CHAPTER

12

Does SPICA Mean PICA?

Pat Holtzapple
Karen Pohlman
Leonard L. LaPointe
Lisa Fletcher Graham

It seems as we approach the end of this decade, our attention is shifting from concerns of "does it work?" to "how long does it take?" A particularly captivating concern is that of cost-effectiveness. A discussion has ensued over the desirability of using a shortened form of the Porch Index of Communicative Ability (PICA) to address some of these issues.

The Porch Index of Communicative Ability, or PICA (Porch, 1967), has been variously shortened into three formats, reported by several authors. DiSimoni, Keith, Holt, and Darley (1975) and DiSimoni, Keith, and Darley (1980) present two short forms: one with four subtests, 10 items; the other with 17 subtests, two items. Phillips and Halpin (1978) and Lincoln and Ells (1980) proposed administering the full 18 subtests with only five items. The hallmark of these versions is a substantial reduction in administration time.

To create the first shortened PICAs in 1975, DiSimoni and colleagues performed a stepwise regression on 222 full PICA administrations, which yielded the two short forms previously mentioned: (1) four subtests, 10 items; (2) 17 subtests, two items. Conversion formulas were derived for these short forms by "extracting" various subtest mean scores. These extracted mean scores were then submitted to various mathematical computations until a combination emerged that would result in an overall score equal to that of a fully administered PICA.

In a follow-up study, DiSimoni and colleagues (1980) administered the full PICA and both short PICA (SPICA) forms to 93 aphasic subjects. The authors reported no significant difference between PICA and SPICA scores, concluding that both versions could predict PICA overall mean scores at acceptable confidence levels. Individual subject results were not provided in these studies.

We have used the DiSimoni and colleagues' (1980) four-subtest SPICA in our clinic for many years. The four-subtest SPICA uses all 10 items from the PICA and only subtests I, VI, VII, and D, which are described in Table 12-1. Ideally, a SPICA score should mirror a PICA score if all conditions are held equal. In the clinic, we assign significance to a test-retest score that falls outside an overall percentile range of ± 5 percentile units,

TABLE 12-1. DESCRIPTION OF SPICA

Items	*Subtests*
10 PICA objects	I — Verbal: describe function
	VI — Auditory: point to function
	VII — Reading: direction + noun
	X — Writing: spelled dictation

as originally suggested by Porch; that is, variability beyond ±5 percentile units would indicate a clinically relevant difference. Disparate overall scores between the shortened and complete PICAs, despite the relative absence of change on other measures, have led us to judicious use and interpretation of the SPICA. Our clinical experience has not substantiated the conclusion by DiSimoni and his colleagues (1980) that the SPICA is an accurate estimate of the PICA.

STATEMENT OF THE PROBLEM

Can we shorten time required for testing? is not the only question. The critical issues are reliability of overall scores, diagnostic and treatment information gained or lost in the process, and the necessity of maintaining the quality of what we do. Can we maximize the gains made from two decades of work with a diagnostic measure and modify it while still being able to apply established norms?

The purpose of this study was to report our own findings on comparison of the PICA and the four-subtest-corrected SPICA by adult aphasic subjects and to examine the risks versus benefits of such a practice.

We posed the following questions:

1. Do our data support the DiSimoni and colleagues (1980) finding that there was no significant difference between PICA and SPICA mean scores?
2. How many subjects will fall within a variability range of ±5 percentile units of compared PICA and SPICA overall percentile scores?
3. Does the difference in PICA and SPICA scores vary relative to severity?

METHODS AND PROCEDURES

Both the PICA and the four-subtest SPICA were administered to 10 aphasic adults in counterbalanced order. All tests were administered on the same day with the exception of three, which were administered within a 48-hour period. All tests were videotaped and double-scored by PICA trained clinicians. Agreement was reached on all responses.

We obtained three overall mean scores from administering the two tests. First, the PICA overall mean was obtained in the usual fashion. Second, the SPICA overall mean was obtained by administering the four sub-

tests mentioned above and submitting the four mean scores to the DiSimoni and colleagues (1980) conversion formula:

$$2.03 + .21(\bar{X}_{I}) + .21(\bar{X}_{VI}) + .19(\bar{X}_{VII}) + .23(\bar{X}_{D})$$

These two overall means (PICA and SPICA) were used in our comparative analyses. A third score, a SPICA "extract" overall mean, was obtained so that we could judge the reliability of comparing an administered SPICA to one merely pulled from a previously obtained PICA. These "extracts" were derived by using the four necessary subtest means from the fully administered PICA and applying the SPICA conversion formula.

