
Memorandum

To: Steven Howell,
B-374 Rayburn House Office Building
Washington, DC 20515

From: Dr. Mark St. John

Date: 8/20/1998

Re: Response to Follow-up Questions from Congress

This memo contains my written responses to the follow-up questions that were derived from the recent hearing conducted by the Subcommittee on Basic Research on the topic of the National Science Foundation (NSF) Systemic Initiatives.

I have kept my responses brief and relatively informal. If you would like more detail or clarification, I would be happy to respond more fully.

1) Your testimony suggests that Systemic Initiatives should be evaluated on the extent to which the Systemic Initiative enabled a school district or state to initiate and sustain a process of continual improvement in the quality of math and science education. How you respond to those critics who would argue that aspects of the Systemic Initiatives – the standards endorsed by NSF, for example – actually hinder good science and math instruction?

My response to this question comes in two parts. In the first part I explore the reason that there may be disagreement with the Standards as a guiding vision, or some other aspect of the Systemic Initiatives. The second part of my response questions the degree to which National Standards are, in fact, the foundational piece of the Systemic Initiatives.

Both the Standards and the Systemic Initiatives are aimed at catalyzing improvements in K-12 mathematics, science and technology education nationwide. If there existed widespread satisfaction with the current status of math and science teaching, and if there was national consensus that the schools were doing well in this area, there would be no need for either Standards or Systemic Initiatives. But the opposite seems true – there is widespread dissatisfaction with the current state of affairs.

But that shared sense of dissatisfaction is where the agreement and national consensus ends. There is not a shared vision for what constitutes “improvement” in the teaching of mathematics, science and technology. And, in my experience, the sources of the disagreement lie not within the domain of research or technical analysis but rather they arise from deep underlying differences in values and beliefs. The arguments about Standards and the Systemic Initiatives represent very different points of view about the nature of good instruction, and even the purposes of schooling.

I was witness to the production of the National Science Standards produced by the National Research Council (NRC). (I am also very familiar with the Mathematics Standards produced by the National Council of the Teachers of Mathematics (NCTM)). In both cases, but particularly in the case of the Science Standards, I think that the documents that emerged are the products of very hard work and fair efforts at balancing very divergent points of view about the nature of “good science instruction.” The Science Standards, for example, lay out a vision of science education that emphasizes the mastery of important content, while at the same time, insisting that the process of science (inquiry) be an integral part of “good instruction.”

Thus, my first response to this question would be to challenge others to do a better job of putting together a vision of mathematics and science education that represents the multiple viewpoints of scientists, educators and the public. I think the current Standards do a relatively sound job of laying out a national vision of “good instruction.”

Because of the differences in perspectives, values and beliefs, differences in the vision of good math and science education are inevitably going to exist. One might argue that it is impossible to proceed until there is agreement about where we are trying to go. I think this point of view will lead to paralysis. I do not think it is wise for the federal government to wait for unanimous agreement of the ultimate goals of reform, or even of science education, before it invests its resources into the improvement of instruction around the country. There is substantial agreement now on what constitutes improvement even if the exact specifics of the goal state remain in dispute.

The second part of my response is to look at the reality of the work being done by the Systemic Initiatives. In contrast to the political debate and the strong ideological positions taken by different groups, the work of the Systemic Initiatives tends to be tempered by practical realities and local context. Thus, I question the degree to which the Systemic Initiatives are, in practice, founded entirely or even primarily on the vision of education laid out in the National Standards. My own experience is that, in practice, the Systemic Initiatives tend to be very pragmatic efforts that refer to state and local standards and to local notions of good instruction: Most Systemic Initiatives use the National Standards as a distant reference point.

