CHAPTER III

Development

This section presents key findings about the process by which the 30 IMD products were developed. The original study design included parallel data collection and analysis related to the non-NSF-funded materials, but the publishers of those materials refused to cooperate with the study. Nonetheless, the findings about efficiency and effectiveness of development of IMD materials provide important information to NSF.

The section begins with an examination of the processes used to develop the materials, including the nature of development teams, approaches to pilot and field testing, and challenges faced by developers.

The Development Process

Almost without exception, the development and field-testing processes were iterative and intertwined. Early versions would be drafted, piloted (most often with teachers from the local area), and revised based upon their feedback. Then, more refined versions of the materials were field tested in multiple national sites and revised still further before being turned over to publishers.

Here is a fairly typical development and field-testing process as exemplified by one project:

1. Identify the important content to be covered within the larger area (e.g., within the geometry segment).
2. Brainstorm about types of stories or powerful concepts that could be featured.
3. Meet with the writers and review the content with them. Generate a first draft.
4. Send the treatment out multiple times ("Drafts went back and forth ad nauseam. We had at least 4 garbage cans full of old drafts!") to the advisory board for feedback regarding content, dialogue, cultural sensitivity, and gender equity.
5. Conduct focus groups of teachers, scholars, and community groups to review the treatments.
6. Pilot an early version of the module in local classrooms.
7. Revise module if necessary, based on pilots.
8. Field-test final version of module in national sites, revising as necessary.

The Development Team

The overall approach to product development relied heavily on the composition of the development teams. All projects involved a wide variety of individuals representing many different professional positions and areas of expertise in the development process. Core teams usually consisted of three to five members, but could include anywhere from 20 to 400+ expert advisors and editors. These teams worked collaboratively to develop and test the materials. An impressive array of content and pedagogical expertise was typically represented within each product's development team. Most frequently, these areas included:

- Science and/or mathematics
- Science and/or mathematics education
- Developmental and/or cognitive psychology
- Curriculum development

- Classroom teaching experience
- Technology and multimedia production

The involvement in development of a broad range of perspectives afforded clear benefits with respect to the accuracy and thoroughness of the content covered by the materials, as evidenced in the expert review of the products. However, only half of the projects included teachers, either using release time, summer work, or a semester- or year-long leave, as members of the development team.

A few developers believed that broad-based involvement of stakeholders helped to diffuse potential controversy and gather support for the products.

> "Development is a very tough process, because you have to please multiple audiences that all have different interests—teachers, state people, political groups, etc. A valuable lesson for us was to have a national advisory board. You can defend the content because of having had their approval." *(Developer Interview, Project 1).*

> "It [development] was a very broad-based process...Each revision involved input from an advisory board that included teachers, people from industry, laboratories, and professors...One thing we learned is how political this process is, especially if you are trying to do something different. It is not enough to do well and throw it out there. You have to work with and build a constituency. You must be vigilant that things can go wrong; you need to stroke, explain, and have patience." *(Developer Interview, Project 2).*

A little over half of the development teams included classroom teachers, and both developers and publishers of the products emanating from such teams repeatedly highlighted their importance in ensuring that materials would fit real-world teaching contexts. They agreed that extensive involvement of classroom teachers in the development process was beneficial, and the consensus was that a lot of involvement is good; more is even better.

> "All of us had lived in a university research environment for a long time...I can't overstate how important their [teachers'] contribution was. One teacher we hired half-time and another 3/5 time. They'd teach and be with us on a daily basis. It was like having a school walk through the door." *(Developer Interview, Project 7).*

> "It is important that teachers be involved with the development of new materials. If materials do not make sense to teachers and are not able to be customized by them, teachers will not use them." *(Developer Interview, Project 5).*

When development teams did not include teachers, it was more likely that the products would fail to meet teachers' criteria related to "usability." For example, two products included "humor" that implementing teachers reported as escaping students. Others required reading levels far higher than students were able to achieve:

> The most vexing problem to the teachers is the reading level of the text. The students, in general, do not seem able to read with comprehension. Teachers say

> they are spending more of their time teaching basic reading and comprehension than anything else. The second problem students have had is that they are so frequently absent that they fall behind and cannot catch up. Falling behind has greater consequences than usual because of how the text structures group work. Teachers wonder, "how are the absent students supposed to fit back in when they come back?" *(Summarized from Focus Group, Project 12).*

In addition, the products that were developed without teacher input at the earliest stages were more likely to demand more understanding of science or mathematical content than most teachers possessed.

