a confounding factor. A final difficulty in this area is that often you can't randomize who is in a class. The solution is to use the classroom as a unit.

When performing statistical tests Campbell and Stanley say that both assignment to groups and sampling must be random. It is acceptable to do a t-test on gain score or use pretest scores as a covariate.

The second design is the Solomon Four-Group design diagrammed below:

$$R \quad O_1 \quad X \quad O_2$$
$$R \quad O_3 \quad \quad O_4$$
$$R \quad \quad X \quad O_5$$
$$R \quad \quad \quad O_6$$
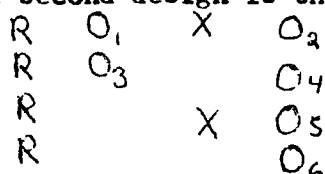
If $O_2 > O_1$, $O_2 > O_4$, $O_5 > O_6$, $O_5 > O_3$ inference is much stronger.(A 2x2 ANOVA can be used for analyzing this design). The main effect of testing and the interaction of testing and X are both determinable. However other problems remain and this method requires a high expenditure of effort that may not be worth it.

Design 6 is shown below:

$$R \quad X \quad O_1$$
$$R \quad \quad O_2$$
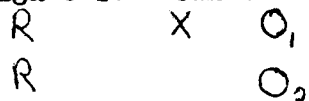
This design recognizes that a pretest is not essential. Sometimes it is recommended that this design be used as a pretest would. However where *be hard to administer or have an effect of its own.* pretest scores are already available the first design should be used since the statistical tests for it are more powerful. The availability of pretest scores also makes the examination of the interaction of X and the pretest ability level possible. For statistical analysis the t-test is often used.

If there are two or more treatment variables, each at several levels this complicates the analysis. Interaction effects can be very important especially for generalizations. Sometimes the main effect is nonexistant but interaction is very important. Nesting means that certain interactions cannot be measured. Ex: teachers cannot be at all schools so teacher school interaction cannot be measured.