There is no doubt that if one examines the work and activities of the NSF Systemic Initiatives, one will see some weaknesses. Observers might well see “inquiry-based” professional development activities that do not always present the content of the discipline in powerful and accurate ways; on the other hand, one might also see very traditional lectures on content that are not effective in either motivating or educating teachers. It is, in fact, relatively easy to gather examples of reform activities that are weak in one dimension or another. But I would argue that the source of such weakness is NOT inherent in the vision or approach of the Standards or Systemic Initiatives, but rather lies in the absence of local capacity to do the work that is required. By local capacity I mean, for example, the expertise of professional development staff; the ability of districts to adopt, implement and support high quality curriculum; or the knowledgeable and proactive support of district administrators for the improvement of math and science teaching. It is somewhat ironic that the very reason one needs the Systemic Initiatives – the absence of local capacity – is the very source of many of the weaknesses of their work.

Let me take this opportunity to raise one other point about a potential weakness of the Systemic Initiatives. Both the National Standards and the Systemic Initiatives do suffer, I think, in that they are “supply-side” reforms. That is, they are being driven by thousands of educators, professional developers, scientists and administrators who have a real interest in and passion for improving the quality of schools along certain lines. And yet, these people who are very earnest and hard working are producing and pushing a kind of high-minded reform that the public and local politicians have not yet embraced. There is not yet a strong and clear demand for the kinds of reform that the Systemic Initiatives have undertaken. To their credit, they recognize this fact and some, at least, have worked hard to engage the public in a dialogue about the need for improved instruction along the lines of the Standards and other reform documents.

In summary, I would say there are, indeed, real differences in what different people with different perspectives see as the current weaknesses in the system and the ways to fix them. And, no doubt, there are some weaknesses in the work of the Systemic Initiatives. But I would argue that these weaknesses come from lack of local capacity and not from some initial mistake in philosophy or ideology.

2) The analogy in your testimony of the Systemic Initiatives to the Apollo Program is an interesting one, but the fact is that the Apollo Program is considered a success we did make it to the moon. Presumably, if a school system is able to sustain continual improvement in the math and science instruction offered, eventually it will be reflected in student test scores. At what point do you think test scores are legitimate measures of the success of a Systemic Initiative?

I will give both a short and long answer to this question. First, the short answer:

The Systemic Initiatives represent a federal investment which is based on an assumption. The assumption is that the amount students learn depends on the quality of the instruction they receive. And the quality of instruction is shaped in complicated ways by all of the critical dimensions of the system that support our schools. Thus, to improve student achievement on a large scale, and in a non-trivial way, the Systemic Initiatives would argue that it is necessary to invest in the upgrading of all the key capacities of the system that support math and science teaching. This means that to help students learn more, there are no shortcuts and one has to work “upstream” and ultimately reform the system that shapes their instruction. Accordingly, I would argue that the NSF Systemic Initiatives have the prime and immediate purpose not of increasing student achievement, but rather of helping states and districts to analyze and improve their own support systems for math and science education.

Thus, it is only appropriate that the first evaluation criteria for the Systemic Initiatives be centered around measuring the degree to which and the ways in which the Initiatives are, in fact, influencing the national, state, and local systems that support math and science instruction. That is, it makes little sense to look at measures of student achievement until there is a satisfactory way to assure oneself that the Systemic Initiatives are significantly contributing to system elements such as state and district frameworks, assessments, professional development practices and curricula. There are many ways that the Systemic Initiatives are helping states and districts become more able to improve the teaching of math and science – the first step toward evaluating these Initiatives should be carefully studying and assessing those system improvements.

When we are convinced that the Systemic Initiatives have made significant contributions to state and local capacities, and when we are convinced that such increases in capacity have, in fact, resulted in significant changes in the quality of instruction that students are receiving, then it might be appropriate to look at actual changes in student achievement.

However, we are a long way from satisfying the first two evaluation criteria, and there are real dangers in trying to take shortcuts that prematurely attempt to link investments in systemic reform with increases in student achievement.

Now, in my longer answer, I will discuss these dangers more extensively.

Let me begin by exploring the Apollo Program analogy a little further. First, I think it is important to note some illuminating differences between going to the moon and improving our schools. (Note that it may well be easier to get to the moon than to significantly improve the instruction that is taking place in the nation's 16,000 independent school districts.)