The Research Base

Without exception, materials were well informed by research on cognitive development, child development, and learning processes. In some instances, materials developed under IMD support were an extension of previously developed curricula or the developer's own research on curriculum, instruction, and the learning process. For others, the IMD program provided the impetus or initial opportunity to produce innovative instructional materials that embody principles of learning that have emerged from work in fields such as developmental psychology and cognitive science.

As evidenced by the high marks given the products by the expert panel reviewers, the purposes, design, and intended uses of the materials developed under the program are in keeping with the larger standards-based reform agenda. Without exception, developers were quick to offer examples of and clearly demonstrate how the content and pedagogy of their products are a direct reflection of the principles underlying reform, such as inquiry-based learning, constructivism, anchored instruction, and authentic assessment.

> [The product] is standards-based. The developers infiltrated the standards-writing process. Its units are cited as an exemplary way of meeting more than half of the standards–54 out of 72. One state listed it as one of the exemplary units for its standards. *(Summarized from Developer Interview, Project 11).*

> "To maximize student understanding of the content, math should arise from realistic situations so that the students have a need and a reason for doing the operation. That was the primary theoretical focus, but we also built in problems that encouraged student interaction, which is in line with social constructivism. For example, in introducing students to similar triangles and the notion of slope, the math problems might be presented in the context of hang gliding or building a bridge. Activities are organized from less formal to more formal, and they begin with what students already know about the particular problem being posed." *(Developer Interview, Project 15).*

Students Served

In keeping with the tenets of reform, products were designed with "all" students in mind within the particular grade level(s) targeted.

> It is intended to be used by all high school students in grades 9-11, and has also

been used with high-ability 8th graders. Materials were designed to increase the performance of students with a variety of aptitudes and talents. The curriculum allows students to look at material in many different ways. *(Summarized from Developer Interview, Project 16).*

"Our first concern was that the materials wouldn't reproduce the bias toward serving suburban communities...So for example, a unit might ask kids to explore cracks in the sidewalk to see what grows there, as opposed to talking about pond life." *(Developer Interview, Project 10).*

One project's attempt to level the playing field among science-oriented students and other students received mixed reviews from consumers:

"Instead of the traditional sequence, we focused on societal issues that involve [science]...The intent was to design a high school...course for students who are going to college, but are not scientifically oriented. The reaction of high school...teachers varied. We have teachers who love us and some who don't love us. Some believe that [science] must be taught in a traditional way. They might say that we watered it down." *(Developer Interview, Project 2).*

Another potential pitfall of attempting to target all students surfaced in connection with a mathematics project: instruction targeting *all* students that does not specifically address different learners' diverse learning needs can inadvertently create the problem of "teaching to the imaginary middle." So rather than one fitting all, it can look more like one size fitting only a few. One product was described as being, "developed for all students," and yet at the same time the developer described the materials as targeting "the middle 60 percent of students in terms of achievement."

Comprehensive and Supplementary Materials

Opinions regarding the relative advantage of comprehensive versus supplementary materials as a stimulus to reform arose with fairly equal incidence. Those favoring the comprehensive approach believed quite strongly that it represented the best way to achieve meaningful reform:

The developers believed very strongly that the material should not be available as modules because it would permit schools to adopt only parts of it. They felt this would not bring about the complete curriculum change that they wanted. *(Summarized from Developer Interview, Project 16).*

"We never considered replacement modules as an option. Always, the idea was: If you're going to do it, do it. It was a philosophical decision. If you're going to change how teachers teach and how children learn math, you have to change everything." *(Developer Interview, Project 15).*

The team moved from supplementary to comprehensive materials in part for reasons having to do with equity. They felt that teachers in urban schools could not easily buy supplemental materials, so a comprehensive curriculum would be more widely available. *(Summarized from Developer Interview, Project 18).*

Less strongly held views favored a supplementary approach, although the reasons for that position differed among respondents. Some viewed supplementary materials as being more responsive to the needs and realities of the school market, while others viewed the development of supplements as the best or at least most feasible first step toward designing comprehensive materials.

