Supporting medical teachers' learning : redesigning a program using characteristics of effective instructional development

Min-Leliveld, M.J.

Chapter 6

Conclusions and Discussion

6. Conclusions and Discussion

6.1 OVERVIEW OF THE STUDY

Instructional development programs can be important tools to facilitate medical faculty in their role as teachers. Although there is a large body of literature on the effectiveness of such programs, various reviews reach different conclusions about their impact (e.g., Levison-Rose & Menges, 1981; Prebble et al., 2004; Steinert et al., 2006). These differences in reported effectiveness might be related to differences in the design characteristics of the programs. Literature on the characteristics of effective instructional development programs (knowledge-for-practice) forms an important source of information for the design of such programs. The characteristics described in the literature may help both teacher educators and program developers, but are often formulated without taking context and specific conditions into account. This makes it difficult to see the relevance of the results, and hinders the implementation of those characteristics in actual teaching practice. The knowledge of teacher educators (practical knowledge) and the preferences of teachers can be useful to bridge this gap between literature and implementation, and to identify characteristics that are not only effective for teachers' learning but are also relevant and appealing to teachers and teacher educators. This leads to the central research question of this thesis:

> *What characteristics of effective instructional development are appealing to medical teachers and relevant for the design of instructional development programs for medical teachers, and what do these teachers learn from a specific program that takes into account those characteristics?*

In order to answer this question we conducted two studies. In the first study, described in Chapters 2 and 3, selecting those aspects that characterized effective instructional development programs was the center of attention; in the second study, reported in Chapters 4 and 5, we focused on teachers' learning in an adapted instructional development program, in which the characteristics of effective instructional development programs selected earlier were used as a framework.

In the first study teachers and teacher educators were asked to indicate which of the 35 effectiveness characteristics derived from the literature on effective instructional development (Guskey, 2003; Steinert et al., 2006) they considered important in instructional development programs. For the study described in Chapter 2 we sent out questionnaires to medical teachers in the Leiden University Medical Center, asking which characteristics would appeal most to them if they were considering participation in instructional development programs. In Chapter 3 we reported a study in which teacher educators from all eight medical schools in the Netherlands were interviewed about the characteristics they considered most relevant for teachers' learning.

In the second study we used the characteristics collected in the first study as a framework to analyse a successful instructional development course (Train the Trainers) in order to understand its effectiveness. Secondly, we developed an additional instructional development course, based on the framework and the information derived from that successful course. We studied teachers' learning in this program in two ways: first (Chapter 4), we used an evaluative questionnaire to report the effects of the program, using Kirkpatrick's four levels (1994). Second (Chapter 5), we constructed an in-depth visualization of the learning processes of four teachers in the various sessions of the adapted instructional development program, using Clarke and Hollingsworth's model. The learning diagrams we constructed on the basis of the interviews informed us about teachers' learning in the various components of the program.

In this concluding chapter we will first describe our conclusions with respect to the research questions, and discuss the outcomes of both studies. Next, we will discuss the strengths and weaknesses of these studies. Finally, we will provide some suggestions for future research and indicate implications of our findings for teachers, teacher educators, program developers, and policy makers.

6.2 CONCLUSIONS

In this section we describe the conclusions for each of the research questions, first for the study into the characteristics of effective instructional development, and second for the analysis of a successful course, the design of an adapted program, and teachers' learning in this program.

6.2.1 First study

In the first study two specific questions were considered. The first of these, addressed in Chapter 2, was:

> *Which characteristics of effective instructional development are most appealing to medical teachers when they consider participating in*

instructional development, and what are the factors underlying these preferences?

Regarding teachers' preferences three underlying factors were identified: (a) facilitating collaboration in educational improvement, (b) individual development as a teacher, and (c) evidence-based education. The first factor was somewhat heterogeneous in character, combining various items related to the design of instructional development programs. A number of these items had to do with collaboration and interaction with colleagues, while others were related to taking the working context into account in the design. The second and third factors were easier to label, using the wording of the high-loading items. Teachers' individual development was the central aspect in the second factor: learning from one's own teaching practice by means of reflection and feedback. In the last of the three factors items were combined that took the evidence from educational research as a foundation for instructional development.

The results showed that although almost all characteristics were important to medical teachers when they considered participating in instructional development, there were marked differences in preferences between individual teachers. None of these differences could systematically be related to various background variables such as time allocated to education or amount of experience, so we assumed that they originated from personal differences in preference. Seven characteristics were found to be relatively more important than the others. Three out of those seven concern the design of instructional development, and four refer to teachers' individual development.

