THE USE OF INNOVATION AND PRACTICE PROFILES IN THE EVALUATION OF CURRICULUM IMPLEMENTATION

Jan van den Akker and Joke Voogt
Faculty of Educational Science and Technology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Introduction

Most generic curriculum reform efforts have to deal with a gap between the innovative aspirations of the initial designers and the daily reality of the intended audience of teachers. That tension is not alarming in itself. One might even say that without it no compelling reason for starting development work would exist. Unfortunately, many evaluation studies on the implementation and impact of curriculum development projects show that this discrepancy does not decrease over time. Apparently, not much improvement is made in detecting and reducing potential implementation problems.

This article presents some conceptual and instrumental guidelines for dealing with these problems, focusing on the use of 'profiles' during evaluation of curriculum materials.

The paper starts with an introduction on the functions of exemplary curriculum materials and their possible representations, on the long road from original designers' ideas to effects of student learning. Next, we will explain the concepts of innovation and practice profiles. We will then provide guidelines for the development and use of such profiles, based on previous research experiences, and illustrated with some specific examples. Finally, we will reflect on the advantages and limitations of working with profiles.

Functions of Curriculum Materials

As has been described by van den Akker and Verloop (this issue), generic curriculum development in The Netherlands usually results in two different kinds of products: general curriculum programs (for planning at school level) and exemplary curriculum materials (for use in instructional practices). Both kinds of products are supposed to be interrelated: the materials offer concrete and illustrative elaborations of the general programs which, in turn, provide the framework for broad categories of materials.
The importance of concrete materials for the realization of content and pedagogical changes at the classroom level is undisputed in our country. However, all curriculum products (as published by SLO, the National Institute for Curriculum Development) have an exemplary, not obligatory, status. Everyone is free to use these products and to adapt them to their own situation and needs. The SLO materials (usually a combination of teacher guides and student materials) may fulfill several functions:

- **Internal validation:** designing, constructing, and evaluating lesson materials may contribute to the specification and testing of more general curriculum ideas and statements;
- **Exemplification:** the materials may serve as concretely specified examples (for other designers and publishers in the educational system) of practical and effective materialization of curriculum proposals, and
- **Training:** the materials may be used as demonstration and practice tools in (both pre- and in-service) teacher education.

These functions benefit from complete elaboration and classroom testing of the materials. This enables diverse audiences to have a clear understanding of the intended classroom practice, and it facilitates utilization in various contexts. A central assumption in this paper is that the quality (but also the credibility) of the materials derives to a great extent from a thorough evaluation on classroom practicality. An iterative design and formative evaluation approach should especially contribute to that goal. Profiles can play a useful role in such activities.

**Curricular Representations**

To get a grasp of the possible uses of profiles it is helpful to consider a typology of different curricular representations (cf. Goodlad, Klein, & Tye, 1979):

- **Ideal curriculum:** the original ideas and intentions of the designers;
- **Formal curriculum:** the written curriculum (documents, materials);
- **Perceived curriculum:** the interpretation of the users (especially the teachers) of the curriculum;
- **Operational curriculum:** the actual instructional process in the classroom, and
- **Experiential curriculum:** the reactions and outcomes of the students.

In many evaluation studies the emphasis has been on comparing the ideal (and sometimes the formal) curriculum with the experiential curriculum, often revealing a large gap. Although such results are an important signal, it is often difficult to draw conclusions about the nature of the causes, let alone to suggest possible solutions, without having accurate data about intermediary stages in the implementation process. Without such information it is also too simplistic to blame specific groups (for example, 'naive' designers or 'reluctant' teachers) for disappointing results.

An insightful illustration of these problems has been offered by Sabar (1986), who made a careful analysis of the route of an Israeli science curriculum innovation, revealing dilutions and distortions on several levels. It appeared that the developers, while constructing curriculum materials, had faced serious problems in reaching a clear and consistent operationalization of their ideals. Moreover, they presumed that
potential users might experience difficulties in realizing the proposals in classroom practice. For those reasons, the developers themselves reduced their innovative aspirations beforehand. As a consequence, the written curriculum only faintly reflected the initially proclaimed ideals. In those cases where the materials arrived in the schools, all kinds of persistent problems hindered an acceptable realization of the instructional process.

