The purposes of this study were (1) to devise two forms of an instrument for measuring the listening ability of first, second, and third graders; (2) to test the instrument on students and teachers; and (3) to measure the reliability of the instrument. To obtain objective data concerning the listening ability of young children (6-9), four subproblems were examined: (1) the relationship of listening ability to intelligence and achievement and the relative listening abilities of (2) males and females, (3) first, second, and third graders, and (4) black and white students. The listening skills measured (comprehension, interpretation, and evaluation) were derived from Bloom's taxonomy. The listening test was administered to 453 white and 255 black primary school children by their own teachers. Achievement was measured by the Stanford Achievement Tests and intelligence by the Kuhlman-Anderson Test. Analysis of the data indicated that listening ability correlated more highly with achievement than with the intelligence measure. There was no difference in the listening scores of males and females. There was a significant, but not unexpected, difference in scores obtained by students in each grade level. An extremely significant difference revealed that black children were about one standard deviation below white children as measured by these listening ability instruments. (MH)
Investigators have agreed that there is a separate listening factor which can be measured objectively. They have further found that listening skills improve with instruction. Some educators have indicated that systematic instruction is as necessary for listening as for reading. While it has been proven that upper elementary grades could be tested objectively for listening ability, little has been known about specific listening abilities in the lower grades because there has been no standardized test published for use with primary grades. A reliable means of measuring listening ability in primary grades seemed to be necessary for teaching listening systematically.

It was the purpose of this study (1) to construct two forms of an instrument for measuring the listening ability of children in grades one, two, and three; (2) to test the instrument by using the students and their teachers; and (3) to analyze the items and measure the reliability of the instrument.

In order to obtain objective data concerning the listening ability of young children, these sub-problems were explored: (1) the relationship of listening ability to intelligence and achievement; and whether there was (2) a significant difference between the listening ability of males and females; (3) students in first, second, and third grades; (4) and Black and White students.

All the primary children, of which thirty-five percent were Black, in a single school district near a large metropolitan area in the southwest, except the absentees on the test dates, participated in the study; item construction involved 181 students. Preliminary tests included another 810, and the final forms of the instrument were tested on the remaining 748 students.

The Pearsonian Coefficient of Linear Correlation was used to measure the reliability.

Because of the insufficiency of direct criteria for evidence of concurrent and predictive validity, it was necessary to rely heavily on construct validity—logical inferences which are drawn from indirect evidence that the test measures what it claims to measure.

The listening skills to be measured were derived in most part from Bloom's Taxonomy of Educational Objectives, the Cognitive Domain. Listening skills to be measured were as follows: (1) Comprehension through identifying main ideas, significant details, sequence of ideas, and denoted meanings; (2) Interpretation, through displaying the implication of ideas, implication of details, and connotative meanings; and (3) Evaluation and application by means of judging the sufficiency of details, recognizing intent, criticizing organization, and judging mood and effect.

1 Presented at the Annual Meeting, American Education Research Association, Minneapolis, Minnesota, March, 1970; and developed under the direction of J. L. Fearing, Texas Woman's University.
The listening selections were chosen materials two or three grade levels above the reading level of the group for which the test is intended. This was done because the listening vocabulary of primary children has been shown to be much larger than their reading vocabulary. The items in the instrument are multiple choice with three options as opposed to the usual five. Research shows that primary children are confused when they are confronted with more than three choices.

Listening materials were chosen for the type of skills being tested. It was not difficult to select material for testing for "main ideas", "details", and other simple comprehension skills. It proved much more difficult to find material to test for "interpretations" and "evaluation". Testing for "implications" was made possible by the use of poetry wherein meanings are often implied. Fables were useful in testing of "intent" and "evaluation".

The items which tested the student's ability to listen to and follow spoken directions were devised by the investigator. Primary children, as a rule, had no trouble in remembering and following one direction. It was the second and third direction which discriminated between "good" and "poor" listeners.

Each teacher participating in the study administered the tests to her own class. It was reasoned that young children would react more freely to a familiar voice than an unfamiliar one. The selections were all read by the teachers themselves. They were instructed to practice the selections beforehand, so as to interpret them with expression and meaning. Both forms of the test were given within one week in order to minimize between-test-learning.

The tests, which had been constructed from pre-tested, discriminating items in the middle range of difficulty, were scored, tallied, and analyzed. An item analysis was done to determine the effectiveness of individual items. None of the items discriminated negatively, and all the distractors functioned.

Achievement was measured by the Stanford Achievement Tests. The available scores were correlated with corresponding listening scores with a resulting .638 for Test I and .650 for Test II. Those correlations indicated a substantial relationship between listening ability and scholastic achievement.

Intelligence was measured the Kuhlman-Anderson Test, the quotients were correlated with the corresponding listening scores with a resulting .422 for Test I and .512 for Test II. Listening as measured in this study, correlated more highly with achievement than with intelligence.

A comparison was made of listening abilities, as measured in the study, of Black and White students. A z ratio was used to determine the significance of the difference in means. The findings of this study indicated a significant difference in the listening ability of the two groups of students.
**Significant at the .01 level.**

There was a difference in means of almost one standard deviation.

A summary of the important comparisons were as follows:

<table>
<thead>
<tr>
<th>Sub-group</th>
<th>Test I</th>
<th>Test II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade One</td>
<td>10.14</td>
<td>10.13</td>
</tr>
<tr>
<td>Grade Two</td>
<td>11.18</td>
<td>11.54</td>
</tr>
<tr>
<td>Grade Three</td>
<td>12.93</td>
<td>13.05</td>
</tr>
<tr>
<td>Male</td>
<td>11.42</td>
<td>11.61</td>
</tr>
<tr>
<td>Female</td>
<td>11.40</td>
<td>11.52</td>
</tr>
<tr>
<td>White</td>
<td>12.29</td>
<td>12.62</td>
</tr>
<tr>
<td>Black</td>
<td>9.85</td>
<td>9.70</td>
</tr>
</tbody>
</table>

When the means were compared and analyzed by using a z score there was found to be no significant difference in the listening scores of males and females. There was, however, a significant but not unexpected difference in the scores obtained by students in each grade level. An extremely significant finding was the difference in means of Black and White students. Certainly more study should be done in this area.

The results of this study showed that the listening ability of children in grades one, two, and three could be measured reliably, may be a more reliable measure of scholastic ability than intelligence tests, and that black children are about one standard deviation below white children as measured by these instruments.