Extensive educational reforms or innovations, even those advocating bottom-up approaches, have usually been initiated at the national or State level by political parties, national educational administrations or researchers and scholars from academe. However, to be successful, the results of such initiatives must be seen at the school level. Having observed educational innovation throughout the world, it is clear to me that if something significant happens at all in education, it does not happen at the ministry of education or at the university; it happens at the school and in the classroom. Reforms can be conceived by policy makers. They can be planned in ministries of education by experts and administrators. But, to have an effect, they must be implemented in the school by teachers and school principals. To change education is to change the school, and evaluation used to improve an education system cannot be effective without bringing the evaluation process—in its full capacity and potential—into the school building. If evaluation is to be used to improve education it has to be dealt with at the school level.

Evaluation has always existed within the school, mainly in its classrooms, but in a very limited capacity. Evaluation was there even before administrators began
to use it for decision making or the monitoring of operations. But evaluation within the school was usually limited to the evaluation of students, executed by means of tests, grades and report cards. The notion of a systematic evaluation of instructors, programmes, projects or the school as a whole, has reached the school building only in recent years, and is still far from being common practice in even the most advanced education systems. Even schools that have adopted innovative approaches to student evaluation, such as performance assessment, portfolios or exhibitions, do not necessarily extend their perception of evaluation beyond the domain of student assessment.

But it is not only the school that has had a narrow perception of evaluation. For many years this was the general perception of evaluation in education (Nevo, 1983, 1995). Educational evaluation was generally equated with student evaluation and the words 'evaluation' and 'measurement' were synonymous. Until the mid-1960s most books carrying titles like Educational evaluation, Measurement and evaluation in education, etc., dealt mainly (in most cases, only) with student evaluation, focusing on test construction, reliability and validity; scoring; and comparing the pros and cons of various types of test items. The change came in the late 1960s, and mainly in the United States, when the concepts of curriculum evaluation and programme evaluation were suggested in the context of newly developed curricular materials and large-scale projects for socially disadvantaged students (Scriven, 1967; Stake, 1967; Stufflebeam et al., 1971).

New ways to perceive evaluation were developed, distinguishing between 'evaluation' and 'measurement,' and extending the scope of evaluation beyond student assessment with the introduction of programme and curriculum evaluation, followed later on by teacher evaluation (Millman, 1981; Millman & Darling-Hammond, 1990). Around this time it was suggested that evaluation could assist with decision-making, classification and accreditation. As a result, evaluation methodology was advanced by combining quantitative and qualitative research methods and was supplemented by other methods, such as judicial procedures (Wolf, 1979) and art criticism (Eisner, 1979). Educational evaluation also became an established profession. University programmes were developed to train professional evaluators at both the M.A. and Ph.D. levels, setting agreed-upon standards to control the quality of the evaluation process (Joint Committee on Standards for Educational Evaluation, 1981, 1988, 1994).

But all this happened—in spirit and in practice—far away from the school building. Evaluation experts, concerned with conceptual problems related to programme and project evaluation, rather than with issues of interest to teachers and school principals, practised their profession with large-scale projects at national and State levels. In the United States, the cradle of modern evaluation and the country with the world's largest testing industry, evaluation provided the conceptual basis and methodological solutions for the accountability movement. It supported the idea of posing evaluation requirements as a means of controlling schools, and it helped create the illusion that instruction could be test-driven. Educational evaluators seemed more interested in working with administrators