First, the Apollo Program was almost entirely funded by NASA – a single federal agency. There was great consensus and clarity about the ultimate goal – getting to the moon. By contrast, the federal investment in K-12 education is on the order of 7% of all education funds – and most of this federal money is for large mainstream programs (such as Title 1), and not for initiatives that aim to bring about improvements through systemic reform. Thus, the federal investments made into the Systemic Initiatives is very, very small relative to the scale of the system they are seeking to influence. (It might be appropriate to imagine a different kind of Apollo program where each state and district was trying to get itself to the moon, and the federal government had very limited resources, and mandate, to help them do a better job of it.)

Also, in going to the moon, it was ultimately clear (at the end of ten years) that the goal had been accomplished. Probably along the way it was more difficult to ascertain whether the program was “on schedule” or making good progress as it developed the technologies and knowledge needed to reach the moon. In education the goal is nowhere near as clear and it is very difficult to know if we are “making progress.”

There are real issues with using test scores as a measure of the progress of the system. Let me list a few.

First, the current tests that are used on a large scale do not reflect the full range of knowledge, skills and attitudes that are specified in documents like the National Standards. I think there are few people who would argue that current tests are a complete and comprehensive measure of the knowledge and capacity of our students. Nor do I think many would argue that they represent the totality of what a student should know and be able to do. Because of economic, technical and political reasons, most state testing systems fall short of assessing the full breadth of their own state standards, not to mention the National Standards. I am not optimistic that this situation will change. (In fact, over time, it appears that most state testing systems have become simpler, less connected with Standards, and more focused on basic knowledge and skills.)

Second, there is also the problem that test scores measure the current status of a student’s knowledge. But I think it is very important to note that student achievement tests do not in any way measure the functioning of the educational system in which the student attends school. Study after study shows that the social and economic status of the student’s family accounts for a huge amount of the variance in test score achievement. Schools that are operating in a poor neighborhood may be offering students good instruction and still have relatively low test scores; similarly the test scores of a school in an affluent neighborhood may be high, but the school may, in fact, be doing a relatively poor job of providing challenging learning experiences for those students. It is the job of the schools to “add value” to students – to improve upon what they already know and are able to do. Today it is all too common to associate the absolute level of test scores with the quality of the educational system where students happen to reside. This, to my mind, is a fundamental misconception that is at the basis of most accountability “systems” and is also fundamentally flawed.

Some states (such as Tennessee) are working hard to develop a system that measures the “value-added” benefits of particular schools or even classrooms. Other states (such as New York) are seeking to compare the test scores of schools

