

EVALUATION - A SERVICE TO CURRICULUM IMPLEMENTATION:  
A PERSPECTIVE FOR PLANNING EVI CIEL\*

HANS BRUGELMANN  
EVI CIEL

1. Background to the project
  - 1.1 The CIEL programme was set up by the Stiftung Volkswagenwerk foundation in 1971 to promote curriculum development for kindergarten and primary schools in West Germany. A number of projects in different curriculum areas were initiated or supported, four of which, at present, are participating in an extensive evaluation and implementation programme. The four projects cover the following areas:
    - natural studies (age group 5-9)
    - integrated infant curriculum (5-7)
    - cognitive social skills (5-7)
    - reading and pre-reading (5-7)
  - 1.2 The projects have started from different political, pedagogical and strategic premises, and they follow different approaches. Some concentrate on the production of materials, while others emphasise in-service education for teachers or involvement of parents. Some favour a structured programme, while others codify their ideas in more open and flexible ways.
  - 1.3 All projects have developed and field-tested at least parts of their final products. Two of them still have up to two years to go. In 1973 the advisory body attached to the CIEL programme suggested mounting an independent evaluation of some projects with specific emphasis on implementation problems. The Stiftung Volkswagenwerk agreed to set up a project with six research, one

\*EVI CIEL is an acronym for Evaluation and Implementation of Curriculum Institutionalised Elementary Education

administrative and three secretarial staff (all full-time) for a four year period which commenced in 1975. The budget amounts to one million dollars and thus the project is one of the largest curriculum evaluation programmes undertaken in West Germany.

1.4 As far as I gathered during my negotiations about the project directorship, different expectations were involved in setting up the project. The basic proposal submitted by a working party of the advisory group, therefore, can be interpreted differently by different parties. In particular, there seems to have been some disagreement about the function of evaluation: to what extent should it yield data contributing to the revision of the curricula on paper and to what extent should it feed directly into dissemination and implementation activities? There was some controversy about the methodology appropriate for such an enterprise: some people favoured an extensive and careful measurement programme, whilst others supported a more holistic and naturalistic approach. In my application for the directorship of the project I strongly pleaded for the second option. As these issues are of principal importance in curriculum evaluation in general, I consider it to be worthwhile to follow them up with reference to some of the more basic ideas, though within the context of the project in question.

## 2. Significant features of the proposal

I shall start my description by outlining some central features of the project. These features are related to four basic issues.<sup>(1)</sup>

- Who shall be the audiences of the evaluation programme?
- What are their information needs?
- What, then, are the tasks of an evaluation programme that attempts to respond to the information needs as defined above?
- What kind of role is appropriate in achieving these tasks?

2.1 Two major groups are important for the success of the innovations; the development teams, and the potential users of, or parties effected by, the programmes. The development teams have been working for a minimum of three years already, and thus, have developed a strong (at least external) identity in terms of what

they want to achieve and how they want to go about implementing their ideas. It did not seem very promising therefore to aim at influencing decisions about the philosophy and the design of the curricula. More generally, one may doubt that a well established development group will respond to interventions from an inspection agency with many resources, but with no obligation to become involved in the construction and production of materials. It is not very probable that they will change what they achieved after long disputes because of the criticism of an outside group with no prestige in the particular curriculum area. This type of unwillingness is strengthened, I assume, by the time pressures under which the projects have to complete their work. Even if they want to use feed-back for revising their programmes they will not have the time for more fundamental changes.

2.2 The most immediate user group will be teachers and children as well as their parents. In addition, innovation research has shown the importance of school administrators and of pre- and in-service educational personnel. Relevant to these groups, then, are questions such as these:

- What is the specific pedagogical potential of the CIEL curricula in comparison to similar programmes?
- Under which conditions of implementation are these curricula likely to be successful and what kind of problems have to be expected under other circumstances?
- What kind of strategies or tactics are available to cope with such problems in implementing the curricula productively in different contexts?

In answering these questions we have to be aware that users vary considerably in their competence and motivation. In particular, we should not exclusively focus our attention on the groups of 'pioneers' and 'early adopters'. On the other hand, experience has shown that one cannot increase the number of classrooms adopting a new curriculum, and especially the number of teachers working creatively with an innovative programme, easily.

