I am pleased to have this opportunity to talk about some recent developments in the methodology of program evaluation and about what I call responsive evaluation.

I feel fortunate to have not only these two days but also some seven months to think about these things. My hosts here at the Göteborg Institute of Educational Research have been most hospitable, but generous also in hearing me out, pointing my head in still another way, weighing the merit of our several notions, and offering occasionally the luxury of a passionate argument.

When Erik or Hans or Sverker or Ulf and I agree, we are struck by the fact that the world is but one world, and the problems of education are universal. When we disagree, they are quick to suggest that the peculiar conditions of education in America have caused me to make peculiar assumptions and perhaps even warped my powers of reasoning. I am sure that some of you here today will share those findings. What I have to say is not only that we in educational research need to be doing some things we have not been doing, but that in doing what we have been doing we are in fact part of the problem.

Our main attention will be on program evaluation. A program may be strictly or loosely defined. It might be as large as all the teacher training in the United States
or it might be as small as a field trip for the pupils of one classroom. The evaluation circumstances will be these: that someone is commissioned in some way to evaluate a program, probably an ongoing program; that he has some clients or audiences to be of assistance to — usually including the educators responsible for the program; and that he has the responsibility for preparing communications with these audiences.

In 1965, Lee Cronbach, then president of the American Educational Research Association (AERA), asked me to chair a committee to prepare a set of standards for evaluation studies, perhaps like the Standards for Educational and Psychological Tests and Manuals, compiled by John French and Bill Michael and published in 1966 by the American Psychological Association. Lee Cronbach, Bob Heath, Tom Hastings, Hulda Grobman, and other educational researchers have worked with many of the U.S. curriculum-reform projects in the 1950s and early 1960s, and have recognized the difficulty of evaluating curricula and the great need for guidance on the design of evaluation studies.

Our committee reported that it was too early to decide upon a particular method or set of criteria for evaluating educational programs, that what educational researchers needed was a period of field work and discussion to gain more experience in how evaluative studies could be done. Ben Bloom, successor to Lee Cronbach in the presidency of the AERA, got the AERA to sponsor a monograph series on curriculum evaluation for the purpose we recommended. The seven volumes completed under AERA sponsorship are shown in the Reference section. The series in effect will continue under sponsorship of the University of California-Los Angeles (UCLA) Center for the Study of Evaluation, whose director, Marv Alkin, was a guest professor here at this Institute for Educational Research two years ago. I think this monograph series can take a good share of the credit, or blame, for the fact that, by count, over 200 sessions at the 1973 AERA annual meeting programs were directly related to the methods and results of program-evaluation studies.

There were two primary models for program evaluation in 1965, and there are two today. One is the informal study, perhaps a self-study, usually using information already available, relying on the insights of professional persons and respected authorities. It is the approach of regional accrediting associations for secondary schools and colleges in the United States and is exemplified by the Flexner report (1916) of medical education in the USA and by the Coleman report (1966) of equality of educational opportunity. In Nine Approaches to Educational Evaluation (see Appendix A), I have ever so briefly described this and other models; this one is referred to there as the Institutional Self-Study by Staff Approach. Most educators are partial to this evaluation model, more so if they can specify who the panel members or examiners are. Researchers do not like it because it relies so much on second-hand information. But there is much good about the model.