In situations like these, there is a risk that student outcomes are measured in classrooms where the curriculum in question has hardly, if at all, been implemented as originally intended (cf. Charters & Jones, 1973). Moreover, it was rather common in the evaluation practices of the 60s and 70s (and it still is not uncommon, especially in large scale assessment studies) for student outcomes to be measured with standardized test instruments that are rarely attuned to specific innovation goals (Walker & Schaffarzick, 1974). However, with the increasing attention for implementation problems - as summarized and stimulated by the work of Fullan (Fullan & Pomfret, 1977; Fullan, 1982) - the need has grown for more direct and precise information about what actually happens in the classroom as a result of the use of new curriculum materials. The study of the operational curriculum, especially, has come more to the forefront.

Measurement Approaches of the Operational Curriculum

Ideas for the procedure and instruments of using profiles for measuring the operational curriculum emerged from several earlier North American innovation studies:

- The work around the 'Concerns Based Adoption Model' in Austin (Texas), focusing on the development of concerns of individuals teachers and their levels of use regarding an innovation (Hall & Loucks, 1977). Within that framework an interview instrument to measure implementation was designed, assessing the various 'configurations of an innovation in practice (Heck, Stiegelbauer, Hall & Loucks, 1981).

- The work of Leithwood and Montgomery from the Ontario Institute for Studies in Education (Toronto, Canada). Elaborating on a distinction in a number of curriculum dimensions (Leithwood, 1981), a procedure for measuring program implementation was designed. For each dimension different representations were described in an 'innovation profile', based on the original intentions of the developers, and varying from complete correspondence to none. Interviews with individual teachers led to an analysis of their stage of implementation with reference to the different dimensions in 'user profiles' (Leithwood & Montgomery, 1982).

- A comparable approach has been used in the DESSI study (Dissemination Efforts Supporting School Improvement) by Crandall and colleagues (1982), who used 'practice profiles to draw up the degree of program implementation (see Loucks, Bauchner, Crandall, Schmidt & Eiseman, 1982).

All these approaches can be characterized as fruitful contributions to systematic evaluation procedures regarding curriculum implementation. Our approach in a number of research projects at the University of Twente was inspired by these procedures, but differed in two respects. First, the primary method of data collection in our research consisted of direct observation (sometimes including videotaping) of lessons in the classroom practice instead of self reports by teachers via interviews after the lessons.
The major advantage is the more direct (and repeatable, when video is used) access to actual lesson events, contributing to the validity and reliability of the data collection and analysis. Second, we chose a more specific and detailed description of curriculum components and elements than the rather broad categories and predetermined stages in the studies mentioned above. That procedure allowed us to pay more attention to the specific intentions of the curriculum designers and their specific interpretation in classroom operations by the teachers.

In the next section we will provide more detailed definitions of our procedure and instruments, and illustrate them with examples from a number of studies.

Applications of Innovation and Practice Profiles

Since the 80s we have been using innovation and practice profiles in several studies (van Aert & van den Akker, 1990; van den Akker, 1988; Keursten, 1994; Kuiper, Kühne & van den Akker, 1989; Voogt 1993). An overview of these studies is presented in Figure 1. The studies differ in domain and scope of the curriculum and in research focus. For example, both a two-hour elementary science lesson as well as a nine-hour social studies curriculum unit were investigated using innovation and practice profiles.