The evaluation programme, therefore, cannot aim at a broad short-term dissemination of the curricula. More effectively, both in

terms of user relevance and of solid, enduring innovation, is a strategy of initially limited implementation gradually spreading on a long-term scale. Educational change is a slow and controversial process.

- 2.3 In the course of negotiations about the design and the strategy of the project, four major expectations emerged that are related to the preceding considerations:
- The evaluation programme should explore the potential of the curricula and make audiences aware of their principal possibilities and difficulties.
  - It should define 'archetypes' of implementation settings by exploring circumstances in schools in the lives of the children that create additional problems or are particularly responsive to the curricula in question.
  - It has to contribute to the central revision or local adaptation of the curricula-on-paper, ie, to increasing their pedagogical potential for use in different settings as far as possible on a general level.
  - It has to contribute to the development of a support structure that makes it easier for the users to work creatively with the programmes in different contexts and to cope with the implementation problems as identified in the first two questions.

2.4 Thus the evaluation programme focuses on two equally important aspects of the projects under investigation; their pedagogical strategies, and, the innovation strategies they use for making the curricula accessible to users in the system. (cf. for a similar perspective M. Eraut, 1975)

### 3. Basic assumptions of the design<sup>(2)</sup>

- 3.1 In designing a strategy for EVI CIEL I started from the following hypotheses:
- People tend to disagree about the objectives of education and about the criteria by which to judge the 'success' or 'failure' of curricula. Evaluators cannot claim any particular authority in this respect. The evaluation of a curriculum has therefore to respond to divergent interests and value conflict as an intrinsic feature of education.
  - The same curriculum-on-paper will have different effects in different settings - it will lead to a broad range of classrooms-in-action. Educational events are unique in a significant sense. Evaluation has therefore to account for the specificity of implementation conditions rather than striving for broad generalisations.

- Even the effects of a curriculum within a specified context may be perceived differently by different participants. So-called 'scientific' theories or methodologies do not provide a basis for validating certain interpretations and for overriding others. Evaluation has therefore to represent discrepant perceptions and to make them accessible to public judgement.

- 3.2 Let me quote a simple analogy for this conception of research and evaluation: When I choose a taxi for going to see my friend, I do not believe that the taxi driver can drive better than I could (far from the truth - though there are exceptions); nor does he always know how to go there better than anybody else (though sometimes he does, of course). I just find it convenient to get rid of the additional burden of finding a parking space, of looking up the house on the map or of keeping an eye on the traffic, especially during rush hours. It is a question of specialisation. There are a number of people, for example, who could make nicer or more useful cupboards than industry offers on average. They mostly find it time-saving, though, not to do so (if they need a special size, they may construct one on their own, however).
- 3.3 There are a number of advantages in the<sup>4</sup> division of labour, in co-operation, in the delegation of responsibilities. It's the same with evaluation. We all do research. Some of us are better in this job than many of the professionals, i.e. we are more perceptive, more rigorous, more creative, or even more systematic. Nevertheless, it seems to be useful to have people specialising in this area as in any other, who develop this particular ability, who try out new instruments, who reserve their time and energy for this task. The distinction between practitioners' 'lore' and scientific 'theory', then, is not a principal one in my opinion. It is a question of specialisation with all its benefits and dangers.
- 3.4 Therefore we need a system of checks and balances through which the process of gathering, processing and presenting information becomes accessible to public scrutiny. The rationale of doing research, then, is not to approximate some kind of 'truth'; it rather should aim at fostering the exchange of experience and at

breaking down monopolies of information and democratising the control of 'knowledge'. In this sense, evaluation becomes a highly political activity. The conventional problem of increasing the methodological precision of evaluation has to be redefined, then, as one of improving its social control (controlled relativity paradigm).<sup>(3)</sup>

4. Procedural implications of the 'controlled relativity' paradigm  
Social control comprises two elements: (a) well defined expectations of what people should do (norms) and (b) effective sanctions enforcing these norms.

- 4.1 Conventionally expectations have been defined in terms of instrumental rules. Such rules prescribe the use of certain procedures within a 'subject-object' relationship. The aim is to increase the degree of technical precision (e.g. laying down how to construct a questionnaire appropriately, or how to use certain statistical techniques).