Demographic data are shown in Table 12-2. Subjects ranged in age from 57 to 88 years ($\bar{X}$ = 68 years) and ranged in months post-onset from 1

TABLE 12-2. SUBJECT DESCRIPTIVE DATA

Subject	*Months post-onset*	*Age (years)*
1	6	68
2	3	58
3	9	60
4	33	88
5	14	59
6	6	67
7	21	70
8	1	80
9	58	67
10	0.50	76
11	0.75	63
12	0.75	59
13	0.25	68
14	96	57
15	0.25	68
16	30	64
17	0.30	69
18	0.20	81
19	102	66
Range	0.20–102	57–88
Mean	20	68

week to 8½ years ($\bar{X}$ = 20 months). All subjects had a history of left-hemisphere cerebrovascular accident.

RESULTS

Table 12-3 presents group data for the PICA and SPICA overall mean and percentile scores for all subjects. Individual PICA and SPICA mean and percentile scores are shown in Table 12-4. To determine whether the obtained mean differences were statistically significant, an analysis of variance with repeated measures was completed. Significant differences were found between the PICA and SPICA overall mean scores ($F = [1, 18] = 25.57$, $p < .001$). Correlation statistics were computed using Pearson's r correlation coefficient. A strong positive relationship was found between group PICA and SPICA overall mean scores ($r = .98$, $p < .001$). No statistically significant difference in order effect was observed.

Visual inspection of the raw data shown in Table 12-4 reveals that 17 of 19 subjects, or 89 percent, scored higher on the SPICA than on the corresponding PICA. Comparison of group means reveals an approximate 5-percent group advantage for the SPICA, which translates to a 5 percentile unit increase, 44 percentile units in the PICA and 49 percentile units on the SPICA.

Table 12-5 depicts the difference between PICA and SPICA percentile scores for each subject, which ranges from −1 to +17 percentile units. For the group, subjects scored 5 percentile units higher on the SPICA, thus falling within a ±5 percentile unit guideline. For individual performance, 10 subjects, or 53 percent, fell within that guideline. Conversely, a full 47 percent failed to fall within our ±5 percentile unit guideline of acceptable variability.

To address our third question, subjects were assigned to one of three severity levels (0–30 percentile, 31–60 percentile, 61–90 percentile) based

TABLE 12-3. GROUP DATA ON PICA AND SPICA OVERALL MEAN AND PERCENTILE SCORES

Measurement	*Range*	*Mean*	*SD*
PICA mean	3.97–12.82	10.03	2.42
PICA percentile	3–78	44	
SPICA mean	5.51–13.64	10.62	2.33
SPICA percentile	9–89	49	

TABLE 12-4. PICA AND SPICA MEAN AND PERCENTILE SCORES

	PICA		*SPICA*	
Subject	*Overall mean*	*Percentile*	*Overall mean*	*Percentile*
5	3.97	3	5.51	9
2	5.62	9	6.33	13
3	6.63	15	7.06	18
6	7.98	26	8.54	31
8	9.32	38	9.39	39
1	9.64	41	10.92	53
4	10.07	45	10.18	46
16	10.23	46	11.09	55
14	10.33	47	10.24	46
19	10.36	47	10.45	48
15	10.60	50	11.41	59
10	11.38	59	11.45	60
17	11.41	59	12.48	74
13	11.74	64	13.06	81
9	11.81	65	12.45	74
18	11.92	67	12.00	68
11	12.29	71	13.06	81
7	12.48	74	12.43	73
12	12.82	78	13.64	89
Range	3.97–12.82	3–78	5.51–13.64	9–89
Grand mean	10.03	44	10.62	49

on their PICA percentile score. An analysis of variance with repeated measures was used to determine statistical significance on PICA and SPICA percentile scores relative to severity of aphasia. Statistically significant differences were found for each group. The F ratios, mean difference, and standard deviation of the differences for each group are displayed in Table 12-6.