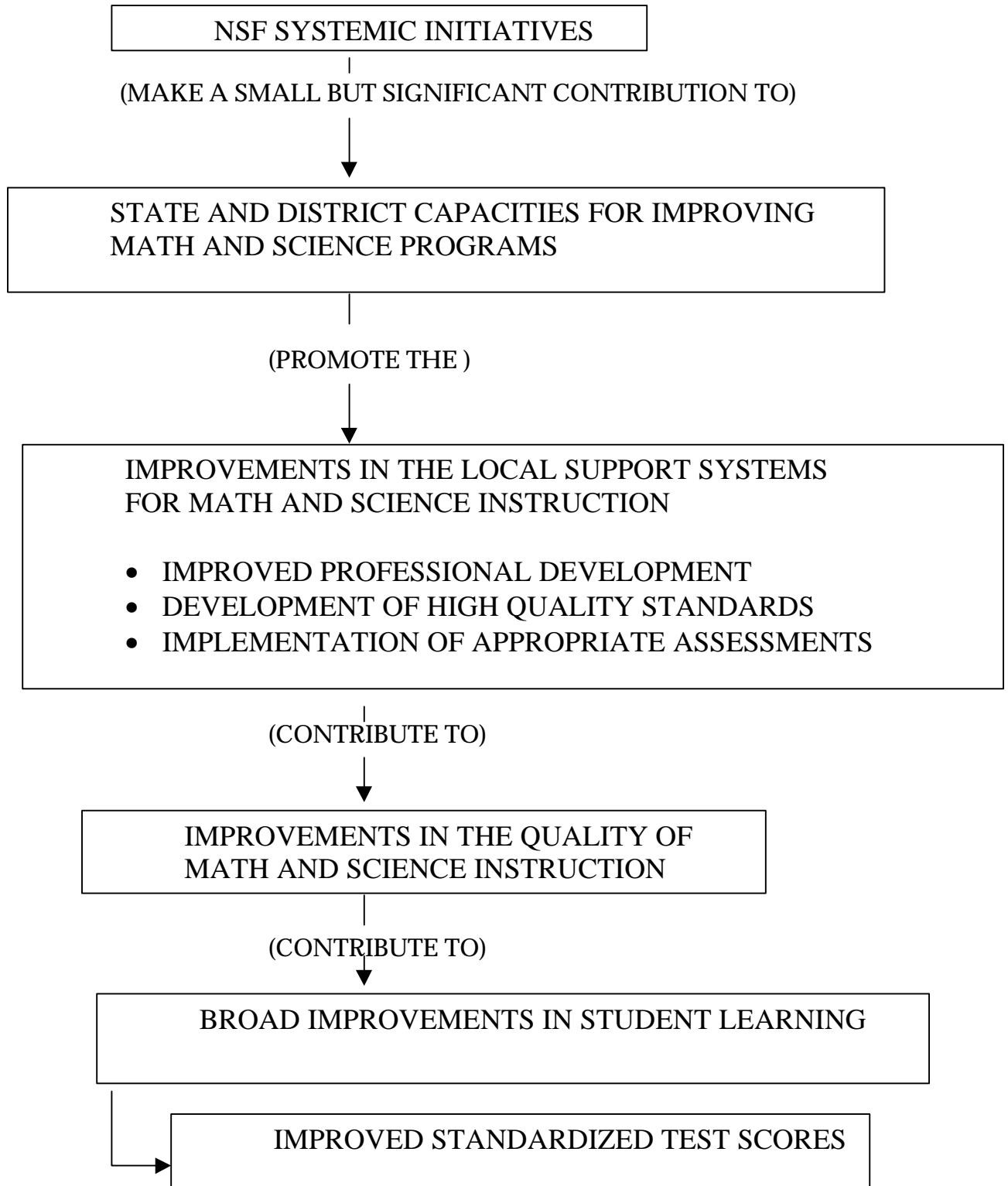
that are similar in the students they are serving. But all of these measures are crude and not yet valid for assessing the quality of schooling that students are receiving. Ultimately one needs to know whether schools are offering high quality instruction and providing students with challenging learning experiences. Test scores are not only a poor proxy for measuring this instructional reality, but they can be (and often are) downright misleading.

Third, in addition to the problem that test scores do not measure what the Standards specify, and in addition to the fact that test scores do not measure the value-added benefits of schooling, there are other difficulties with using test scores to measure the benefits of the Systemic initiatives. At best, the NSF Systemic Initiatives increase the capacity of states and districts to improve their own educational infrastructure. That is, they help states and districts develop better local standards, improve the quality of their assessments, train professional developers, adopt better curricula, etc. The Systemic Initiatives represent, then, a long-term and upstream investment. One would not expect to see results for many years; nor would one expect to see direct and visible connections between investments and improvements in student achievement as measured by very limited testing procedures.

Perhaps one more analogy would be useful here. The Systemic Initiatives are not unlike investments made in the infrastructure of “third world” and developing countries. If we invest, for example, in new irrigation equipment in Somalia, we would not look for an immediate decrease in the infant mortality rate. If we invest in economics education for emerging Russian entrepreneurs, we would not expect to see the GNP of the country immediately rise. I would argue that it is very important to conceptualize the Systemic Initiatives as similar kinds of long-term investments in infrastructure and to evaluate them on that basis.

Finally, I will make one more point here about the difficulty of using test scores as any kind of measure of the effectiveness of a Systemic Initiative. Not only does the Systemic Initiative work at the system level (far removed from most students), but also the scale of this investment is tiny relative to the scale of the system. The investments are typically at a level of less than a dollar per student per year. Even if there was a direct connection between the Systemic Initiative and the improvement of student achievement on standardized tests, the effect would likely be very small indeed.