If we accept that there are no substantive or procedural criteria available by which to discern 'true' or 'right' interpretations of reality, we have to look for some means to regulate the process of gathering, processing and reporting information. The example of law is particularly interesting since it has to cope with a comparable situation. Thus we could try to lay down standards that should govern the process of research by regulating the social relationships involved; we have to develop a code of conduct for evaluators. In a previous paper I have called for the establishment of a system of 'checks and balances' that can help to limit the risks of deliberate or unconscious distortion. Authority, group pressures, institutionalised power, traditions, they all can endanger a free exchange of experience and pluralism of values. Organised properly, on the other hand, they can help to stabilize an open system.

For reasons of space, I have to summarise very briefly some standards that seem to be particularly important.<sup>(4)</sup> The first five standards

focus on the process of evaluation, while the second group refers to the presentation of results.

- unobtrusiveness: the investigation should interfere as little as possible with the work in the field; the evaluator should refrain from creating problems in the solution of which he is not willing to participate;
- transparency: the work of the evaluator should be accessible to the scrutiny of participants and audiences; it must be possible for them to evaluate the evaluation as it were;
- impartiality: the evaluator should not take sides but represent different points of view and respond to different information needs, criteria and interpretations;
- confidentiality: 'people own the facts of their lives' (SAFARI 1974), i.e. the researcher has to respect the integrity of his informants; they have to decide what kind of information should be released and when;
- research as a 'learning experience' for all participants: the evaluator should serve participants rather than exploiting them - he has obligations in terms of fast and useful feed-back;
- relevance: evidence should be reported because it promises to be useful to audiences rather than because of academic merits; different audiences might need different kinds of data or they might need different formats of presentation;
- transferability: evidence has to be reported in such a way that it can be used to understand other situations; the presentation should support this instrumental use of reports;
- inconclusiveness: reports should not pre-empt decisions, they should feed them; they therefore must contain enough (excess) information to enable audiences to draw their own conclusions;
- accuracy: if real events are described, the report has to be authentic and fair; if the report is fictive, it still has to be detailed and of high verisimilitude;
- anonymity: informants should be protected against the curiosity of outsiders.

4.2 These standards resemble legal rules, as I said. The degree of their implementation in a particular research activity cannot easily be determined. Impartiality, relevance or accuracy cannot be measured against a yardstick. If we want to enforce these standards, or at least minimise the number of cases where they will be violated, we have to devise a social structure that

supports adherence to them.

The simplest device used in the legal models is the differentiation of roles and tasks and the distribution of power. It is useful, for instance, to split the research process up into stages and to link feed-back obligations to these stages. This gives participants a chance to calculate their information risks. Another means would be to give different responsibilities to different people. Thus, one person might be responsible for gathering the data while another will prepare the final presentation. Or one might ask two investigators to provide evidence from two different points of view as in the 'adversary model'<sup>(5)</sup>.

There are cases, however, when a distribution of rights and obligations between evaluators, participants and audiences is not sufficient. What can be done, for instance, when one group of participants feels that their point of view has been misrepresented in the final report? It is here that an arbitrary body might be useful (often its mere existence will prevent people from violating the rules). Thus in EVI CIEL we have established a formal procedure and a board for settling such disputes. The board comprises representatives from the project team, from the development groups, from the CIEL advisory body and from trial schools. Another complication may arise when an outside teacher, for example, wants to check the accuracy of a statement by an informant to whom anonymity has been granted. Here, it might be useful to have nominated an independent confidential person who will have access to the files with the consent of all participants. Thus, one could satisfy information needs from audiences who want to check evidence and at the same time preserve the confidentiality granted to informants. ,

Though such formalised mechanisms are important in providing a way out in borderline cases, the feasibility of the 'controlled relativity' paradigm rests on the informal relationships built up in the course of the work. The distribution of responsibilities as mentioned above provides the groundwork. The structure cannot

become effective, however, without participants being interested in the exchange of information. The evaluator will not get any relevant information as long as he cannot convince people that participation is worthwhile. The best means he has, is fulfilling his promises. This to me seems to be a sound basis for co-operation between researchers and practitioners.