The second specific research question addressed in the first study, covered in Chapter 3, was:

Which characteristics of effective instructional development do teacher educators consider most relevant when designing actual instructional development programs in medical schools?

To answer this research question we conducted semi-structured interviews with teacher educators from all eight medical schools in the Netherlands. All of them were experts in the design and implementation of instructional development programs. As a result of their experience as teacher educators they had practical knowledge about educational practices, about learners (i.e., medical teachers), and about how these learners learn. This knowledge is relevant for the design of instructional development programs. To explore the teacher educators' practical knowledge we asked them to identify relevant characteristics from the list compiled in the first study. Furthermore, they were asked to describe effective instructional development in their own

medical school, focusing on a specific instructional development program that they themselves considered to be an example of a successful "best practice". Fifteen characteristics were selected by the teacher educators as most relevant for teachers' learning.

The interviews were used to identify contextualized specifications of those fifteen characteristics. For example, "inclusion of alternative practices" was described by teacher educators as including new ideas in instructional development programs, such portfolios, observation of teaching sessions so that teachers can receive feedback, and new formats such as online sessions, role play, individual coaching, and peer group sessions.

6.2.2 Overview: which characteristics were selected by both teachers and teacher educators?

An overview of the selected characteristics of effective instructional development is displayed in Figure 6-1. It shows the sixteen characteristics identified by both the teachers and teacher educators: seven characteristics were selected by the teachers as most appealing, and fifteen were selected by the teacher educators as most relevant. Six characteristics were selected by both groups.

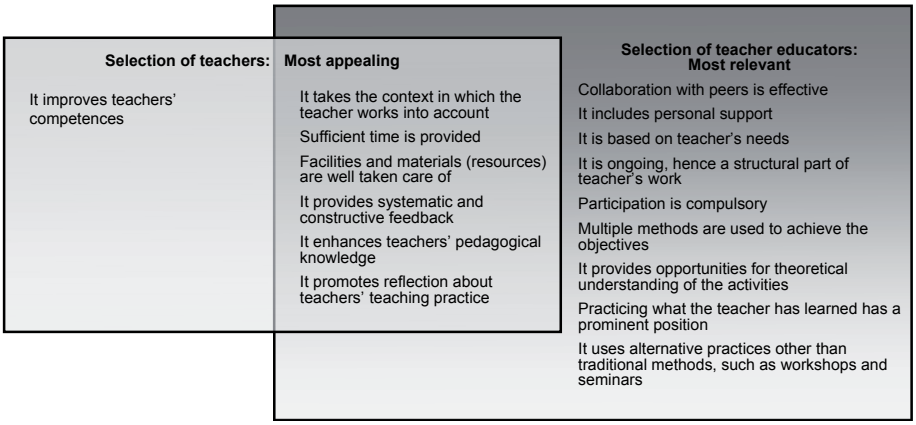

Figure 6-1. Overview of the characteristics selected in both studies

6.2.3 Second study

The first specific research question covered in the second study was:

Can characteristics of effective instructional development be used as a framework by which to understand why a specific short course is successful?

The successful short course "Train the Trainers" was selected as the subject of this study, because it is a popular short course in medical education in the Netherlands as well as in other countries such as the United Kingdom and Denmark. The course is generally rated highly satisfactory by participants, and is a good example of a popular instructional development program in higher medical education, probably because of its design as a short workshop. In order to find indications for the reason of its success, we analyzed the course using the characteristics found in the first study as a framework. Most of the characteristics of effective instructional development programs (ten out of sixteen) were found to be well-implemented in this course, in particular those selected by the teachers as most appealing.

We found that six characteristics from the list of sixteen had not been implemented well in Train the Trainers course. These six characteristics were related to the format of the course (personal support, use of alternative practices, inclusion of theory and practicing) and the time needed to follow the course (sufficient time and being ongoing). Interestingly, these six characteristics were among those that the teacher educators had selected as most relevant for teachers' learning, while most of the six were absent from the list of characteristics selected by the teachers as most appealing. Hence, it seems that the course is more in line with what is appealing to teachers than with what the teacher educators consider most relevant for teachers' learning.

The second research question covered in this second study was:

What do participants report to have learned from an additional course that included all characteristics selected?

An additional course (called the "Plus Course") was designed, based on all sixteen characteristics of effective instructional development (see Figure 6-1). The new program consisted of various sessions: three workshops (360° feedback session, session on video vignettes, and peer group discussion) and two one-hour web seminars (optional), scheduled over a five-month period. The Plus Course mainly focused on improving teachers' knowledge and skills concerning feedback, and on creating more awareness of their roles as teachers. A prominent aspect of the Plus Course was the great amount of time devoted to practicing with what the teacher had learned, by means of assignments that had to be completed in the daily work context.

