship between PMD on a combined measure of subtest A and B and overall change. Scores are not equally distributed, but rather reflect low PMD values with proportionately greater overall score change. Patients found the harder graphic subtests difficult at 2 MPO yet their overall test scores changed significantly. On the other hand, there were four patients who exhibited large PMDs, yet showed little change in overall scores. These patients attained high PICA scores on the initial testing, which limited the amount of score change possible. Yet they still showed varied ability to complete successfully the harder subtests.

Finally, we examined individual subtest variance and overall change. These correlations are plotted by subtest in Figure 4. Note that the subtests are ordered according to the PICA's ranked response summary sheet. PMD on the more difficult subtests of the PICA resulted in negative correlations while PMD on the easier subtests produced positive correlations. Our initial thought was that lack of variance on the harder subtests might explain this finding. But a closer look at mean variance, shown in the upper half of this figure, fails to support this position. Many of the

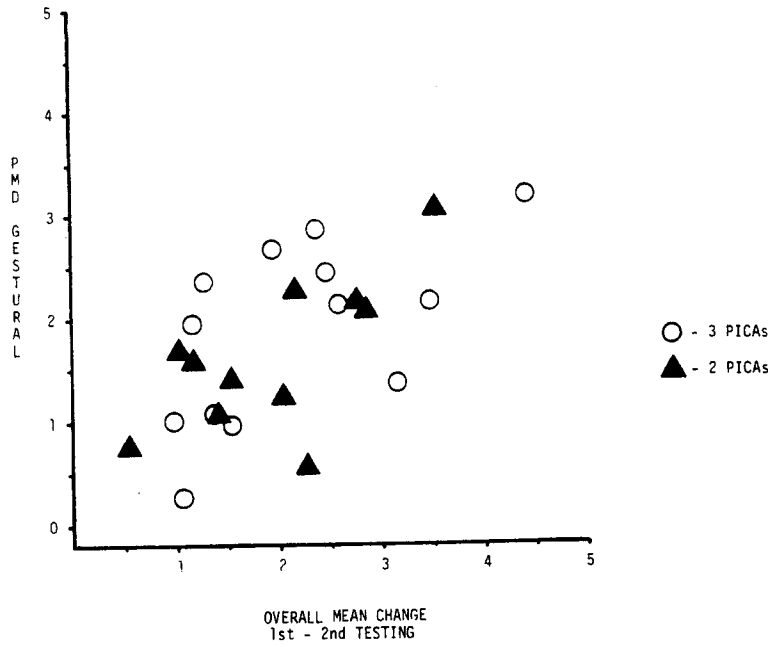

Figure 2. A Scatter Plot of PMD-Gestural And Overall Mean Change On PICA Subtests.

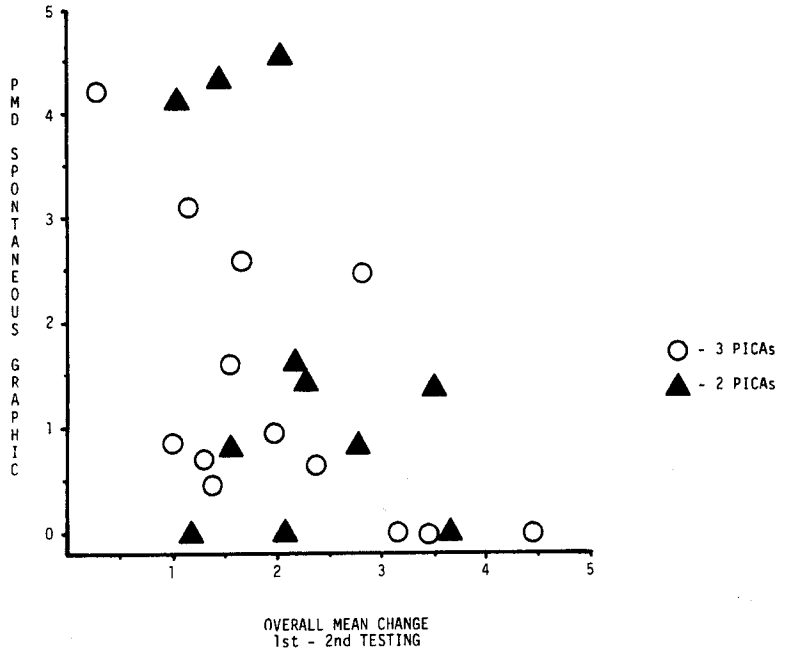

Figure 3. A Scatter Plot Of PMD-Spontaneous Graphic And Overall Mean Change on PICA Subtests.