5. Relevant methodological options<sup>(6)</sup>

5.1 The evidence gathered about curricula and their fate as innovations usually is either poor or puzzling. As mentioned earlier, educational events seem to be highly context-dependent and experience does not lend itself easily to generalisations. Interpretations often contradict each other; theories so far seem not to have been particularly useful in practice. The inconclusiveness of evaluation findings points to the need of closer observation of reality. I suggest that studies of specific instances leading to densely interrelated descriptions will contribute more to our understanding of curriculum innovations than representative surveys focussing on few selected variables. Such an approach presupposes a methodology of research that is perceptive to the complexity of the issues under investigation.<sup>(7)</sup> However, the present state of the art in educational research does not allow for designing an approach based solely on one tradition. The definition of roles and the selection of procedures depend on many circumstances, e.g. the purpose of gathering data, the kind of evidence to be processed, the resources available, the interests and responsiveness of participants. Thus, I would like to explicate a range of options for approaching the issues in different settings; and this can be seen as an eclectic rather than uniform approach to methodological design.

In looking for constituent elements in different approaches we should try to discover their respective advantages and shortcomings: what can we gain by adopting a standardised observation scheme, for instance - and where do we lose information? My assumption is this: one cannot achieve both precision and richness, focus and scope, reliability and situational validity by adopting

one methodology only. We should, therefore, try to find out more about the conditions under which these options become significant. The following distinctions could serve as a starting point for exploring these issues.

- 5.2 In deciding about the number of cases one wants to investigate, there are two dimensions relevant.
- 5.2:1 Similarity vs. Heterogeneity Every social event is similar to others but also unique in an important sense. The extent to which one can presuppose that social situations are comparable and stable will determine the increasing utility of generalisations over space and time (e.g. in evaluating 'elementary skills programmes'). Means, frequencies and representativity will then become adequate criteria. If, on the other hand, the specificity of events and the flexibility of human actions become prevalent (and this depends on the complexity of the situation as well as on one's own interest), it will be more useful to portray the idiosyncracies of selected instances. The 'pathological case' (e.g. the slow learner) may even be more interesting than an average one. Divergence rather than frequency becomes the criterion for selecting examples.
- 5.2:2 Analysis vs. Gestalt One interest of research is describing more precisely the character of cause-effect relationships. This leads to isolating identifiable factors, relating them to other selected variables and investigating their impact over a greater number of constellations. The more influential these factors are (like social status in a rigid hierarchical society) the more successful the model; the more complex their interaction, however, the more context-dependent they become. In the second case (e.g. in social learning) we cannot assume the constancy of frame conditions and neglect or neutralize their impact. We have to represent as much complexity as possible, i.e. we have to describe instances in action and context. The more facts we want to take account of, the fewer examples we can choose. Thus, we can achieve more representativity in terms of numbers or of features of cases, only.

- 5.3 The issue of complexity (cf. 5.2:2) leads to the question of how many perspectives one can (and should) take into account. Again, two dimensions have to be distinguished.
- 5.3:1 Focus vs. Scope The more we know about an area, the more specific our questions can be (e.g. in research on reading skills). We, then, can concentrate our attention on pre-specified hypotheses that have to be tested rigorously. If we perceive our knowledge as limited, however, we have to keep the design flexible and open - for unexpected issues and perceptions. The enrichment and specification of hypotheses rather than their corroboration or falsification would be the aim of research (e.g. in cognitive learning).
- 5.3:2 Authority vs. Pluralism Not only the number of issues, but also the range of interpretations taken up show a significant difference between research designs. Sometimes criteria for judging events are widely shared. This allows for their application over different situations and makes them more accessible to an operationalization in advance (e.g. writing ability). The more controversial issues become, the more important that the design be responsive to diverging interpretations (as in the case of 'discovery learning'). Under such circumstances it becomes difficult to define a homogeneous set of criteria and to impose them on actions and statements of research participants.
- 5.4 The issue of responsiveness (5.2:1, 5.3:1, and 5.3:2) leads to the question of standardization.
- 5.4:1 Reliability vs. Situational Validity This is a well-known tension in classical research which becomes more significant, however, under conditions such as described in 5.3:2. If one is interested in observer-independent descriptions, one has to select surface features that can easily be allocated to pre-specified categories. This pre-supposes a well established explanatory power of surface indicators, i.e. a successful theory. In many cases, however, there is no simple one-to-one relationship between the deep (social, psychic) structure and overt behaviour. Understanding

an event then implies understanding the specific symbolic meanings, underlying norms and interpretative patterns in a certain situation (e.g. moral actions). In these cases it becomes important to achieve some sort of situational validity by becoming accustomed with the hidden intentions and assumptions that reveal themselves to an outsider through intensive observations over time.