> "We made the decision to produce modules rather than a comprehensive curriculum...Mathematics teachers are fiercely independent and like to do things in their own way...By producing separate modules, the teachers would be free to choose whatever materials they felt would be helpful and could better integrate these modules into existing curricula." *(Developer Interview, Project 8).*

> The market partly directed this decision. People want to be able to pick and choose. Teachers need flexible, supplemental materials, and the developers wanted to provide something that teachers could add to their existing plans. *(Summarized from Developer Interview, Project 12).*

Role of Assessment

There were vast differences across products in terms of the extent to which they incorporated student assessment within their design. Somewhat surprisingly, roughly 20 percent of the projects included no student assessment at all, at least not initially.

> Assessments are not included and were "not a heavy emphasis" in the eyes of the developers. It was assumed that teachers using the modules would "create their own tests." *(Summarized from Developer Interview, Project 1).*

> Originally, they had no vision of student assessment. After field testing, however, teachers expressed a strong desire for assessment items. Developers then responded by creating the Teacher's Resources Package. *(Summarized from Developer Interview, Project 19).*

> In the words of a developer, "We are not in the assessment business." The products carry the suggestion that students develop portfolios based on the materials, but there was no intention to build in assessment as part of the product. *(Summarized from Developer Interview, Project 20).*

When assessments were included in the materials, nearly all tended to emphasize curriculum-specific, performance-based assessment as opposed to more traditional measures, although here were exceptions to this.

> Most traditional assessment, according to the Co-PI, tries to find out what students don't know. Their vision of student assessment, according to all three interviewed, is to allow students as many ways as possible to demonstrate what they have learned in class, which in turn gives teachers many more ways of understanding what their students know and understand. They try to make teachers see that assessment is something that you do every day, and is distinct from grading. The teacher and student materials offer multiple modes of assessment:

> oral presentations, self-assessment, problems-of-the-week, and take-home assessments that extend the ideas embedded within a unit to a different context. *(Summarized from Developer Interview, Project 21).*

> "We offered suggestions for teachers on how to make observations, suggestions for helping kids put together portfolios and cumulative folders, and guidelines for helping kids decide what to showcase that would show improvement over the course of the module. Kids don't consider them tests, but these activities give the teacher an idea of where kids are." (*Developer Interview, Project 4).*

Pilot and Field Testing

Development included pilot and field testing. Pilot tests were generally of early versions of the materials and involved teachers who were geographically close to the developers, and field tests included a wider variety of sites and teachers. Pilot-test teachers gave feedback to developers, who also frequently observed in classrooms. The teachers provided information about how easily the materials were used, student reaction to them, and problems encountered. The value of the feedback is demonstrated by the changes developers made following pilot testing. However, we note that pilot-test teachers were drawn from those with whom developers had an earlier relationship, either through professional associations or prior work with the school districts that employed them. Consequently, they were likely to be sophisticated, reform-oriented teachers who had participated in earlier professional development activities.

Products were field tested in a wide enough range of settings representing students with a variety of backgrounds. Local and national field-test sites included students from all ethnic backgrounds; ability levels; geographic areas; and from an appropriate mix of urban, rural, and suburban settings. There were two exceptions to this. One developer felt that too few urban schools were included in the field tests, which reportedly, "led to a myth that the materials were only good for suburban and higher level students." The other noted that the project would have benefited from, "better representation in the field-test groups of high-ability and accelerated students."

The realities of school schedules, governance, and agendas often resulted in delays and other challenges to the field-testing process.

> Field testing suffered from many problems related to school independence. Schools were doing so on a voluntary basis, and implemented the materials to meet their own needs and commitments. Implementation varied across schools. Five of the ten schools put their best students into the control group and below average students into the project group. One school put all their students into the project group. Sites used different standardized tests and different testing schedules. Also, many schools did not have adequate technology to implement all the activities. *(Summarized from Developer Interview, Project 24).*

Delays experienced in the field-testing process contributed to the widely held desire on the part of developers for funding to be granted for longer periods of time.

> The publisher recommends that developers be funded for a 6-7 year cycle, rather than the more typical 3-4. This would cover two years of research and development to get the materials field-test ready, three years for field testing, and then a

> year or two to get the materials market-ready. He thinks that NSF needs to be better aware of the fact that curricular materials can't really be developed any faster than students and teachers can try them out and developers can redesign them. *(Summarized from Developer Interview, Project 6).*

The vast majority of data collected during field testing consisted of teachers' and students' perceptions of the usefulness of the materials, anecdotal evidence of students' engagement with or apparent benefits received as a result of the use of the materials, and measures of teacher or student satisfaction and attitudes regarding science and mathematics.