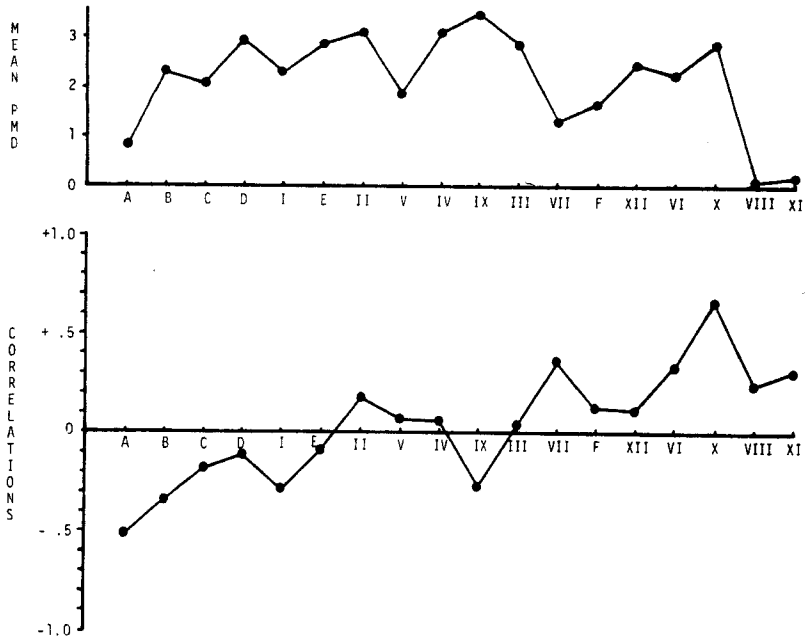

Figure 4. Mean Subtest PMD And Correlations Of Subtest PMD With Subtest Change.

difficult subtests exhibit average or above average PMD levels. Evidently, test scores on hard subtests do vary, but not in accordance with overall test score change. Variance on easier subtests, on the other hand, does relate directly with overall change, the best example being subtest X (.672). Still, none of these correlations could serve as a reliable predictor of overall change.

If PMD measures are unable to predict accurately overall change in this format, perhaps they would be more effective in predicting change on gestural, verbal, or graphic modality subtests. Table 2 shows the highest PMD correlate for test changes in each of the output modalities. In most cases, the best indicator of score change is variance on that same measure. For instance, gestural modality change correlates highest with gestural modality variance. The same trend holds true for auditory and visual subtest change. Even when the highest correlate was not PMD for the same PICA group measure, it often was PMD of a subtest within that group. For instance, PMD on subtest VII correlated highest with reading change from 1st to 3rd testings. Correlations were significantly higher when using these output modality measures than when using overall change. Gestural categories produced the highest correlations. But with the exception of visual subtests, the predictive value of these scores is questionable.

Next we examined subtest change and subtest PMD. These correlations are shown, as before, in the lower half of Figure 5 and arranged according to subtest difficulty. Correlation values did increase over earlier values, supporting the notion that variance within a measure best predicts change on that same measure. Mean PMD values did not vary with subtest correlations. Of the graphic subtests, D produced the highest (.425), for the verbal subtests, XII had the highest correlation (.384),

as before, the highest values occurred on the gestural subtests XI - .991, VIII - .877, X - .875 and VII - .672. The high correlations on VIII and XI reflect minimal variance and subsequent minimal change and should be interpreted as such.

To summarize these data, the five measures of variance incorporated here yielded similar correlations, the highest correlation values being attained with PMD. None of the PMD measures proved to be a sensitive indicator of overall change. Although higher correlations were found when PMD was correlated with score change on the same PICA measures, their predictive value was confined to visual and auditory subtests and they were of no value in judging change on verbal or graphic measures.

Table 2. Highest PMD Correlate For Groupings Of Gestural Verbal and Graphic PICA Subtests.

	Correlate			
PICA Subtests	1st - 2nd Testing		1st - 3rd Testing	
Gestural	Gestural	(+.699)	Gestural	(+.775)
Auditory	Auditory	(+.764)	X	(+.943)
Visual	Visual	(+.914)	Visual	(+.921)
Reading	Reading	(+.539)	VII	(+.741)
Gestural	III	(+.534)	VIII	(+.588)
Verbal	Verbal-cued	(+.471)	II	(+.438)
Spontaneous	Verbal-cued	(+.476)	IX	(+.665)
Cued	X	(+.511)	Auditory	(+.659)
Graphic	XI	(+.320)	XI	(+.444)
Spontaneous	D	(+.474)	D	(+.548)
Copying	X	(+.620)	X	(+.682)