<table>
<thead>
<tr>
<th>Study</th>
<th>Scope/curriculum Domain</th>
<th>Design</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>van den Akker, 1988</td>
<td>two hours / primary science</td>
<td>experimental</td>
<td>effect of written teacher material, using procedural specifications, on lesson execution of teachers</td>
</tr>
<tr>
<td>Kuiper et al., 1989</td>
<td>nine hours / secondary social studies</td>
<td>descriptive/evaluative</td>
<td>use of exemplary curriculum materials in lesson practices</td>
</tr>
<tr>
<td>van Aert &amp; van den Akker, 1990</td>
<td>32 lesson series / primary science</td>
<td>descriptive/evaluative</td>
<td>determining discrepancies between intended, perceived, and operational curriculum</td>
</tr>
<tr>
<td>Voogt, 1993</td>
<td>ten hours / secondary science hour</td>
<td>case study</td>
<td>effect of courseware characteristics on teacher behaviour</td>
</tr>
<tr>
<td>Keursten, 1994</td>
<td>five hours / geography</td>
<td>experimental</td>
<td>effect of teacher material (written and videotape), using procedural specifications, on courseware use</td>
</tr>
</tbody>
</table>

Figure 1: Studies Using Innovation and Practice Profiles

Some studies used innovation and practice profiles for operationalizing implementation as a dependent variable in an experimental design (van den Akker, 1988; Keursten, 1994) in which the experimental group of teachers was compared with a control group. Other studies (van Aert & van den Akker, 1990; Kuiper et al., 1989) had a descriptive and evaluative character. In these studies the degree of implementation by teachers of curriculum materials was established using innovation and practice profiles. Voogt (1993) used innovation and practice profiles in a case...
study approach, in which she emphasized the way teachers mastered specific pedagogical content knowledge necessary for using inquiry-based courseware for science education.

How do we define these profiles? An innovation profile of a (unit within) a curriculum consists of operational descriptions of:

1) the main components by the designers consider essential for the realization of the intended curriculum (unit), and
2) the extent to which different ways of realizing these components correspond with the intentions of the developers.

In drawing up such a profile, first, each of the essential components is operationalized in lesson elements that are considered necessary for a 'more or less satisfying' (or 'acceptable') implementation from the designers' perspective. These are called 'threshold elements. Furthermore, for each component lesson elements are operationalized which represent possible realizations of the curriculum (unit) at stake. These lesson elements may either strengthen ('ideal' elements) or weaken ('unacceptable' elements) the implementation of the curriculum according to the developers' intentions. Moreover, scores may be assigned to every threshold, ideal and unacceptable lesson element. In assigning these scores, ideal and unacceptable elements can only be scored if the threshold score is met. Hence, the threshold score is a conditional score. The rationale for assigning scores is that it offers the opportunity of weighing the relative importance of different components and elements. The complete innovation profile as such represents the ideal curriculum of the developers.

The use of innovation profiles focuses on the actual use of (innovative) curriculum materials in the classroom, the operational curriculum. While the teacher is primarily involved in the operational curriculum, the profile mainly emphasizes the (instructional) behaviour of the teacher. The extent to which teachers realize the innovation profile can be established by crediting scores based on classroom observations. This results in a practice profile per teacher, reflecting the degree of actual implementation of the innovative curriculum. In this way a quantitative measure of the degree of implementation is obtained. In the next sections we will discuss the use of innovation and practice profiles, with special reference to the studies of van den Akker (1988), Voogt (1993) and Keursten (1994).

Some Examples

The innovation profile consists of a limited number of components (usually four or five). Together they reflect the core of the curriculum innovation. For an inquiry-based science curriculum, Voogt (1993) considered the following to be essential components of the innovative curriculum: planning an investigation; executing an investigation; building a theoretical framework; deriving physical rules; and general skills (such as stimulating cooperation within a group of students). Van den Akker (1988) distinguished among context, discovery, inquiry, and evaluation as essential components of a elementary school science curriculum. Keursten (1994) used lesson start, teacher guidance, independent work, and lesson summary as the basic characteristics of a computer integrated geography curriculum.

The essential components are refined by operationalizations in threshold, ideal and unacceptable elements. The different operationalizations of threshold elements
show the main concept of the innovation, such as the emphasis on specific pedagogical content knowledge of the teacher for inquiry-based science in Voogt's study, and the importance of general instructional strategies in Keursten's approach.