An evaluation questionnaire geared to the learning outcomes of this new program, based on Kirkpatrick's four levels, was developed and distributed among the participants. The majority of the respondents were less satisfied with the Plus Course than with the Basic Course. They did, however, report a positive change in their learning, behavior, and the learning climate in actual teaching practice. Participants reported that they had become more aware of their role as a teacher and were more focused on their students; that they managed to create a more effective learning environment by providing more structured and more positive feedback; and that they were interacting more with students about their prior knowledge and skills.

Thus, the adapted program may be considered effective in terms of teachers' learning, since changes in learning, behavior, and learning climate were reported. There appears to be a tension between "what is best" according to the literature on effective characteristics, and "what is most desired" as measured by teacher satisfaction: although teachers reported that their teaching behavior had changed on the basis of the program, they were less satisfied with the Plus Course than with the Basic Course.

The final research questions covered in the second study were:

> *How can teachers' learning in the adapted instructional development program be visualized? What kind of learning sequences can be recognized in the various components of the program?*

To answer these questions, in-depth interviews were conducted with four participants of both the Basic Course and the Plus Course. They were questioned about what they had learned from the various sessions of the instructional development program: the Basic Course session, the 360° feedback session, the session on video vignettes, and the peer group discussion. The teachers mentioned various learning outcomes in the interviews, such as being more aware of their role as a teacher and acquiring new pedagogical knowledge and skills.

The participants' learning processes were analyzed using Clarke and Hollingsworth's model. In this model four domains are distinguished in which teachers' learning can take place. Diagrams are used to visualize different learning patterns as pathways through the domains. Some sessions resulted in patterns that included only one or two domains, while others included all four. The External domain, where the instructional development program is located, was the starting point for all diagrams constructed in this study. All teachers reported having learned from the program.

According to Clarke and Hollingsworth (2002) the complexity of the diagrams indicates the complexity of teachers' learning. A diagram with many

arrows (reflections and/or enactments) and including many domains suggests that a complex learning process took place. The diagrams resulting from the interviews indicated that the session that included feedback from students (Session 360° feedback) contributed most to teachers' learning. The most complex diagrams were found when there was a combination of theory and actual practice (practicing in the workplace), and when the student (located in the Domain of consequence) was actually included in the sessions.

6.3 MAIN CONCLUSIONS OF THE STUDIES

In this section we will address the main research question and summarize the general conclusions of our research.

> *What characteristics of effective instructional development are appealing to medical teachers and relevant for the design of instructional development programs for medical teachers, and what do these teachers learn from a specific program that takes into account those characteristics?*

Figure 6-1 depicts the sixteen characteristics that were identified in our first study as effective as well as appealing and relevant in medical education. They were selected on the basis of the literature (knowledge-for-practice), practical knowledge (knowledge-in-practice) of teacher educators, and by asking teachers about their preferences.

An existing instructional development program was analysed and expanded, an operation for which the sixteen effectiveness characteristics were used as a framework. Extending the program meant that more time was available for practicing in the workplace with what the participants had learned. Teachers reported having learned new skills about feedback, having become more aware of their role as a teacher, and being more focused on their students. By using Kirkpatrick's four levels and Clarke and Hollingsworth's model as frames of reference, it was possible to focus our analysis on what teachers had learned in the program and what learning processes could be distinguished in specific sessions of the program. Teachers reported what they had learned from the various sessions, and these data enabled us to construct diagrams that visualized the learning processes. In the diagrams, reflections and actions were included (see Section 6.3.4), as well as the different domains mentioned by the teachers. Including student feedback in the sessions resulted in diagrams indicating more complex and rich forms of learning, and this was also the case when theory and practice were integrated.

6.3.1 General conclusions

Our general conclusions can be divided into conclusions about characteristics of effective instructional development programs and about teachers' learning. Although the conclusions in this section are presented in units, they can be properly interpreted and understood only in combination with the information presented in the rest of this thesis.

Regarding characteristics of effective instructional development:

- Sixteen characteristics of instructional development were identified that can be used to design effective instructional development programs in the medical setting (Chapters 2 and 3).
- Combining this empirical research knowledge from effectiveness studies (*knowledge-for-practice*) with the practical knowledge of teacher educators and with teachers' own preferences, is a way to implement these characteristics into a specific training context in medical education (Chapters 2 and 3).