DISCUSSION

Do our data support the DiSimoni and colleagues' (1980) finding that there was no statistically significant difference between PICA and SPICA

TABLE 12-5. DIFFERENCE BETWEEN PICA AND SPICA PERCENTILE SCORES

Subject	*PICA percentile*	*SPICA percentile*	*SPICA/PICA difference*
5	3	9	6
2	9	13	4
3	15	18	3
6	26	31	5
8	38	39	1
1	41	53	12
4	45	46	1
16	46	55	9
14	47	46	−1
19	47	48	1
15	50	59	9
10	59	60	1
17	59	74	15
13	64	81	17
9	65	74	9
18	67	68	1
11	71	81	10
7	74	73	−1
12	78	89	11
Range	3–78	9–89	−1 to +17

TABLE 12-6. ANALYSIS OF VARIANCE WITH REPEATED MEASURES FOR PICA AND SPICA PERCENTILE SCORES RELATIVE TO SEVERITY

Rating	*df*	*F ratio*	*Alpha*	$\bar{X}$ *Difference*	*SD*
Severe	1,3	47.88	<.01	4.50	1.29
Moderate	1,8	7.29	<.05	5.33	5.92
Mild	1,5	8.17	<.05	7.83	6.71

mean scores? Our answer to question one is no. In our study, a statistically significant difference between PICA and SPICA mean scores was observed.

The answer to our third question, "Does the difference in PICA and SPICA scores vary relative to severity?" is also no. When we separated the

groups by severity level, a statistically significant difference between PICA and SPICA percentile scores for each severity level was found.

In response to the question, "How many subjects will fall within a variability range of ±5 percentile units of compared PICA and SPICA overall percentile scores?" our answer is not enough! Remember, the number of subjects whose overall percentiles fell within a ±5 percentile range represented only 10 of the 19 subjects, or 53 percent. This result parallels our clinical experience, which has been that often a difference between PICA and SPICA scores would occur in the absence of any other noticeable change. Using the ±5 percentile unit standard, nine, or 47 percent, of our subjects' scores surpassed a clinically relevant point. At obvious issue here is the change in score one deems clinically significant.

When these discrepancies were observed in the clinic, we considered the possibility of extracting a SPICA from a previously administered PICA. To accomplish this, we took the mean scores of subtests, I, VI, VII, and D from the PICA, applied the correction formula, then obtained an overall mean score, so that we could compare a SPICA extract with the administered SPICA. Our desire was to compare the same subtests; Table 12-7 depicts those comparisons. Seven of the 19 subjects exceeded ±5 percentile units of variability. As is evident in the scores, some subjects mirrored their score, while one subject was off by a troublesome 24 percentile units. An analysis of variance with repeated measures applied to these scores also yielded statistically significant differences ($F\ [1, 18] = 6.20$, $p < .05$).

In addition, when an analysis of variance with repeated measures is applied to the extracted SPICA scores versus the full PICA, a statistically significant difference obtains ($F\ [1, 18] = 4.79$, $p < .05$). See Table 12-8 for subject scores. This result calls into question the conversion formula itself.

CONCLUSION

The purpose of the SPICA was originally to create a shortened test whose overall score could quickly and reliably predict that of a more familiar, comprehensive diagnostic tool. The authors of those original papers made no recommendations as to its use, its relative advantages or disadvantages, or even its overall desirability. However, their conclusions of no significant difference would seem to encourage PICA and SPICA comparisons. Our clinical observations, supported by the significant difference on all measures in this study, suggest that interchanging PICA-SPICA scores may be a dangerous practice. We stongly caution against the assumption that, for any given individual subject, a SPICA score would be equivalent to the same subject's performance on the PICA.

TABLE 12-7. DIFFERENCE BETWEEN SPICA AND SPICA EXTRACT PERCENTILE SCORES

Subject	*SPICA*	*SPICA extract*	*SPICA/SPICA extract difference*
5	9	9	0
2	13	12	+ 1
3	18	20	− 2
6	31	29	+ 2
8	39	40	− 1
1	53	43	+10
4	46	53	− 7
16	55	51	+ 4
14	46	48	− 2
19	48	48	0
15	59	47	+12
10	60	57	+ 3
17	74	64	+10
13	81	78	+ 3
9	74	50	+24
18	68	70	− 2
11	81	68	+13
7	73	72	+ 1
12	89	78	+11

Further study is obviously needed to address several issues, including the following: (1) Would a replication of this study with a larger number of subjects have resulted in the same outcome? (2) Might some correction in the conversion formula result in a closer match with the PICA overall? (3) Does the SPICA provide test-retest stability that would merit its use independent of the PICA?