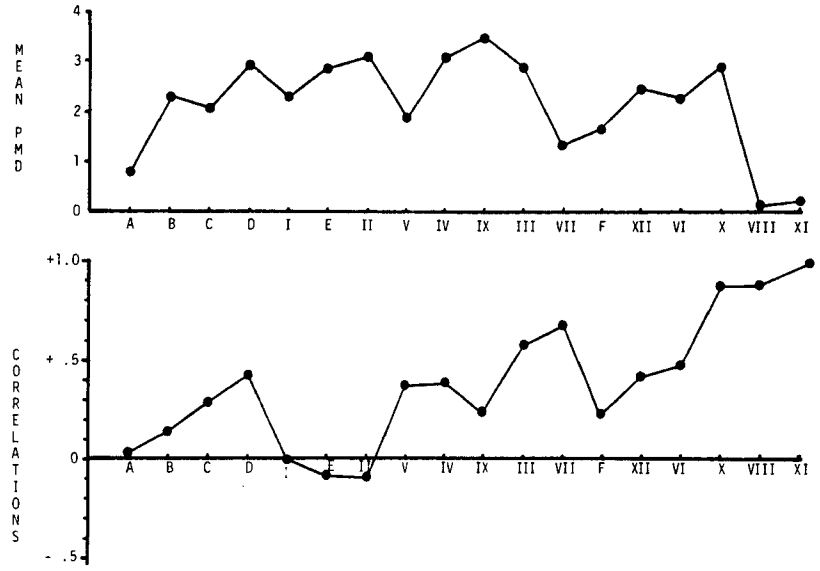

Figure 5. Mean Subtest PMD And Correlations Of Subtest PMD With Overall Mean Change.

DISCUSSION

The minimal relatedness between certain measures of subtest variance and change in test scores suggests that varied patient performance within a task does not predict later gains. Before accepting such a conclusion, several questions need to be examined within the concept of variance itself: (1) Are variance measures of value in other than a correlative sense? (2) Would variance measures be more sensitive within subgroups of the patient sample? and (3) Is the sensitivity of variance measures limited by the test itself?

Porch (1978) has suggested a more general use of PMD. He has established a cut-off level for this measure, maintaining that PMD scores below this level indicate that the patient has reached his maximal potential. All the patients in this study had PMD scores above the cut-off level and all did show test score improvement, supporting Porch's position. Yet variation in PMD above the cut-off level did not correspond to variation in test score change.

If change had been predicted by variance, then there would be reason to believe that the patient parameters mentioned by Darley (1972) were accounted for in the variance measures. Since this was not the case, an in-depth analysis of these parameters is warranted. Such factors as initial severity of aphasia, type of aphasia, and degree and type of therapy need to be examined individually.

There may be factors within the PICA which also limit the sensitivity of these variance measures. The primary purpose of the PICA is to provide a psychometric measure which is sensitive to patient change on specific communicative tasks (McNeil, in press). The development and assessment of communicative tasks which validly reflect variance within such tasks is a different purpose. The following factors, in reference to the present PICA format, need to be considered.

Establishment of a Dynamic Test Range. Variance on tests must be attainable. Tasks where optimal scores are achieved with ease, or just the opposite, tasks that are too difficult to do, negate variance-score change relationships.

Inclusion of Homogeneous Test Items According to Salient Features Within Each Subtest. Homogeneity on the PICA means that test items are common to the experience of adults of both sexes, capable of being demonstrated gesturally, and approximately equal in difficulty across all tests (Porch, 1967). Homogeneity as it applies to subject variance on a task must satisfy the condition that all test items are of equal difficulty for the task being measured. It is hard to imagine that matches, toothbrush, and cigarette complement pen, key, and comb on verbal repetition or spelling by name subtests.

Uniformity in Scaling Patient Responses. The issue of whether PICA scores can be treated statistically has been raised elsewhere (Silverman, 1974; McNeil, Prescott and Chang, 1975; Martin, 1977; Duffy, 1977). The basis of this debate hinges on the fact that PICA scores are not interval scores, yet are treated as such. A score change from 6 (error) to 7 (related response) on Subtest IV may represent a larger functional and communicative gain than a change from 8 (cued) to 9 (repeat). Unequal interval score comparisons also occur across subtests.

A Measure of Patient Potential on Tasks. The initial instructions on the PICA were devised to provide the patient with the amount of stimulation necessary for a normal person to do the task. Based upon a cybernetic theory of brain function, such a procedure may be well justified. However, to validly assess variance within subtests, we need to know the full range of the patient's capabilities, thus arguing for demonstration items.

To conclude, we have presented some factors above that may have reduced the significance of variance as a direct indicator of language change. The fact remains, though, that such variance measures, even with these modifications, may not conform to the basic premise that one directly predicts the other.

REFERENCES

Darley, F. The efficacy of language rehabilitation in aphasia. JSHD 37,1: 3-20 (1972).