Regarding teachers' learning:

- The popularity of the Train the Trainer course among medical teachers can be ascribed to the fact that it has many characteristics of effective instructional development that are also important for medical teachers when they consider participating in instructional development (Chapter 4).
- In instructional development programs there appears to be a tension between what is best according to teacher educators (as indicated by the characteristics), and what is most desired by participants (as indicated by participation rates and satisfaction) (Chapter 4).
- Although teachers report less satisfaction with a course that was more consistent with teacher educators' preferences, the same teachers also report changes in their behavior, learning, and learning climate in interactions with their students.
- The Clarke and Hollingsworth model (2002) is a helpful frame of reference in which to represent learning sequences in the field of instructional development in medical education.
- The impact of instructional development programs on teachers' learning might be improved by including characteristics that relate teachers' learning to their Domain of practice and their Domain of consequence (Chapter 5).

6.4 DISCUSSION

With regard to the results of the studies in this research project several points of discussion can be raised. In this section we will discuss our findings regarding the characteristics of effective instructional development, the instructional development program, and teachers' learning.

6.4.1 Characteristics of effective instructional development

Medical instructional development in comparison with the general educational field

Although the field of medical education seems to have its own specific approach to facilitate medical faculty in their roles as teachers, our results suggest that instructional development in medical education is not essentially different from instructional development in other fields. In the studies described in Chapters 2 and 3 we found that from a list of characteristics, compiled from reviews of both medical and general instructional development programs, medical teachers as well as medical teacher educators selected characteristics from both fields as appealing (teachers) or relevant (teacher educators). However, it is not possible either to say that there was a clear preference for the characteristics derived from research studies in the medical field, as brought together in the review by Steinert et al. (2006), over characteristics from the general empirical field (Guskey 2003). This suggests that instructional development in the medical context could benefit from results from the general educational field regarding the design of instruction programs. Since both fields have their own focus and research traditions, increased interaction between these disciplines will be an enrichment for both. Our findings are in line with those of Steinert et al. (2006), who in their review of the medical education literature state that many of their findings are similar to what has been found in literature reviews on the training of university teachers in general. They advised researchers investigating instructional development in medical education to explore and learn from the relevant literature outside the medical field, incorporate those findings and methodologies into new research in the context of medical education, and collaborate with researchers in the field of higher education. On the other hand, findings in medical education could also be of interest for general educational research (cf., Weimer and Lenze, 1997).

Using the characteristics of effective instructional development as a framework

The method of combining results from the literature with teachers' preferences and the selection made by teacher educators, thus using their practical knowledge, provided us with additional information relevant to the medical context of our research. The sixteen characteristics selected were used as a framework to study an existing successful program and later to adapt that program. According to Guskey (2003) it would be unrealistic to assume that one single list of characteristics would emerge, because of the complexity of the context. He states that by agreeing on criteria for "effectiveness" and by providing clear descriptions of important contextual elements of instructional development programs, this type of research would improve in quality. The framework of our research is close to what he describes: it offers the elements which are important, and leaves space for contextual adaptations.

This framework proved useful as it showed that in a well-known course that is widely considered to be successful, most of the characteristics of effective instructional development were well-implemented. Interestingly, almost all characteristics that were considered important by teachers were among those that were well implemented in this successful course, whereas some of the characteristics labelled most relevant by teacher educators were less well implemented. Apparently, the design of the existing successful course was mainly in line with teachers' preferences.

Tension between "best" and "what is most desired" by teachers in instructional development

The characteristics of effective instructional development that were found to need more attention in the course were mainly those that the teacher educators had indicated as relevant for teachers' learning. We found that if those characteristics were taken into account in the design of a program (the Plus Course), the teachers reported having learned on the Learning, Behavior, and Results levels in the Kirkpatricks' model, but that their satisfaction with the course was lower than with the Basic Course. This suggests a difference between "what is best" in the design of instructional development programs according to the results from the literature, and "what is desired" by medical teachers as indicated by the preferences and satisfaction reported.

Given the fact that Steinert et al. (2006) found that almost all articles studied in the medical context reported a high satisfaction on the part of teachers, it is likely that in the design of instructional development programs in medical education the emphasis is more on "what is desired" than on "what is

best". Teacher educators and program developers in the medical context may consider the attractiveness of these programs relatively more important than their counterparts in the general educational field, because it is more difficult to convince medical teachers to participate. Unlike teachers in general education, many medical teachers have never participated in educational instructional development programs (in our study 53% of medical teachers had not yet participated in any such program). By designing a program that is appealing to teachers the developers possibly hope to attract those new teachers. However, if teacher educators only focus on teachers' preferences they will miss opportunities to design a program that is as effective as possible. Nevertheless, an instructional development program will have to strike a compromise between "attractive to teachers" and '"as effective as possible", because teacher educators want not only an effective program in terms of teachers' learning, but also a high participation rate.