The Development of Innovation Profiles

The development of the profile and the assignment of scores to the different parts of the profile needs ample deliberation. In some studies (e.g., van den Akker, 1988; Kuiper et al., 1989) the operationalization and assignment of scores took place in close cooperation with the developers of the materials in order to guarantee a valid interpretation of the latters' intentions. In these cases the development of the profile and the assignment of scores were the result of deliberation within the curriculum development team and between developers and evaluators/researchers. The primary task of the researcher was to stimulate (e.g., by asking questions for clarification) and manage (e.g., by editing drafts of the profile) that process of discussion and negotiation. This was done by identifying relevant intentions of the developers via analysis of previous project documents and through interviews with team members. Afterwards, group discussions were held to discuss and refine the provisional statements. The ultimate result of this deliberation was a consensually operationalized ideal curriculum.

In other studies (Keursten, 1994; Voogt, 1993), in which the roles of developer and researcher were combined, the profile was elaborated by researchers themselves (often based on literature research). In these studies validation of the profile was done by carefully checking and revision through field try-outs and consultation with experts in the specific curriculum domain.

The assignment of scores to the different elements of the curriculum has, of course, consequences for the interpretation of the collected data. In some studies (van den Akker, 1988; Keursten, 1994; Kuiper et al., 1989) the total threshold score to be obtained was about 55% of the total score, thereby reflecting a 'more or less satisfying' execution of the curriculum. This appeared helpful in valuing the interpretation of results, because it referred to the grading system in (Dutch) schools. Other studies used a much lower threshold score. For example, Voogt (1993) used a 17% score as threshold, which indicated a minimal way of meeting the requirements which could be set for teaching innovative inquiry-based science. This approach appeared to be very useful for case studies in which the differences between the planning and coaching of individual teachers were of primary importance.

From Innovation Profiles to Practice Profiles: Data Collection and Interpretation

When using innovation and practice profiles, the collection of data by lesson observations is essential, using a form with categories derived from the innovation profile. The advantage is a direct and systematic registration of the operational curriculum, instead of data collection via retrospective information by teachers. After all, in the latter case the perceived and operational curriculum can be easily mixed up.

Some studies added video- or audiotaping techniques for direct observations, which made it possible to follow the teacher more closely in his/her execution of the
lessons. This is particularly important when the innovative curriculum is characterized by a lot of individual or group work of students. Teachers' interventions can be followed more precisely in this way. Furthermore, from a research perspective, audio- and videotaping is helpful in improving (via training) and assessing the reliability of scoring because of easy reproducibility. Also, in van den Akker's study the viewing of videotaped lessons appeared to be very stimulating and clarifying in discussions with the designers about the innovation profile.

From the studies of Voogt and Keursten it became clear that audiotaping is a good alternative to videotaping in situations where already quite a lot of apparatus is in the classroom, which is the case in a computer-integrated curriculum. After training, the interrater-reliability between independent raters appeared to be high in all three studies.

Most studies also applied other techniques for data collection as a supplement to lesson observations, such as interviews with teachers (giving information about the perceived curriculum), testing of students or analyzing students' worksheet (providing information about the experiential curriculum).

In the studies of van den Akker and Keursten the quantification of the data made it possible to compare an experimental with a control group. In both studies the use of specially designed teacher materials with concrete procedural suggestions for preparing and executing the innovative curriculum were compared with common teacher materials. It appeared that the teachers working with the experimental materials outperformed the teachers in the control group with respect to the implementation of the curriculum. Moreover, Keursten found that the students of the experimental teachers obtained higher results in the test administered after the curriculum compared with their peers in the control group.

Voogt's study on inquiry-based science illustrates the way in which innovation and practice profiles can contribute detailed information about the different planning and coaching behaviour of teachers. The teacher materials (with specific pedagogical suggestions) played a crucial role in the degree of implementation. It became clear that teachers who used the teacher guide reached a higher degree of implementation, as established in the practice profiles, compared with teachers who did not use the teacher guide, but chose only to use student materials for lesson preparation and execution. Moreover, the partition in type of interventions (planning an investigation, executing an investigation, building a theoretical framework, deriving physical rules, and general skills) of the teachers who used the guide agreed with the 'ideal partition of interventions. The other teachers spent too much of their interventions on general skills. In combination with information from the interviews with the teachers and an analysis of students' worksheets, the results could explain this different teacher behaviour, providing insight into the obstacles teachers have to overcome when implementing innovations.