- 5.4:2 Expertise vs. Experience In 5.3:2 and 5.4:1 we have already touched upon the issue of expert authority. Specialization is necessary. It can become successful only, though, as long as it does not lose contact with other specialized traditions and with the everyday world, in particular. It may be possible to increase abstract knowledge more rapidly by concentrating on technical precision. This is a useful perspective for basic research, for example. On the other hand, this knowledge will not become effective (i.e. general knowledge) if it cannot be understood by the people who have to base decisions on it. Thus, in defining the degree of technicality required, evaluation has to give priority to the needs of audiences and their ability to handle information.
- 5.5 I do not want to comment on these dichotomies at length. However, it seems apparent to me that a tendency towards the first pole in each dimension will relieve the practitioner from responsibility, by providing generalisations that reduce the complexity of everyday situations. They give security by focusing on particularly important factors or widely applicable models. The second option in each dimension draws attention to the risks of such generalisations and reductions. In particular, they make us aware of the problems of translating models into practical decisions. Moreover, they try to balance the power of experts by giving more weight to the values and interpretations of laymen.
- 5.6 With respect to the area of work described in section 5.3, there are three roles that the project team has to consider for achieving a balance between different expectations:

- A study role: i.e. the team would attempt to understand problems related to the curricula in particular settings and to build interpretations out of close observation of such instances.
- A support role: i.e. the team would offer co-operation to the people and agencies they study, of the kind which has to be negotiated so as to be useful to participants and not to interfere with the study role of the team.
- A dissemination role: i.e. the team would take over the responsibility for making experience gathered during the course of the study available to others interested.

It should be noted that not all forms of intervention would fit the tasks comprised by the study role. There is a tension, for example, between commitment to a specific political or pedagogical mission and the impartiality needed for a comprehensive documentation of actions: the 'honest broker' could become suspect if he co-operated with one party more intensively than with another, and this might lead to a loss of access to some interpretations that might be important. The features of the team's role, therefore, would have to be negotiated with each group in advance. Problems still might arise between different groups or programmes if they relate to each other within the system (e.g. one group co-operating might be a school, another a state agency in this province and a third the development group). This implies that the team cannot accept a change-agent role. The amount of co-operation has to be limited so as not to interfere with the research responsibility.

One of the significant features of the project is that the subject and the medium of study are closely related. It is concerned with the transfer of ideas and experience from adults to children, from developers to users and from the evaluation programme to reference groups. Thus the team deliberately takes over the responsibility for making experience accumulated in the course of its work accessible to all parties interested (dissemination role). This may be important within a specific programme (exchange of perceptions and interpretations), between different programmes (cross fertilisation of ideas and experiences) and finally between the participating programmes and other groups. It is one of the main reasons for

establishing a separate team that the needs of transfer have to be permanently considered.

I would like to draw attention to two research traditions that seem to be particularly relevant to the tasks implicit in this cluster of roles: the ethnomethodological and the action-research paradigms. The first offers suggestions for how to transfer experience from one setting to another. The second tradition is relevant to the problems of linking a study role to a support function. I have outlined a possible relationship between these traditions in an earlier paper. (8)

## 6. A tentative network of activities

An evaluation design has to be flexible. It should be adaptable to different situations and perceptive of changes in the area under study. However, to be able to learn from our experiences in the course of the study, we need a framework within which to interpret such experience. We need a wall against which to play the ball. I see the following outline as an instrument of that kind. It is a preliminary concretisation of the principles outlined in the previous section.

- 6.1 The familiarisation stage is characterised by a simple dilemma: to be useful to the development groups, feed-back of data has to start as soon as possible - to be credible and valid it needs thorough preparation. Though I have argued in section 5.3 that evaluation has to focus on the needs of users of the programme, this does not mean that we can afford missing the chance to improve the curricula-on-paper as far as possible. Thus we need an instrument that can fulfil the following tasks simultaneously:
- it should familiarise the team (and other participants in the evaluation) with the programmes;
  - it should help to locate key issues for the investigation phase;
  - it should help identify criteria for the interpretation of data;
  - it should provide significant feed-back to the development groups very early on;

- it should contribute to the development of a problem solving capacity in the system.