6.4.2 Improving instructional development

Including students in instructional development

In our study (Chapter 5) we found that including students (i.e., including the Domain of consequence) in an instructional development program led to teachers' reporting complex teacher learning, including reflections and actions (enactment). It was especially the students' individual feedback to their teacher that resulted in complex learning diagrams. As students are the target group of teaching, it is not surprising that they can be a useful source of information for teachers. From the literature it is known that student ratings (as well as other input from students) can provide teachers with feedback, advice, and support by which to improve teaching (Prebble et al. 2004). These student evaluations might be useful, valid, and reliable (Menges & Austin 2001). Weimer and Lenze (1997) state in their review that consultation with a staff member over student ratings has been found to have a significant and positive impact on teachers' learning.

Interestingly, characteristics that refer to the inclusion of students in instructional development programs were not among those selected by the teachers and teacher educators in the sixteen effectiveness characteristics. This is in line with the review by Steinert et al. (2006), in which the importance of including students is not highlighted either. However, this finding goes against Guskey's review (2003), in which two characteristics that include students are formulated: one indicating that data on student learning should be used, and the other indicating that students' backgrounds and interests are important for teachers. Steinert et al. (2006) do not include the students because many publications in medical education literature do not focus on students. This may

be explained by the fact that medical specialists' first concern is patient care rather than student learning (Dolmans et al., 2004), which results in clinical teachers being more focused on their patients than on their students. In this thesis we found that student results are important for teachers' learning. It would, therefore, be advisable to include the students in future instructional development programs. This could take the form of practical assignments, in which teachers ask their students for feedback.

Linking theory and practice, including assignments

After the adapted instructional development course, teachers reported having learned new feedback skills, having become more aware of their role as a teacher, and having become more focused on their students. Teachers seemed to have shifted to a more student-centered approach after participating in the adapted course. The increased focus on the students is in line with Gibbs and Coffey (2004), who uses the Approaches to Teaching Inventory (Trigwell et al., 1999) in their study. They show that after participating in a 4-18 month training program, teachers became less teacher- and more student-centered. Postareff et al. (2007) have also, in their one-year pedagogical training, found a change towards more student-centeredness by the participants, although they show that this change occurred slowly, and their results imply that an intensive pedagogical training (in their course: 10-12 EC) is needed before positive changes in approaches to teaching emerge.

In Chapter 5 we focused on specific sessions and compared them to the learning outcomes. It was especially the sessions combining theory and practice, and those that included assignments involving students, that led to complex diagrams, indicating that teachers had reported having learned a great deal. We expect the sessions that enable teachers to link their new knowledge and skills to teaching practice (Domain of practice) and to the students (Domain of consequence) to be important for obtaining complex learning results. Characteristics that might stimulate this link could be related to characteristics that change teachers' Personal domains (e.g., promotion of reflection, the inclusion of pedagogical knowledge) and their Domains of practice (e.g., provision of feedback, inclusion of experiments). We also think that the scheduling of the program over a longer period was important, as teachers had more time to experiment with what had been learned in actual practice. Holton et al. (2000) identify three factors that are considered primary variables whose interaction affects the transfer of learning from the training environment to the work environment: (a) The ability of participants to use the skills learned in the work setting, (b) their motivation to use them, and (c) the support from the work environment in the use of these skills. The scheduling over a longer period and

the integration of theory and practice are in line with the first variable presented by Holton et al. (2000). The motivation to use what is learned (second variable) will be stimulated if teachers see the effects of their change of teaching behavior, e.g., more motivated students or better results. Holton et al.'s (2000) last variable, about the supportive working environment, is more difficult to influence by teacher educators, as this support is the responsibility of policy makers.

Evaluating teachers' learning

In order to improve instructional development programs it is important to evaluate the impact of those programs and to study the underlying learning processes. In general, Kirkpatrick's Reaction level is the only level that is measured in the evaluation of instructional development programs. This evaluation usually takes the form of a "happiness rating" measuring participants' satisfaction with the experience and their appraisal of the usefulness of the program in their work (Sparks, 1994), with no attempt to measure higher levels of effectiveness (Prebble et al., 2004), and without addressing the issue of achieving change (Guskey, 2000).

In this study we applied two other techniques to evaluate an instructional development program: we used not only Kirkpatrick's levels as a format for measuring impact on more than the Reaction level, but also the Clarke and Hollingsworth model, which made it possible to visualize the teachers' learning processes in the various sessions of the program, enabling us to see which sessions led to complex learning. These other methods of evaluating can provide teacher educators with additional information about teachers' learning and about underlying factors affecting this learning that cannot be derived from regular evaluation forms.