Even among the more recently funded and most commercially successful projects, there was a conspicuous lack of emphasis on evaluating the instructional materials' effect on student learning, either immediately following the experience or to assess student retention of concepts, information, and skills. Indeed, some of the quotations below indicate that developers believed that the sole focus of the required formative evaluation was on how readily teachers could use the materials. While perhaps disconcerting, it is not surprising that approximately 70 percent of the projects conducted no evaluation of student outcomes (comparative or otherwise) that would enable them to measure the impact of the materials on student learning of specific content.

> "It [outcome evaluation] was not required by NSF. Measuring student gains wasn't required; it wasn't necessary. We collected qualitative information only—observations and comments, input from developers, teachers, trial coordinators, and observers...No other evaluation was performed. NSF doesn't pay for assessment and data analysis...so the real research on the materials will have to come independent of us...." *(Developer Interview, Project 4).*

> The evaluation did not include attention to student outcomes. One team member did an evaluation of student attitudes early, after the first commercial edition (1988-89). Other results were anecdotal (e.g., impressions of the number of students continued in another course and were successful.). There was no concerted follow up to track the relative success of students. *(Summarized from Documents, Project 2).*

> There is constant evaluation and feedback from teachers through the developers' 1-800 phone number. [Name] has been the external evaluator for several years. While she has done a number of evaluations, they all focus on teachers rather than the actual material. They looked at how well the product worked as teaching material, with a secondary focus on students. *(Summarized from Documents, Project 23).*

> An evaluator was named in the proposal and carried out a very small scale evaluation, mostly because it was obviously required in the RFP. There was very little money to do evaluation—just a few days for evaluation and expenses—his role was more like an auditor's. He would come in and be present at key points and would provide feedback to the developers on what he was hearing and observing. He focused on the development process, not on the content. One bit of advice the evaluator offers to NSF: 10 days over three years is not adequate. *(Summarized from Evaluator Interview, Project 11).*

There were exceptions—some projects used part of their IMD funding or were able to secure additional funding to conduct or contract for evaluations of the impact of the materials on student learning. Several project evaluations indicated that the materials had a positive effect.

> The developers believe that there is no purpose in developing materials that are unproven. If schools are going to adopt a program, there must be evidence that the program improves student achievement. The program received a third-party evaluation that indicated that students in the treatment group performed significantly higher than students in the control group, and that treatment students scored higher than the control students regardless of pretest difference in their achievement levels. *(Summarized from Developer Interview and Project Documents, Project 17).*

> The staff completed both a formative and summative evaluation. Students who used the materials had higher raw scores than students in the control group. The treatment students answered more test questions and answered more of them correctly than students in the control group. On standardized tests such as the California Achievement Test, students in the treatment group made gains of one and a half to one full grade-level over scores from the pretest. Attitude shifts about science were also monitored for girls, Latinos and African American students. Positive shifts were found for female students from all ethnic backgrounds. The subgroup samples of Latino and African American students were too small to generate any significant data. *(Summarized from Documents, Project 19).*

> "Students using the materials outperformed control group students on measures of complex problem-solving and did better on some of the subtests. On computation and application of skills, the students did as well or better than the controls." *(Summarized from Developer Interview, Project 1).*

> Field testing involved 12 schools with 6 treatment and 6 control schools. Impact evaluation was done during the field testing phase. Among other impacts, the evaluation assessed whether there were differences in student outcomes related to gender, SES, or cultural background. The group's mean achievement score was significantly higher than that of the comparison group, and the difference was significant at the .04 level. Student attitudes showed no significant differences between treatment and comparison groups at both the pretest and post-test. *(Summarized from Project Documents, Project 5).*

Evaluation of student learning did not uniformly favor IMD products. Some materials evidenced little or no differences between IMD and other materials.