The diagram below summarizes the difficulty of making the connection between the investments in Systemic Initiatives and improvements on student test scores:



3) Your testimony suggests that NSF needs to improve the way it provides local reformers with outside experts. What advice would you give on ways to do this?

I think the National Science Foundation has made real efforts to help states and districts gain access to experts. It has funded Westat, EDC and others to provide such assistance. But I think that NSF's success to date has been limited. I think it is important to note that the Systemic Initiatives represent a very different way of doing business. The traditional recipients of NSF grants – scientists, science educators, university faculty, etc. – are now being asked to play different roles. That is, they are being asked to play supportive rather than central roles. Thus, I think it is important to give NSF credit for trying to learn how to engineer these new relationships.

While I do not have the definitive answer to this question, I can think of several different ways to provide assistance to states and districts as they work on their Systemic Initiative programs.

a) Most of the leaders of the Systemic Initiatives would tell you, I believe, that one of the most valuable types of assistance they received was in having the opportunity to talk with and learn from others doing similar work. Thus, the NSF Principle Investigator meetings that allowed for the sharing of ideas, problems and successes were seen as useful. Probably more could be done to specifically engineer such opportunities not only for the Initiative leaders (the PIs) but also for the evaluators, professional developers, and others involved in the Initiatives.

b) One key role that the NSF plays is to provide states and districts with knowledge of and access to national expertise and national resources. For example, through the Systemic Initiatives many districts have learned about the NSF-funded and other innovative curricula that are now available in mathematics and science. They have also had the opportunity to consider the whole question of professional development design, and to examine issues of equity in mathematics and science education.

There are several ways to provide states and districts with knowledge of and access to national expertise and national resources. One is to bring in experts to talk to Systemic Initiative leaders. Another is to offer states and districts a chance to ask for help from such expertise.

One of the most effective ways that I have seen to make expertise and knowledge truly available and successful is to identify people and projects that are very good at doing a particular kind of work and then fund them to

serve as “national centers” for the Systemic Initiatives. Let me give some examples. For years the Exploratorium Science Museum in San Francisco has offered workshops that helped teachers understand and implement an inquiry approach in their classrooms. Recently, NSF funded these people to work with professional developers around the country so that they, too, could similarly help teachers with inquiry approaches in their own localities.

Similarly, NSF funded a center (EMELI) at the University of California, Santa Barbara, allowing them to draw upon their years of experience in providing workshops that helped educators identify and address issues of equity in their own educational settings. Now they are a resource for the nation that can provide in-depth training for state and district leaders, where the goal is to empower these leaders to address issues of equity in mathematics education.

In another example, NSF funded a group of scientist at the California Institute of Technology to support a science education center. The purpose of this Center (CAPSI) is to help other districts implement an elementary science model similar to a successful program that has evolved in Pasadena, CA. Here the work that the university scientists and district leaders did over a decade has resulted in a model of district change that is now being shared with other California districts.

Each of these examples is similar in that NSF identified people already doing good work along a particular dimension of systemic reform, and provided them with a grant so that they could develop programs that would share this work with the leaders of many other programs and Systemic Initiatives.

c) Networks are another mechanism that I would suggest that NSF consider as a mechanism for supporting the leaders and workers engaged in the Systemic Initiatives. For example, NSF could provide funding for some talented group to oversee and coordinate a long-term network of district science and math specialists. This network, which might function along the lines of the National Writing Project Network, would allow for the sharing of experience, the collection and documentation of successful practice, and most importantly, the development of leadership capacity at the district level. I think it is important that the networks NSF funds be more than a mechanism for doing the bidding of NSF or any other agency; rather they should be guided by the interests of the members of the network and should serve the needs of those members.

Networks might be created for teacher leaders, principals, state leaders and university faculty. (The New York State Systemic Initiative has a very interesting network of higher education faculty interested in inquiry and involved in K-12 reform work.) Networks can also be created for whole districts involved in reform. (The Education Development Center (EDC) currently conducts an NSF-supported network of urban school districts engaged in elementary science reform (CUSER). This Network allows leaders to meet with each other as well as learn from national experts.) Networks such as these, if funded over the long term, could become part of the national infrastructure that supports states and districts as they seek to improve their own programs. (The California Science Project and the California Math Project are good examples of long-term networks that serve as a state-level infrastructure and greatly assist California in promoting instructional improvements on a local level.)