Benefits gained from the use of a full diagnostic examination such as the PICA range from its focus across diverse modalities to its treatment implications and clinical familiarity. An obvious drawback is the time required for administration, if time is one's primary concern. As we indicated at the beginning of this chapter, we frequently use the SPICA. For screening purposes, we administer a SPICA followed by subtests IV (naming) and XII (repetition), which we feel gives us an adequate check of relevant modalities when time permits no more. If circumstances dictate that a screening is all we get in a quickly recovering patient, then we readminister the SPICA and compare SPICA to SPICA.

TABLE 12-8. DIFFERENCE BETWEEN PICA AND SPICA EXTRACT PERCENTILE SCORES

Subject	*PICA*	*SPICA extract*	*Difference*
5	3	9	− 6
2	9	12	− 3
3	15	20	− 5
6	26	29	− 3
8	38	40	− 2
1	41	43	− 2
4	45	53	− 8
16	46	51	− 5
14	47	48	− 1
19	47	48	− 1
15	50	47	+ 3
10	59	57	+ 2
17	59	64	− 5
13	64	78	−14
9	65	50	+15
18	67	70	− 3
11	71	68	+ 3
7	74	72	+ 2
12	78	78	0

Our experience with the SPICA has led us to the bias that while it is like the PICA, it is not the PICA. We believe that as difficult as it is to obtain reliable retest results using the same measure, it is demonstrably more difficult when you eliminate a good portion of the test and add an *S* to its name. Our current recommendations are for caution.

REFERENCES

DiSimoni, F., Keith, R., and Darley, F. (1980). Prediction of PICA overall score by short versions of the test. *Journal of Speech and Hearing Research, 23,* 511–516.

DiSimoni, F., Keith, R., Holt, D., and Darley, F. (1975). Practicality of shortening the Porch Index of Communicative Ability. *Journal of Speech and Hearing Research, 18,* 491–497.

Lincoln, N., and Ells, P. (1980). A shortened version of the PICA. *British Journal of Disorders of Communication, 15,* 183–187.
Phillips, P., and Halpin, G. (1978). Language impairment evaluation in aphasic patients: Developing more efficient measures. *Archives of Physical Medicine Rehabilitation, 59,* 329.
Porch, B. E. (1967). *Porch Index of Communicative Ability* (Vol. 2). Palo Alto, CA: Consulting Psychologists Press.

DISCUSSION

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

Q. Just a point of information, where does the ±5 percentile figure come from?
A. That is an arbitrary figure originally suggested by Porch indicating that a score less than 5 percentile represents test-retest variability.

Q. It seems to be fair to the SPICA, wouldn't you have to establish that test-retest of the PICA is better than what you got for the SPICA? It seems that using an arbitrary cutoff point without establishing that the PICA doesn't vary more than that is pretty important to this whole issue.
A. Yes. A lot of work would need to be done to establish the SPICA as a psychometrically sound test. Our original intent was to use the SPICA to shortcut a PICA so that we could then tap into the PICA scores and use the same standard that we were using for the PICA. If we are saying that they are the same measure, then we can use the same cutoff. There are a lot of variables there that don't let you do that.

Q. Does Porch report test-retest stability?
A. He reports retest data on 40 subjects. The difference was 41 to 44 percentile, so it was a range of 3 percentile.

Q. In looking for an explanation for your findings, what made the one guy on the extract 24 percentile units different? What were the peculiarities about the test pattern of those that were real different; did they have particular characteristics as aphasic people that maybe separated them from the others.?
A. We looked at those, and basically there were just random kinds of changes. We could not identify anything in particular that accounted

for that. The patient who differed by 24 percentile between SPICA and SPICA extract only differed by 9 percentile between PICA and SPICA.

Q. Didn't you do a study a few years ago that looked at the variability of the PICA, and didn't you find that it was extremely variable?