> The developers' philosophy is that all activities that have measurable outcomes need to be assessed. Both formative and summative evaluations were done during both the pilot and field-test phases. A control group design was used to measure basic skills, higher-order learning skills, and science attitude between students using the materials and those in regular classes. Standardized achievement tests were used along with problem-solving measures. The major result of the evaluation was that there was no significant difference on content knowledge gained

between students using the product and those in regular classes. *(Summarized from Interviews and Documents, Project 24).*

Development and Marketing

For virtually all products, the field-test sites turned out to serve as initial target markets. The involvement and participation of teachers as product developers led quite naturally to their buy-in and subsequent interest in adopting the materials.

> "...what they did was to bring together teachers to participate in training workshops during the field-testing even before the products were published. This created a cadre of people who keep coming back. It's been a fabulous way to get a stronghold in the market...You've got ready customers out there because of the field testing, and those contacts require care and feeding. You've got to manage the demand that the field testing creates, as opposed to just waiting to recontact them two or three years later." *(Summary from Publisher Interview, Project 27).*

> Purchasers have primarily included teachers, schools, and districts that have been directly involved in field testing. *(Summarized from Publisher Interview, Project 15).*

> "The field test touched on Texas, California, New York and Colorado, and a lot of the initial sales moved out from there. There were teachers there ready and willing to adopt the course and materials." (*Publisher Interview, Project 20).*

Role of Developer in Marketing

For many projects, developers played a critical role either directly or indirectly in assisting their publishers to identify target markets and the strategies for reaching them. Both developers and publishers agreed that IMD developers—given their still forming market niche—played a greater role in developing marketing strategies than is typically the case between a publisher and developer in the traditional textbook market.

> "[After some preliminary difficulties], NSF then gave us 90 days to find a publisher. All of the other programs that had made their initial agreements with standard textbook publishers had found that those publishers stayed with them until the materials were completed, but then they would eventually walk away without signing contracts. We had been seeing what was happening, so we had deliberately avoided them...Eventually, we approached one company with the idea that [we] would put them back in the marketplace. Up till then, they'd only been selling movie films, and had been losing money. I showed them how to go into video lessons, laser discs, computers, plus hands-on science materials. They got involved in 1987, and the [grades] 3-6 materials turned them around to where they started making a profit, and were later willing to invest more in the [K-2] materials." *(Developer Interview, Project 4).*

> "In marketing the materials, it also adds credibility to be associated with prestigious groups. The recognized name and credentials, combined with NSF funding,

> is stronger than NSF funding by itself." *(Publisher Interview, Project 4).*

> The publisher did not actively seek out or otherwise proactively identify members of the target market. Rather, they have relied largely on the training and professional development activities undertaken by the developers themselves to generate interest in such alternative curriculum materials. The publisher also relies quite heavily on word of mouth referrals between users, and on the fact that NSF's IMD program is known within the math community and people have been awaiting the products funded within it. *(Summarized from Publisher Interview, Project 7).*

In a few instances and for different reasons, developers doubled as the publishers of their own materials. For a small number of projects (less than 10 percent), the decision to self publish was a conscious one designed to circumvent the commercial publishing arena in order to make the materials available to teachers free or at minimal cost. For another small set of projects (again less than 10 percent), the decision to self publish was driven by difficulties encountered in finding or keeping a publisher to carry the materials.

In roughly 20 percent of projects, the developers played a far less active role in marketing, choosing instead to leave the marketing to the publishers once development was completed. This approach has met with mixed results in those instances where publishers were less experienced or otherwise hampered in their capacity to carry out their agreed-upon functions.

> Marketing and dissemination were problematic. These functions ended up with a group that was just starting in the elementary school market and had neither the knowledge nor the enthusiasm to market effectively. In the words of the marketer, "We did not have a lot of experience in the elementary school market. The sales force wasn't comfortable with the materials and they didn't put much energy into it. They didn't follow suggestions." *(Summarized from Publisher Interview, Project 20).*

Conclusion

Although the vast majority of the development teams receiving IMD funding produced high-quality materials that meet content standards, the development process itself contains the seeds of ultimate success or problems. For example, the composition of the development team influenced the degree to which materials were judged usable by adopters. More positively, field-test sites became early adopters and champions of the materials and the fit with content standards was frequently cited by adopters as influencing their decision.

Few projects collected data related to student outcomes, so there is little evidence besides expert judgment that the materials will help students meet high standards. On the other hand, few adopters asked for information about student achievement to influence their decision to adopt, although parent and teachers concerns during implementation frequently raised questions about how well students would perform.