4) Recently it was reported that more than half of the prospective teachers – graduates from American education colleges – in Massachusetts did not pass a standardized test of basic knowledge. Clearly, that is unacceptable. It doesn't seem unreasonable to insist that teachers in math and science have more than just a passing acquaintance with the subjects they teach. The Systemic Initiatives seek to improve teacher knowledge through professional development.

4.1) How would you rate this aspect of the Systemic Initiative program?

In short, I think that one of the major contributions of the Systemic Initiatives lies in the area of professional development. First, the resources that the Systemic Initiatives provide states and districts allow for a much higher level of professional development activity than otherwise would be possible. Second, the Systemic Initiatives have influenced the nature and quality of the professional development that is offered. There is no doubt that because of the NSF Systemic Initiatives many thousands of teachers have had learning opportunities that they otherwise would not have had.

However, I do have two reservations about the professional development aspect of the Systemic Initiatives. The first is that, at times, states and districts have tried to do too much too fast. Largely under NSF pressure, they sought to “scale up” their efforts before they had the professional development expertise and designs needed to assure high quality activities. Second, I think that at times the professional development component was not done “systemically,” that it was an isolated activity and not tied in with curricular, assessment or policy changes.

Finally, I think it is important to note again that the professional development offered by the Systemic Initiatives is a relatively small part of all the professional development offered by a state and district. The Systemic Initiatives contribute on a relatively small scale to the knowledge and teaching skills of teachers in a state or a district. (Note that even if the California State Systemic Initiative used all of its money for professional development, the expenditure would amount to less than \$10 per teacher per year. Thus, again I would argue that the Systemic Initiatives should be conceptualized as an effort that builds the capacity of California to do high quality professional development, rather than the doing of all the work itself.)

4.2) In your experience, does the professional development component reach into the education colleges in the states where it operates and what has been the impact on new education graduates?

I have less experience here. From what I have seen, the Systemic Initiatives have been most successful in working with teacher leaders, somewhat less successful in working with the “average” teacher, and even less influential in working with teacher candidates. Most efforts to combine professional development with pre-service programs have been, I think, largely experimental in nature. I think there are probably some success stories within the Systemic Initiatives which could be documented and shared, but I think there are few large-scale successes in this domain. (In fact, I believe NSF has a separate initiative which targets the improvement of pre-service education.)

4.3) How can this aspect of the program be improved?

My own study of pre-service reform is that it is a very difficult, even gridlocked domain. Pre-service programs mostly operate under state teacher credentialing laws, university degree and course requirements, and school district constraints. (For example, in one program I studied it took two to three years to get a new course approved by a faculty senate and included in the college catalogue.) There are also often great divides between the academic faculty who teach the undergraduate science and math courses, the education faculty who teach the methods courses and provide clinical supervision, and the master teachers who work in the schools.

I think that the Systemic Initiatives could help build the capacity of the system in the following ways:

- a) I think the most important thing that the NSF Systemic Initiatives might attempt to do is to influence who comes into science and math teaching. Teaching as a profession is losing its appeal and consequently draws poorly from those who graduate with degrees in mathematics and science. If NSF

could create programs, incentives and rewards so that successful undergraduates in mathematics and science could be drawn into teaching careers, that would go a long way toward improving the capacity of the system. As it is, all of the pre-service training and professional development in the world can not overcome a basic lack of talent.

b) NSF Initiatives can also involve university-level math and science faculty in K-12 reform activities, not only so that teachers and others can benefit from their disciplinary expertise, but also so that they may become interested in reforming their own undergraduate courses. There needs to be incentives and rewards for faculty to create appropriate learning experiences for undergraduates who are non-majors and who are likely to be teacher candidates.

But overall, I am not optimistic about the Systemic Initiatives having a major impact on pre-service education. Most of the Urban, State, and Rural Systemic Initiatives have their hands full simply in trying to work with their local school systems. While there is a need for fundamental reform in pre-service education, I am not sure it is wise to ask the Systemic Initiatives to spread themselves even thinner by taking on this immense challenge as well.