The only device I know of that can meet these requirements and that will work without too long a preparation is curriculum analysis. The best instrument available at present is the 'Sussex Scheme' developed by Eraut et al. (1975). Curriculum as suggested by the Sussex group can best be done by groups of people representing different backgrounds and different kinds of interest in the programmes. Such groups would review the programmes in a workshop setting, building up a number of perspectives that seem to be relevant to different user groups. In Scriven's terms this activity would count as 'intrinsic evaluation'. It would concentrate on aspects of the materials that can be evaluated without looking at its implementation results (e.g. what kind of normative decisions have influenced the programme? which learning theory, view of the subject matter, assumptions about relevance, etc. have been the basis for programme development? is the programme manageable, what kind of costs are involved, how innovative is it, etc?). Of course, these judgements may have to be altered during the evaluation of programme implementation. But the analysis of materials can provide useful hypotheses and questions. And it can reveal reactions of audiences to central features.

Participants in these workshops would be members of the evaluation team, teachers, parents, administrators, teacher education personnel and academic specialists. Some groups might be mixed, others might include people of certain groups only.

Since these workshops would have to be rather intensive, not many people can participate. It therefore might be useful to complement these reviews with more general surveys eliciting issues and criteria from teachers and parents. Widely read journals like 'The Primary School' and 'Parents' offer a medium for approaching these groups and for giving feed-back. At the same time, this would help to make future users aware of the forthcoming innovations.

- 6.2 The systematic investigation stage focusses on the implementation of the programmes in selected settings. The task is to approach more extensively and intensively issues that have tentatively been defined during the first stage. The design, however, still has to be perceptive to variations and changes and it should be possible to make use of additional forms of describing and analysing what happens.

The main goal is to gradually build up case histories of children, teachers, whole classes, schools, and perhaps even some aspects of the programme as a whole. Such archetypes should cover a sufficient range of implementation possibilities and problems that they can serve as signposts for future users of the programmes in exploring the potential of the curricula with respect to their own situations.

- 6.3 The final reporting stage will help to pull together and to reassess presentations that have been used during the second phase. An attempt will have to be made to write separate reports appropriate for different audiences. This should feed into systematic dissemination of results. To be effective the infrastructure for dissemination has to be built up in the course of the work and appropriate strategies for dissemination have to be explored perhaps through participation in in-service courses for teachers.

7. Some personal afterthoughts (December 1976)

Working on EVI CIEL has been a fascinating experience, but nevertheless wearisome. It might be more useful here to mention some of the problems than to comment on the benefits of this enterprise so far.

- 7.1 There have been problems of a political nature. EVI CIEL came into existence at a time when educational innovation had lost its impetus - not only in West Germany. But in the province of Northrhine-Westphalia the school and kindergarten authorities were particularly anxious to get their institutions settled after a five year period of public debate and experiments with new ideas. Moreover, there was a hidden battle going on about the responsibility

for pre-school education between the school ministry on the one hand and private associations, especially the churches and the welfare ministry, on the other; thus, we had to face suspicions about the role of a project working in the five to seven years age range. Moreover, the significance of the evaluation, both in terms of its focus and its magnitude, was interpreted as a potential disturbance rather than as a reason for supporting public investment. On the other hand, there were individuals at different levels, especially in schools and in the regional administration, who backed the project and helped it to survive the one year negotiation period. The advantage, then, may be that participants do not see the project as being forced on them from above. In the area of schools, the approval of the ministry does legitimize the evaluation in legal terms, but it is not the motivating factor. The kindergarten authorities still have not granted access to their institutions.

- 7.2 Another problem has been the understanding of "evaluation" in this project. The central administration, for example, has demanded that the "quality" of the curricula be demonstrated before embarking on their evaluation in practice. In a sense we were asked to tell in advance what was supposed to be the result of our work. In particular, it was difficult to make administrators recognize the value of curricula as being hypotheses for experimental teaching; they were more interested in consolidating the present situation than in increasing the range of options open to teachers. Researchers, on the other hand, have interpreted the evaluation proposal as careless, amateurish, or just ingenious in hiding our political aspirations. Not only for them, but also for members of the evaluation team it was difficult to live with this pluralistic and evolutionary approach, i.e. with the strain of accepting divergent interests and perceptions in the course of the evaluation instead of declaring one's loyalty to one point of view from the beginning.
- 7.3 I was very much surprised, however, by the supportive stance of some development groups and the disinterest of others - and the

lack of anxiety. On the other hand, our work was heavily delayed by the late arrival of the materials from the development groups (though deadlines had been postponed several times in some cases). Thus, we had to start the curriculum analysis work with other materials available for achieving at least some of the goals of the preparation phase (we still have not received the conceptual framework from three of the groups). This meant a lot of reconstructive work which was very time-consuming, but also forced us to analyse the materials very carefully.