Duffy, J.R. and Dale, B.J. The PICA scoring scale: Do its statistical shortcomings cause clinical problems? In: Clinical Aphasiology Conference Proceedings 1977, R.H. Brookshire (Ed.), Minneapolis, MN, BRK Publishers (1977).

Eisenson, J. Prognostic factors related to language rehabilitation in aphasic patients. JSHD 14:262-264 (1949).

Martin ,A.D. Aphasia testing: A second look at the Porch Index of Communicative Ability. JSHD 42,4:547-562 (1977).

McNeil, M.R. Porch Index of Communicative Ability (PICA). In: Evaluation of Assessment Techniques in Speech Pathology, F.L. Darley (Ed.), Reading, MS, Addison-Wesley Publishing Company (In Press).

Porch, B.E. Porch Index of Communicative Ability: Theory and Development, Vol I. Palo Alto, Consulting Psychologist Press (1967).

Porch, B.E. Porch Index of Communicative Ability: Administration, Scoring and Interpretation , Vol II. Palo Alto, Consulting Psychologist Press (1967).

Porch, B.E. Multidimensional scoring in aphasia testing. JSHD 14,4:776-792 (1971).

Porch, B.E. Diagnosis of Aphasia and Brain Dysfunction. Albuquerque, NM, VA Satillite Television Presentation (1978).

Porch, B.E., Collins, M.J. and Wertz, R.T. Statistical and clinical procedures for predicting recovery from aphasia. In: Clinical Aphasiology Conference 1974, 27-37 (1974).

Silverman, F.H. The Porch Index of Communicative Ability (PICA); A psychometric problem and its solution. JSHD 39,2:225-226 (1974).

Wepman, J.M. Recovery from Aphasia, New York, Ronald Press (1951).

DISCUSSION

Q. Did you use total PMD when correlating it with change over time? We completed a similar study last winter where we looked at patients early post onset, 1 to 2 months, 2 to 6 months and beyond 6 months. We correlated PMD, that is total PMD, with change and found in those 33 patients that the amount of time was inversely related to the PMD score. PMD starts high, as I pointed out this morning, and the overall starts low and then the pattern changes over time. And these were all significant. So maybe we're looking at the data in different ways.

I don't know if you get a washout effect by using variance instead of totals or by neglecting to differentiate between earlier stages and greater variance within months post onset.

A. I'm not sure either. However, the issue of using different ways of tabulating an overall PMD does not appear valid. By multiplying a constant of 180 to the overall PMD values used in this study, you obtain the identical overall PMD values used in the study cited above. Multiplying values in a correlation by a constant does not alter the correlation score. Did you look at percentiles in your study?

Q. No. What did you find?

A. Another way of looking at change, beside test score difference, would be to look at the difference in percentiles. I think such a measure is more valid for the overall, verbal, gestural and graphic test scores where every percentile value can be translated into a given test score. If you go to Appendix C in the PICA scoring manual, there are some subtest scores that represent a percentile range rather than a single percentile. But what we did find when we correlated overall PMD with overall percentile change was a .398 value compared to .023 when using test scores. The other modality percentiles correlated within a tenth of the correlations obtained when using test scores, thus no major difference. Why the overall percentile correlation differed from the overall test score is unclear.

Q. What was your argument for demonstration items?

A. First, I would like to make it clear that I'm not trying to argue against the PICA format as it stands as a diagnostic measure. The issue I'm trying to address here is the best way of getting valid measures of subtest variance. If that is your goal, then it seems to me that you need to know the patient's potential or range of performance. Dr. Porch's philosophy has been to provide the patient with the stimuli needed for a "normal" person to do the task. I feel it would be of value to determine if a patient could do a gestural task if you put an object in a patient's hand, assisted him through the desired response and then came back to the object moments later. If you score this behavior, you should get a different range.

Q. That sounds to me like you would be getting a greater effect at the upper ends of the response continuum than the lower. Is that what you want?

A. I don't know if that would be better or worse. If prognosis for change rests on what the patient is capable of doing, it might give us a better picture. I don't know, but I'd like to see its effect.

Q. There is some evidence that the prestimulation makes things worse.

A. No response.

Q. The question has been asked why demonstration items are not included in the PICA. The point is that normals don't need demonstration items and aphasic patients do. The second point is that the question of homogeneity of test items keeps coming up on the PICA. And people keep assuming that key and comb are easier than toothbrush and

cigarette and that's not the case. That's an a priori assumption, but you have to ask the patient's system whether that is true or not. With some patients cigarette and toothbrush are much easier than key and comb. If they have more trouble turning on their auditory system, then the longer words are better than the shorter. So if you take all those items across the board for all aphasic patients they are of equal difficulty. Darley and Disimoni have established this.

A. I'm not concerned with differences across subtests but rather within subtests.