There is an approach to the evaluation of social action programs which seems so sensible that it has been accepted without question. The underlying assumption is that action programs are designed to achieve specific ends and that their success can be established by demonstrating cause-effect relationships between the programs and their aims. In consequence, the preferred research design is an experimental one in which aspects of the situation to be changed are measured before and after implementation of the action program. To support the argument that the program is responsible for the observed changes, the anticipated effects may be measured simultaneously in a control situation that does not receive the program (Campbell & Stanley, 1966). This plausible approach misleads when the action programs have broad aims and take unstandardized forms.

The term broad-aim program is intended to describe programs that hope to achieve nonspecific forms of change-for-the-better and which also, because of their ambition and magnitude, involve unstandardized, large-scale interventions and are evaluated in only a few sites. These characteristics have been shared by a
number of social action programs launched during the 1960s — the delinquency prevention and grey area programs, portions of the poverty program, and the model cities planning program. These programs attempted to produce increased community competence, increased participation of low-income citizens in community action, and more effective utilization of existing institutions. All these aims could be realized in many alternative ways, and none of these programs could be judged by whether one particular end was achieved.

In the evaluation of broad-aim programs, experimental design creates technical and administrative problems so severe as to make the evaluation of questionable value. A more historically oriented, more qualitative evaluation has greater value. The experimental model has been criticized by Stufflebeam (1968), Suchman (1968), and Schulberg and Baker (1968). However, only Schulberg and Baker have agreed with the position taken here that the experimental model is intrinsically unsuitable to the evaluation of broad-aim programs. Even among critics of experimental designs for program evaluation, this is a minority position.

A Case in Point

The difficulties encountered in evaluation by experimental design may be illustrated by a case study. The failures and frustrations encountered in this case, we will argue, derived from the inability of experimental design to respond to critical issues in the evaluation of broad-aim programs. Certainly they cannot be attributed to incompetence or indolence among the research staff; the project directors were well-trained, committed, and hard-working, and, in addition, in the first year of the study they incorporated in their planning the advice of several men who are greatly respected as experts in social research.

Let us call the program the Neighborhood Benefit Program. Its aim was to change existing community institutions — social agencies, schools, employment services — so that they might be more useful to citizens of the community and especially to underprivileged youth. It was a broad-aim program, as we use the term: general aims, unstandardized intervention, virtually unreplicated evaluation.

Criterion Development

An evaluation study was required by the federal agency that funded the program. The agency, acting on the advice of its scientific consultants, recommended that the evaluation study be as methodologically rigorous as possible. In addition, the agency made clear that the program should be construed as primarily a test of its form of intervention in functioning systems, rather than as a remedy specific to the ills of the city in which it was being introduced. This meant that the research group