The two models seem to overlap, so it would be interesting to compare the Kirkpatrick model with the Clarke and Hollingsworth model in order to identify similarities and differences. The first model (Kirkpatrick's four levels) is displayed in Figure 6-2, and the second in Figure 6-3. The models have different goals: the first was constructed to identify the effects of a program, and the second to visualize a more "complete" picture of the learning processes, including the possible effects and the various factors that influence teachers' learning. This explains why in the second model (Clarke & Hollingsworth, 2002), teacher satisfaction (Reaction level) is not taken into account, because this does not provide information about the learning process. Instead of this level the model includes an External domain, related to information that may have initiated the learning processes, such as a specific instructional development program. The Clarke and Hollingsworth model is a non-linear model which can be used to visualize teachers' learning. Kirkpatrick's model contains four levels, which are

not linked to each other, although in various publications (e.g., Steinert et al., 2006) the levels are used hierarchically.

Although the two models differ, they also seem to have many similarities. They both display change (teachers' learning) and their four different levels/ domains seem to be comparable to each other. In Figures 6-2 and 6-3 the levels/domains that seem to be similar are indicated by the same shading. It becomes clear that the Learning level is comparable to the Personal domain, the Behavior level to the Domain of Practice, and the Results level to the Domain of consequence.

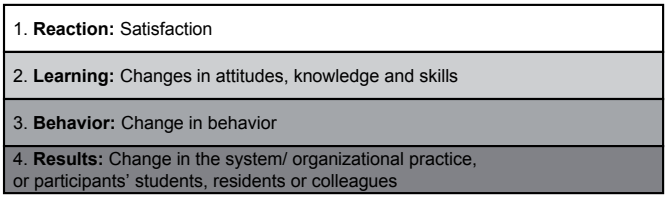

Figure 6-2. Overview of Kirkpatrick's levels, adapted by Steinert et al. (2006)

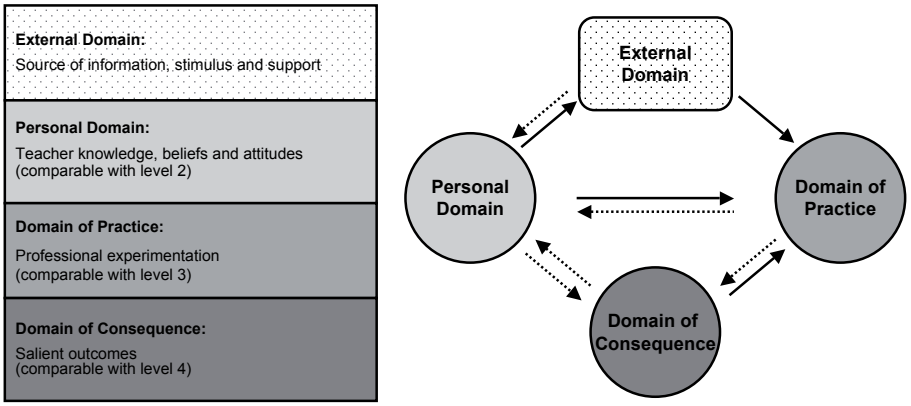

Figure 6-3. Model of teacher's professional growth (Clarke & Hollingsworth, 2002)

Clarke and Hollingsworth's model is the most attractive for understanding underlying processes, because besides learning effects it also takes learning processes into account. It is generally known that the Results level (e.g., students) is difficult to attain in instructional development programs (Stes, Min-Leliveld, et al., 2010). The Clarke and Hollingsworth model might explain this: a change in the Domain of consequence (e.g., students), as facilitated by an instructional

development program (External domain) needs to be preceded by changes in the Personal domain and/or the Domain of practice.

Teacher satisfaction (Reaction level) does not reflect teachers' growth, and hence is not included in the Clarke and Hollingsworth model. Weimer and Lenze (1997) already concluded that participants' reactions do not contribute to a clear picture of the real impact of instructional development, and that it is questionable whether this level can be used as a measure of impact. Thus, in reviews on the evaluation of effects of instructional development programs this level is not always included (e.g., Stes, Min-Leliveld, et al., 2010). Unlike Stes, Min-Leliveld at al. (2010) we did include the Reaction level (satisfaction) in our study, as we think that the information on this level might alert us to possible reasons why teachers would or would not report teacher learning. This level might be considered a pre-condition for attaining the other three levels of Kirkpatrick (1994).

6.5 STRENGTHS AND LIMITATIONS OF OUR RESEARCH

For this research project several strengths and limitations can be identified.