- 7.4 The lack of background materials made us more prone to accept a support role and responsibility for interpreting the curricula to administrators, teachers and other audiences. We could not let them speak for themselves. This shift in our role which was necessary to get access to the field was supported by some members of the evaluation team who favoured a more interventionist stance as well as by expectations from the teachers themselves who wanted to exploit our expertise. These expectations are clearly linked to the traditional role of in-service trainers and have already influenced our analysis workshop sessions. Moreover, this was seen as one side of the bargain by the teachers who were interested in co-operating with the project in the trials.

Though I still feel that there are reasons for declining such a role to a certain extent with teachers, I could not do this with children. This, perhaps, is the most moving experience: individual children expressing their need for personal attention and particular help because of their special difficulties. Although helping them subtracts energy and attention from planned activities, it helps us to remain aware of the main purpose of our work: To promote a better understanding of learning difficulties in educational institutions and to provide aides for implementing such an understanding in everyday situations. On the other hand, the past fifteen months have strengthened my belief that individual perceptions, situational differences, the need for personal and institutional survival, gradual developments, informal relationships, and common

experiences are of central importance in the process of educational innovation. Clearly, there are objective frame factors; they become effective in different ways, however, through the mediating interpretations of different people. In this context, the approach of an 'evaluation through dissemination' as outlined earlier in this paper seems to have some potential as it attempts to relate experimental work in selected situations to the clarification of needs of different audiences from the beginning.

Footnotes

- 1 Cf. Stake, R.E. Program Evaluation, Particularly Responsive Evaluation. In: New Trends in Evaluation. Reports from the Institute of Education, University of Goteborg, No. 35, 1974. and Stufflebeam, D. Meta-Evaluation. Evaluation Center, College of Education, Western Michigan University. Occasional Paper No. 3, Kalamazoo, 1974.
- 2 For a fuller account see MacDonald, B. The evaluation of the Humanities Curriculum Project: a holistic approach. In: Theory into Practice, 3, 1971. 163; Briefing Decision-Makers. In: Schools Council (Ed.) Evaluation in Curriculum Development: Twelve Case Studies, London 1973; Evaluation and the Control of Education. In: SAFARI 1974 Some Interim Papers, 7.
- 3 Brugelmann, H. 'Towards Checks and Balances in Educational Evaluation: on the use of social control in research design' In SAFARI 1974 Some Interim Papers
- 4 For a fuller elaboration see Simons, H. Innovation and the Case-study of Schools. In: Cambridge Journal of Education, 3, 1971. 118
- 5 See Levine, M. Scientific Method and the Adversary Model: Some Preliminary Suggestions. In: Evaluation Comment, 2, 1975. 1
- 6 I should like to thank J. Rudduck who has contributed a great deal to an earlier version of part of this section that we had prepared for a joint project proposal in the field of innovation research: Encounters in Curriculum Innovation. Mimeo CARE, Norwich 1974.
- 7 For a fuller account see Walker, R. The Conduct of Educational Case-Study. Ethics, Theory and Procedures. In: SAFARI 1974. Some Interim Papers. 67
8. op. cit. 3, especially p. 45

Bibliography

- Bannister, D. and Fransella, Inquiring man. The Theory of Personal Constructs: Harmondsworth, 1971.
- Elliott, J. and Adelman, C. Theory from Practice. A final report to The Ford Foundation. mimeo, Ford Teaching Project, Centre for Applied Research in Education, University of East Anglia, Norwich, 1974.
- Eraut, M. Review of Some Recent Evaluation Studies in the U.K. In: Schools Council, 1975 (in prin
- Eraut, M. et al. The Analysis of Curriculum Materials. University of Sussex Education Area. Occasional Paper 2, Brighton, 1975.
- SAFARI Innovation, Evaluation, Research and the Problem of Control Some Interim Papers. Workshop Curriculum No. 1, Norwich/Tubingen 1974.
- Stenhouse, L. An Introduction to Curriculum Research and Development London: Heinemann Educational Books, 1975.