Sequential Model for Appraising Instructional Superiority of a Revised Material

Subramaniam Kandaswamy

The effectiveness of a revised instructional material compared with its prototype usually is determined by subjects' performance on criterion-referenced tests. The alternative method outlined in this article—sequential analysis—appears to have several advantages over the traditional method. For one thing, it permits estimation of how much better (or worse) the revised material might be, in terms of both educational and statistical significance. In addition, it provides advance estimates of the number of subjects required (usually 50% fewer than traditionally), allows immediate analysis of data, and offers a scheme for controlling both Type I and Type II errors.

The traditional method of examining whether a revised instructional material is better than its prototype has been to assign the prototype and the revised material to separate groups and compare the groups' performances, usually on a criterion-referenced test, by means of a t test or a similar test. Several studies (Abedor, 1972; Gropper & Lumsdaine, 1961; Kandaswamy, Stolovitch, & Thiagarajan, 1976; Robeck, 1965; Rosen, 1968) have used this approach.

All these studies indicate that the revised materials were instructionally more effective than the prototype. They do not, however, tell us how much better. The difference in effectiveness between the prototype and the revised material could have been statistically significant—favoring the revised material—and yet not educationally significant. Recently educators have begun to question statistical significance and strive for educational significance (e.g., Cohen & Hyman, 1979).

The procedure outlined in this article, known as sequential analysis (Wald, 1945), enables one to estimate how much better the revised material is in terms of both educational and statistical significance. Thus one can set up a minimum standard, in terms of probability of success, and adopt the revised material only if it meets the standard.

Sequential analysis has many advantages over traditional statistical procedures in that it: (a) normally allows 50% fewer subjects and observations; (b) allows the immediate analysis of data (instead of waiting until the end of the experiment); (c) offers a scheme for controlling both Type I and Type II errors; (d) provides, in ad-
vance, estimates of the required number of subjects under different conditions; and (e) eliminates the task of computation on the part of the researcher.

SEQUENTIAL MODEL FOR APPRAISING REVISED MATERIAL

Sequential analytic techniques have been applied successfully in a variety of fields, including military science and medicine (Wald, 1947; Wetherill, 1975). Education is not an exception. Anastasi (1953), Burgess (1955), and Walker (1949) used sequential analysis to select test items. Cowden (1946) and Moonan (1950) employed it for appraising performance on achievement tests. Ferguson (1970) and Kandaswamy (1979) applied sequential analytic techniques to domain-referenced testing. Kimball (1950) used sequential analysis for checking accuracy of scoring. The purpose of this article is to illustrate how sequential analysis can be used efficiently to assess the instructional effectiveness of revised materials. Although any two types of instructional materials (or methods) might be compared by sequential analysis, this paper focuses on a comparison of the instructional effectiveness of a formatively evaluated and revised material with its prototype.

PROCEDURE

In order to compare the instructional effectiveness of two materials, the traditional statistical procedure requires a sample of fixed size, usually determined before the experiment. However, in sequential analysis, the actual size of the sample is determined during the course of the experiment. The two materials to be tested are tried with one pair of subjects at a time and, on the basis of the subjects' performance, one of the following three courses of action is taken:

1. Testing is terminated when the prototype material is superior to the revised material.

2. Testing is terminated when the revised material is superior to the prototype.

3. Testing is continued if there is insufficient evidence of the superiority of any material over the other.

If testing continues, the materials are tested with a second pair of subjects. On the basis of evidence up to that point (i.e., the subjects' performance on the first two trials taken together), one of the three decisions is again made. The testing continues with subsequent pairs of subjects until a decision on superiority of one material over the other is established.

To examine whether the revised material is instructionally superior to its prototype, the evaluator specifies the following values in advance:

- \( p \) the acceptable level of performance on the posttest for either material.
- \( u_1 \), the standard below which the prototype is superior to its revised counterpart.
- \( u_2 \), the standard above which the revised material is superior to its prototype.
- \( \alpha \), the probability of erroneously deciding in favor of the revised material when, in fact, the prototype is superior.
- \( \beta \), the probability of erroneously deciding in favor of the prototype when the revised material is superior.

A pair of straight lines drawn with four of the above parameters determines the criterion for selecting one of the three courses of action mentioned earlier. The two straight lines are derived as follows:

\[
L_1 = \frac{\log \beta}{\log u_2 - \log u_1} + \frac{\log 1 + u_2}{\log u_2 - \log u_1} \cdot n \quad (1)
\]

\[
L_2 = \frac{\log 1 - \beta}{\alpha} + \frac{\log 1 + u_2}{\log u_2 - \log u_1} \cdot n \quad (2)
\]

\( ^1 \)In fact, the term standard stands for the odds ratio, defined as the ratio of the proportion of successes to failures of one material over the other.