6.5.1 Strengths

A first strong point of the research project is that we have combined results from the literature (knowledge-for-practice) with the knowledge and preferences of medical teachers and teacher educators (knowledge-in-practice) in order to adapt the available information to the context of medical education. This resulted in a selection of characteristics that appealed to teachers, were considered relevant by teacher educators, and were described as effective in the literature. Integrating this knowledge-for-practice with practical knowledge and knowledge of teachers' preferences is something that is not yet common in the instructional development literature, but that we recommend because it might help to translate theory into actual practice (and so bridge a gap).

A second strength is that we have combined the results found in general education literature (e.g., research on teacher education and research on higher education) with those from medical education literature. In general, those two research fields are worlds apart: they have their own conferences, their own journals, their own research traditions, and their own terminology. Combining results from both fields makes it possible to find overlapping themes and to combine bodies of knowledge. The literature on teaching in higher education and teacher education in general has a longer research tradition in the field of instructional development, so that this literature may contribute to the development of theoretical knowledge in medical education. In the medical educational field the topic of instructional development is very popular at the

moment, and much research is being conducted internationally. An advantage of the research in this field is that, contrary to the situation in the general educational field, the curricula of the medical schools are quite similar in many countries, facilitating the comparison of results.

A third strength of our research is that we have used alternative ways to evaluate an instructional development program. Using the Kirkpatrick model as a framework has produced a variety of information on the effects of the program on teachers' learning, as, besides satisfaction, it also focuses on actual learning, behavioral changes, and changes in the student. This model might be integrated into the regular evaluation forms. The Clarke and Hollingsworth model can be used to provide in-depth insight into the underlying learning processes of teachers. The sessions that contributed to complex learning patterns should be considered the components most powerful for learning. Teacher educators could use this information in the development of adapted programs.

6.5.2 Limitations

There are several issues that limit the scope and, consequently, the conclusions from our study. A first constraint of the study was that we combined two reviews (Guskey, 2003; Steinert et al., 2006). This raises the question whether the reviews selected (from two different fields) are the most relevant and complete overviews of characteristics of effective instructional development. Also, combining two reviews results in overlap between the various characteristics. We combined some overlapping items, and used teachers' preferences and teacher educators' practical knowledge as a filter to finally select sixteen characteristics from this list, in order to make sure that the final selection was relevant for this context. This still does not guarantee that the selection is complete. New reviews may introduce new characteristics that have not been taken into account in our selection. We should also be careful not to use specific effectiveness characteristics in isolation, as this can lead to a fragmented and mechanistic view which does not do justice to the complexity (Doyle, 1990). We have to take into consideration that whether or not the characteristics contribute to effective learning also depends on the individual context.

A second limitation was the fact that the study presented in Chapters 4 and 5 only included participants in one specific instructional development program who were all affiliated with the Leiden University Medical Center. This means that, strictly speaking, conclusions can be drawn only about this instructional development program, in this center, and in the Netherlands. The chosen instructional development program (Train the Trainers), however, is used in almost all Dutch medical schools in a comparable format, so similar results are to be expected in other schools. Similar courses also exist abroad.

Although the differences between the medical centers in the Netherlands and other international medical schools are much smaller than the similarities, we still should be cautious about generalizing our findings.

A third limitation of our study was that it depended on medical teachers' self-reports. We did include feedback from students, but did not include results related to the actual behavior of teachers. The choice to rely on teachers' self-reports was made in order to limit the time investment asked from the busy medical teachers. Our program was set up in such a way that it took more of the teaching staff's time than usual, and researching their behavior would have stressed their time schedule even more. Moreover, collecting data on the actual behavior of the specialists in learning-by-doing situations would unavoidably have involved patient contacts, which could have led to precarious situations. Finally, we may mention that we have the impression that medical teachers as a group are quite direct in their answers and comments, so that there was a relatively low risk of our receiving only socially desirable answers.

6.6 IMPLICATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

In this section we will discuss ideas for future research, and practical implications for policy makers, teacher educators, and teachers.

6.6.1 Future research

In our explorative study we adapted a popular instructional development program (Chapter 4). We were able to show that this adapted program facilitated teacher learning. In future research we recommend setting up design-based research, using the sixteen characteristics identified in this study in other medical schools in order to further adapt and improve instructional development programs. A suggestion could be to use the Basic Course as a starting point, but other programs could also be used. We recommend including various centers in the data collection, so that a "second phase study" (Borko, 2004) may be carried out. Borko (2004) distinguished various phases in research, related to a time sequence. In phase one, research on a new type of program is started on a single program on one site (as in this thesis). The research proposed above would be a second-phase study in which a single program is analyzed on multiple sites with multiple facilitators. The third phase would include comparative field studies, including multiple programs on multiple sites.