5) Mr. Baird's testimony suggested that the administrative burdens NSF imposes on grant recipients, particularly inconsistent and changing directives from NSF, made implementing the program more difficult in Florida. Based on your extensive experience assessing these programs:

5.1) Is this a problem that you have observed in other Systemic Initiatives?

Yes, this definitely has been a problem. Many of the leaders of Systemic Initiatives I have talked with have complained strongly about the loss of time and resources because of the NSF demands for accountability. Perhaps it was not necessary, but I have seen some Systemic Initiatives devote so much time and energy to meeting reporting requirements, mid-point reviews, and site visits, that they nearly dropped the programs they were supposed to be concentrating on.

To be clear, there are several problems here. One, which I have already discussed, is that NSF sometimes demands evidence of results in domains where the Systemic Initiatives should not be held accountable for results (e.g. student achievement). Another is that NSF at different times has demanded a lot of different things – annual reports, program performance reviews, management reports, strategic plans, evaluation reports along the dimensions of the “NSF drivers,” etc. The sheer volume of demands is, in itself, a problem.

Finally, there is a problem that results from changing expectations. NSF will freely admit that they change their demands of the Systemic Initiatives as they evolve. Even though they may approve a proposal at year zero, the NSF may well ask for a given program to do more as the program evolves – and as NSF’s own understanding of systemic change evolves.

5.2) As the Systemic Initiative program has matured over the years, has this become less of a problem?

Yes, I think so. The issue of accountability is complex. I have seen NSF interfere with projects and hinder their work. But I have also seen NSF interfere with Initiatives that were either poorly designed and/or implemented, and which needed strong intervention. Some Initiatives were “de-funded” which shows that NSF is not afraid of making hard decisions and cutting its losses. Other Initiatives were supported that probably should never have been funded. Some Principle Investigators have told me that they benefited greatly from their reviews and site visits even if they did not like them at the time. So the whole issue of accountability and managing these contracts is a complex one.

Let me suggest that there are actually some systemic issues that underlie the whole question of the relationship between NSF and its grantees. One is that the Initiative funding scheme may have contributed to the problem. For example, it may have been better to give cities and states up to three years of planning and development grant funding. (Many leaders of the Systemic Initiatives will tell you that they were really only ready to begin serious work after two to three years.) Also, NSF program officers are unable to travel to and spend time in the sites they are supposed to oversee. Consequently, NSF hired monitors (APT) and evaluators (SRI) to document the work. They did good work, but this added to the complexity of the relationship. Moreover, it was not atypical for a Systemic Initiative to have three different program officers over the life of the Initiative. All of these things led to difficulties in the relationships between grantees and NSF.

I think NSF is learning in this domain. I would recommend that NSF be encouraged to re-conceptualize its relationship to its grantees so that there are truly “cooperative” agreements between itself and the states and districts it funds. Too often in the past, NSF was more eager to hold “the feet of its grantees to the fire” than to support their success. This may have come from either real or imagined pressure from Congress to produce immediate and concrete results. There is a kind of “kicking the dog” syndrome in action here where each level in the hierarchical system places on the level below accountability demands which are sometimes unreasonable and even counterproductive. Clearly, the success of this Initiative

requires a positive and mutually supportive relationship between the funder and the grantees; more needs to be done to make sure such a relationship exists.

Finally, I would recommend that there be some sort of broad oversight of the Systemic Initiatives and how they are doing vis-à-vis their grantees and vis-à-vis their goals. To date the evaluation of the Systemic Initiatives has focused largely on the success of its individual grantees. This is not unlike evaluating a football team by evaluating the play of each of its team members. But ultimately, the team needs to be evaluated on the basis of its overall play, on the game plan of the coach, and the degree to which the whole team can work together. The Systemic Initiatives are, I think, a worthy experiment. Both the funder and the grantees are quite serious about doing good work, and about improving the quality of math and science instruction. I think these Initiatives should be both supported and studied from that point of view.