

Chapter II Methodology

2.1 Description of the Sample

The student population at the Earl C. McGraw School in Hampden consisted of approximately 110 kindergarten children who attended school for half-days and 110 first grade students.

The population was white middle class. Within each grade thirty students were randomly selected from the population to take each of the seven tests and the Learner Perception Interview Schedule (LPIS).¹ The single constraint on the sampling process was that a student should not be tested more than three times.

The student population in the control group consisted of 30 kindergarten students who attended school for half-days and 30 first grade students. The population was comparable to the Hampden population in terms of socio-economic status. As the population was rather small, each student was scheduled to take each test.

2.2 Experimental Design

The experimental design utilized in this study is known as a static group comparison design (Campbell and Stanley, 1966). The design consists of making observations on previously existing experimental and control

¹The LPIS is explained in Chapter III.

groups at the end of the school year. The choice of this design was dictated by two constraints. First, it was impossible to randomly assign students to the ANISA group and the control group. Thus, previously existing intact groups had to be utilized. Second, as the necessary measurement instruments did not exist at the beginning of the school year, it was impossible to conduct a testing program at that time.

2.3 Test Administration

All tests and student interviews were individually administered. The tests of the processes were administered by second and third grade teachers and student teachers from the Hampden school, staff members of the Laboratory of Psychometric and Evaluative Research, and community volunteers. All test administrators tested children in both the ANISA and control schools. The Learner Perception Interview Schedule (LPIS) was administered to students by staff members of the Laboratory of Psychometric and Evaluative Research. These staff members also completed the Learning Environment Observer Rating Scale (LEORS).¹

2.4 Limitations of the Methodology

Some of the limitations of our work were inherent in the static-group comparison design used in the study. These limitations were that we could not randomly assign students to the ANISA and control groups or administer the tests at the beginning of the year. (Among the

¹The LEORS is explained in Chapter III.

be sure that the groups were equivalent at the beginning of the study. As a result, it is possible that the observed results were caused by pre-existing differences in the two school groups.

Another limitation was that the resources and time available for the evaluation during the first year were not sufficient to enable us to conduct extensive validation studies on the measuring instruments developed in our work. Thus, we cannot be absolutely sure that our instruments measure what they are intended to measure.

Because of the above limitations, along with others to be mentioned in later chapters, the results of this year's evaluation need to be interpreted with extreme caution. The results should be considered to be quite preliminary.

Finally, it is important to note that the ultimate success of the ANISA approach rests on the summation of small incremental gains on hundreds of different dimensions -- gains which are cumulative in their effects over time, even though each small gain may not prove statistically significant by itself. During this first year, it was possible to sample performance on just a few dimensions. Thus, much more has taken place than is reflected in the analysis of the test results. For example, we collected no data related directly to the extensive efforts to individualizing instruction, the provision of a wider variety of environments, the attention given to the affective or emotional life of the child, or the care given to make certain that each child will love the formal learning experience. As performance on more dimensions can be examined over longer periods of time, a more complete picture of the model's efficacy will emerge.