A. Yes, but those results can only be applied to patients "out in time" in a chronic phase. One was 6 years post-onset, the other 3 years, and neither were in treatment; we saw a 15 percentile unit change on retest. But that is different from a patient who is early on. In that first study, the patient delayed a lot, a strategy that served him well in the real world; he also asked for repeats, even on subtest E, a copying task, whereas on the second administration, he took the pencil and immediately started copying.

Q. Can the SPICA still be used as a descriptive tool if we are doing research and want to tell folks a little about our patients? Do you have any more feeling of caution in that situation, given the data you presented today?

A. The rule in our clinic is that you don't compare a SPICA to a PICA; you can compare a SPICA to a SPICA. We don't have any data on that, but following our observations after giving it serially to several patients, we have a comfort level there that obviously needs to be tested in a more rigid manner. It gives you a ballpark figure and so identifies a range. But to say that this specific score on the SPICA can be interpreted as a PICA score and then add onto that all the interpretations you might make from that score could be misleading.

Q. I think you demonstrated well that the mean for SPICA and the mean for PICA are not the same, well beyond chance. But it seems to me that you can predict PICA from SPICA, or the other way around, based on a .98 correlation if you plot the regression coefficient; if you want to predict X from Y with .98 correlation you have accounted for 95 percent of the variance. Why can't you predict one from the other, if that's your purpose, and maybe that's not your purpose, if you do it in the appropriate fashion with a correlation of .98?

A. Our correlation was robust at .98 and does indicate a relationship, but when you pool all of the evidence, the ANOVA with repeated measures said that there was a significant difference and clinically, half of the patients did not meet our requirement.

C. Yes, the means were not equal. Six did not equal seven. But because the correlation is high does not indicate necessarily that the means will be equal. It means that if one is high the other is high, and so

forth, but because the correlation is high I think if you apply the proper formula you can predict one from the other with pretty good accuracy.

A. We did, out of a natural curiosity, look into that. We used a regression analysis formula and found that we still had seven subjects who exceeded our ±5 percentile guideline.

Q. Why do you suppose your findings differed from DiSimoni and his colleagues? Can you talk about specific methodological differences? The second question is, can you describe a little bit how you deter mined the conversion formula?

A. To answer the second question, we did not determine the conversion formula; that was from the original DiSimoni article out of the regression formula. The first question, what is the difference, we can only speculate on differences; we would love to see his subject data. That was not available. We know that he had 93 subjects, but we don't know what the range of severity was, we don't know time post-onset, we don't know anything about the subjects other than that they were left-hemisphere aphasic subjects. There were some differences in administration in the sence that he administered the PICA and two SPICAs, we administered a PICA and one SPICA, we did ours on one day, he did his over three successive half-day intervals. There are differences; it is not a replication. In his study, the SPICA scores came out 1 percent lower than the PICA scores, and our SPICA scores were 5 percent higher.

Q. I want to go back to a previous question. Your correlation was amazingly high, it really doesn't matter what the differences between the means are as long as you can predict one from the other. I guess my question is, what do you want? You didn't like it that 7 out of 19 didn't fall within your percentile, but what would you accept? I think it predicts it pretty well at .98; what do you want?

A. The .98 is great, but clinically we want 100 percent to fall within the ±5 percentile.

C. Then you'll just have to stop using a statistical method because that is not going to happen.

C. I would like to reinforce a previous question, as well. It seems to me, that if you did the regression and you still got 7 out of the 19 subjects who didn't, I don't understand how you can do that if you use the proper regression. The regression will define the .98 correlation; you had to have misclassified maybe 1 of the 19, not 7. I don't understand where that came from. If you do the regression, you regress on *X* to predict *Y* for the correlation of .98; you have to do .95 percent of them correctly, which would be one that you wouldn't. And

all you have to know is that regression formula to predict them all perfectly. You could make the little correction the way DiSimoni tried to do.

C. I think it depends on the purpose; things can be significantly related and significantly different. I think what you're suggesting is that they are significantly different and so you may not want to mix them if you're using them as a change measure. Or, they may give you different impressions of overall severity.

C. It depends on your purpose, you may or may not want to make that prediction. I think she addressed that, but if you do want to make it, I think you can.

C. Part of the problem on this percentile thing is that you are correlating raw scores and then going back to percentiles, and that causes some of the problems that we've been talking about; but you can predict the overall scores, if not the percentiles. So correlate your percentiles, and see what happens.