Further research should also include a control group, or the "internal referencing strategy" (Haccoun & Hamtiaux, 1994). In this strategy one single group is used, but specific test items are identified that would not be expected to change from pre-test to post-test. Comparing the changes in these training-

irrelevant items with the changes in items that are considered relevant would make it possible to study the effectiveness of the training.

Further research might also include longitudinal studies, in order to determine if changes as found in our research would last over a longer period. Most research focuses on the direct rather than the long-term effects of instructional development programs. Rubak et al. (2008) studied the effects of the Train the Trainers course after six months, and found positive outcomes. In future research it would be important not only to rely on teachers' self-reports, but also to include observations of actual behavior and results from the students, for example by using the Cleveland clinic's clinical teaching effectiveness instrument (Copeland & Hewson, 2000), the Pheem questionnaire (Roff, McAleer, & Skinner, 2005), or the Student course experience questionnaire (Ginns, Prosser, & Barrie, 2007). These questionnaires would require a larger group of respondents, so for future research a larger-scale study is recommended.

We did not focus on the professional development of staff members regarding, for example, their leadership roles. Steinert (2000) indicates that, due to changes in medical education, focusing exclusively on the development of medical staff in their role as teachers will no longer be sufficient. Future research might focus on more than instructional development only, for example by including the professional development of staff members regarding leadership capacities.

In the study described in Chapters 5 and 6 we found that participants were less satisfied with the adapted program, but reported having learned much. A final direction for future research might be to investigate the relation between satisfaction (Kirkpatrick's Reaction level) and actual learning. Kessels (2010) distinguishes two aims in an induction program for teachers: one being teachers' well-being (satisfaction), and the other the actual learning of teachers. It would be interesting to investigate if those two are really separate, as Kessels (2010) found, or that satisfaction is a prerequisite for participation in those programs, as we supposed in the Introduction.

6.6.2 Practical implications

In this section the possible practical implications of our findings for higher education are discussed, and recommendations are formulated for teachers as well as for policy makers, program developers, and teacher educators.

Implications for teachers

In this research project we selected sixteen characteristics (Figure 6-1) that should be part of effective instructional development programs. As part of this selection procedure we looked at teachers' preferences, and found that an

existing successful course fitted these preferences well but that improvements in the program, based on the effectiveness characteristics, could lead to more effect on teachers' learning. An example of such an improvement is to include characteristics mentioned in Figure 6-1 as needing more attention in the implementation, such as the introduction of practicing in teaching practice, personal support, and the extension of the program over time. When selecting an instructional development program, teachers are advised to look for programs that include these characteristics. In this way they can be more sure that their valuable time is well spent, as they will participate in program that is both well-designed and effective.

In Chapter 5 we indicated that it is important to include information from the students or regarding their performance in the instructional development program. For example, feedback from the students led to more awareness about the participants' roles as teachers and also stimulated teachers' reflection. Teachers are advised to ask students for specific feedback and discuss the feedback with the students in order to improve teaching quality.

In our research we identified relevant knowledge-for-practice. We asked teachers about their preferences when participating in instructional development programs, and teacher educators about their practical knowledge about the design of these programs. Involving stakeholders in instructional development in the design of a program is a good way to ensure it specifically fits the local context. Teachers are, therefore, encouraged to join task forces that advise their institution on instructional development.

Implications for teacher educators, program developers, and policy makers

The sixteen characteristics (Figure 6-1) of effective instructional development can be used as a framework for the analysis of existing instructional development programs, as well as for the design of new ones. In Chapter 4 we showed that teachers participating in a program that was adapted according to this framework reported having learned much, even though they claimed to be less satisfied with the additional course than with the Basic Course. Therefore, it is important to monitor participation rate and satisfaction as well as actual learning in the newly constructed programs, .

In their daily working environment medical teachers are normally more focused on their patients than on their students, even if they are in their role as a supervisor for their students. Programs for medical teachers should therefore be aimed at shifting the teachers' focus in the workplace from the patient to the student during supervision. Including the students in the instructional development programs, for example in video vignettes or by collecting feedback from them, could accelerate this learning process (see also Chapter 5).

In Chapters 4 and 5 we described two different methods that we used to study teachers' learning. These two methods provide teacher educators and policy makers with complementary information on teachers' learning. The type of information gathered by studying the Kirkpatrick levels and the Clarke and Hollingsworth model can be used in combination, in order to further improve instructional development programs. By strengthening components that are reported to be useful, and by changing those aspects that did not lead to learning, important improvements can be obtained. Our questionnaire, based on Kirkpatrick's four levels, could be a starting point for the further development of instruments for the evaluation of instructional development programs on more than just the Reaction level.