Discussion

In discussions on the merits of applying these kind of profiles, it is sometimes argued that this evaluation method reflects a specific vision on innovation approaches, that is, a preference for 'fidelity' above 'mutual adaptation' or 'enactment' perspectives (cf. Snyder, Bolin, & Zumwalt, 1992). In fact, this measurement approach does seem
most appropriate if one is interested in the correspondence between the original intentions of the curriculum designers and the teachers' actual classroom behaviour. However, such an interest does not necessarily contradict the importance of active adaptations by the teacher, exploring the potential of curriculum materials (cf. Ben-Peretz, 1990).

From our studies we concluded (van den Akker, 1994) that a stepwise strategy is advisable in which teachers are stimulated to engage in new teaching situations with the help of materials that contain a lot of procedural specifications. Such materials can reduce the personal uncertainty that is almost inevitable at the early implementation of substantial changes. In addition, they may reduce the chances of premature adaptation occurring, even before concrete (and hopefully successful) experiences with the essentials of the innovation and its effect on students have been gained. This learning-by-doing can then be reinforced by reflection, exchange of experiences with colleagues, feedback and follow-up support. Adaptive activities, for example in additional and site-specific materials development, seem most appropriate in those later stages for strengthening the individual and social learning processes of the teachers involved.

Whatever the specific innovation approach, during the process of implementation evaluative information from practice profiles can be very helpful for developers, both in revising original intentions as well as improving materials and focusing of additional support activities.

This conclusion suggests that profiles are not only useful in summative evaluation, but also, and perhaps even more, in formative stages of development projects. For formative purposes (in view of improving draft versions) informal and efficient methods of data collection and analysis are usually more feasible and necessary than in summative evaluations. For example, in formative situations there is little need to attach scores to the different elements of an innovation profile (this being a time consuming task). Also, less observation sites are needed for obtaining information that can generate improvement suggestions. The profile itself may be more tentatively formulated and less detailed. Such a provisional profile may not only be helpful as a tool for observation, but also for prestructuring other evaluation instruments like interview checklists and teacher logbooks. Experience with such combined evaluation methods, with the innovation profile serving as a focal point, has been gained in several projects. Moreover, some designers even found the task itself of making a profile to be useful and stimulating in early stages of development activities because it challenged them to articulate their intentions from the very beginning. Although it is a demanding task, the purposiveness of the development activities may certainly profit from such 'profiling'.

Our conclusion, to be validated in further research, is that the use of profiles in curriculum evaluation has a good potential for revealing, but also increasing, the consistency between images of curriculum designers, perceptions of teachers, and classroom realities. And that seems a relevant aim in the often rather elusive processes of curriculum design and implementation.

Note

1. For examples of profiles developed by Voogt and Keursien, apply to the authors of this article.
References


Keursten, P. (1994). *Courseware-ontwikkeling met het oog op implementatie: de docent centraal* [Courseware development with a view to implementation: focusing on the teacher]. Enschede: University of Twente, Faculty of Educational Science and Technology.

Kuijper, W., Kühne, E., & van den Akker, J. (1989). *De implementatie van een maatschappijleercurriculum* [The implementation of a social studies curriculum]. Enschede: University of Twente: OCTO.


Voogt, J. (1993). *Courseware for an inquiry-based science curriculum. An implementation perspective*. Enschede: University of Twente, Faculty of Educational Science and Technology.


The Authors

JAN VAN DEN AKKER is Associate Professor in the Department of Curriculum, Faculty of Educational Science and Technology, University of Twente, Enschede, The Netherlands.

JOKE VOOGT is Assistant Professor in the Department of Curriculum, Faculty of Educational Science and Technology, University of Twente, Enschede, The Netherlands.