

This article was downloaded by: [Stanford University]

On: 27 September 2008

Access details: Access Details: [subscription number 776101540]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

APPLIED  
MEASUREMENT  
IN EDUCATION

## Applied Measurement in Education

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t775653631>

### Lessons Learned from the Process of Curriculum Developers' and Assessment Developers' Collaboration on the Development of Embedded Formative Assessments

Paul R. Brandon <sup>a</sup>; Donald B. Young <sup>a</sup>; Richard J. Shavelson <sup>b</sup>; Rachael Jones <sup>1 a</sup>; Carlos C. Ayala <sup>c</sup>; Maria Araceli Ruiz-Primo <sup>d</sup>; Yue Yin <sup>e</sup>; Miki K. Tomita <sup>b9</sup>; Erin Marie Furtak <sup>h</sup>

<sup>a</sup> College of Education, University of Hawaii, <sup>b</sup> School of Education, Stanford University, <sup>c</sup> School of Education, Sonoma State University, <sup>d</sup> University of Colorado at Denver and Health Sciences Center, <sup>e</sup> College of Education, University of Illinois, Chicago <sup>f</sup> College of Education, University of Hawaii, Manoa <sup>g</sup> The Curriculum Research and Development Group, University of Hawaii, <sup>h</sup> Max Planck Institute for Human Development,

Online Publication Date: 01 October 2008

EA LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS  
Mahwah, New Jersey London

**To cite this Article** Brandon, Paul R., Young, Donald B., Shavelson, Richard J., Jones <sup>1</sup>, Rachael, Ayala, Carlos C., Ruiz-Primo, Maria Araceli, Yin, Yue, Tomita, Miki K. and Furtak, Erin Marie(2008)'Lessons Learned from the Process of Curriculum Developers' and Assessment Developers' Collaboration on the Development of Embedded Formative Assessments',Applied Measurement in Education,21:4,390 — 402

**To link to this Article:** DOI: 10.1080/08957340802347886

**URL:** <http://dx.doi.org/10.1080/08957340802347886>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

# Lessons Learned from the Process of Curriculum Developers' and Assessment Developers' Collaboration on the Development of Embedded Formative Assessments

Paul R. Brandon and Donald B. Young  
*College of Education*  
*University of Hawaii*

Richard J. Shavelson  
*School of Education*  
*Stanford University*

Rachael Jones<sup>1</sup>  
*College of Education*  
*University of Hawaii*

Carlos C. Ayala  
*School of Education*  
*Sonoma State University*

Maria Araceli Ruiz-Primo  
*University of Colorado at Denver and Health Sciences Center*

Yue Yin  
*College of Education*  
*University of Illinois at Chicago*  
*College of Education*  
*University of Hawaii at Manoa*

---

<sup>1</sup>Rachel Jones is currently affiliated with Moanalua High School, Honolulu, HI.

Correspondence should be addressed to Paul R. Brandon, University of Hawaii, Curriculum Research and Development Group, 1776 University Avenue, Honolulu, HI 96822-2463.  
E-mail: brandon@hawaii.edu

Miki K. Tomita  
*School of Education*  
*Stanford University*  
*The Curriculum Research and Development Group*  
*University of Hawaii*

Erin Marie Furtak  
*Max Planck Institute for Human Development*

Our project to embed formative student assessments in the Foundational Approaches in Science Teaching curriculum required a close collaboration between curriculum developers at the Curriculum Research & Development Group (CRDG) and assessment developers at the Stanford Educational Assessment Laboratory (SEAL). This was a new endeavor for each organization, and throughout the project, many lessons were learned about embedding assessments and about the collaboration process. In this article, we discuss what we learned about the strengths and weaknesses of the collaboration up to the beginning of the randomized experiment. What we found comported with the literature on research collaborations. For example, past collaborations between CRDG and SEAL facilitated moving the project forward and sustained the collaboration. That said, the physical distance between the groups gave rise to some misunderstandings and led to a commitment to meet face-to-face on a regular basis; we found that conferencing software did not suffice. Moreover, in our zeal to implement formative assessments, the voices of teachers and teacher trainers got muffled until a pilot study confirmed their advice.

In the project described in this special issue, the Stanford Educational Assessment Laboratory (SEAL) at Stanford University and the Curriculum Research & Development Group (CRDG) at the University of Hawaii contributed largely disparate resources, experience, knowledge, and skills at a considerable distance from each other. These are conditions likely to be found in future collaborations among assessment developers and curriculum developers preparing formative assessments. In this article, we report on the process and decision points where the study took shape—something rarely reported. We believe research and evaluation would improve faster if retrospective accounts of design choices and their consequences were made available to the wider community. These accounts should consider how questions were chosen, how resources were deployed, how quality of data was controlled, how decisions were made, and how observations were assembled and communicated.

In this article, we draw on the literature about collaborative efforts among independent organizations and describe the major steps of our project during the

period until the implementation of the experiment in the six schools (see Furtak et al. and Yin et al., this issue), when the collaboration essentially ended. We show how the collaboration aspects identified in the literature were manifested in our project and present the lessons that we learned about conducting a collaboration to develop embedded formative student assessments.

## METHODS

The description and conclusions presented in this article are the results of a “study within a study.” That is, within the study of the development, implementation, and outcomes of embedded formative student assessments in the Foundational Approaches in Science Teaching (FAST) program, we examined the SEAL-CRDG collaboration and documented the features and events that helped and hindered the collaboration’s success. The first and fourth authors served as observers and recorders from the beginning steps of the project up to implementing the randomized experiment. The observer/recorders documented the process of the project as it progressed and wrote up the results in stages, while regularly obtaining feedback from the Assessment Development Team (ADT) team members about the thoroughness and accuracy of the description. At the end of the project, the team members discussed the lessons learned about the collaboration that might be helpful for future collaborators preparing embedded student formative assessments. We present these lessons in this article.

## INTER-ORGANIZATIONAL COLLABORATIONS

Researchers examining multi-organizational and multi-disciplinary collaborations that draw on the disparate skills and knowledge of the participating parties have defined them in common ways. Mattessich and Monsey (1992) defined an inter-organizational collaboration as “a mutually beneficial and well-defined relationship entered into by two or more organizations to achieve common goals” (p. 7; see also Hara, Solomon, Kim, & Sonnenwald, 2003). Hardy, Phillips, and Lawrence (2003) defined it as a “cooperative, inter-organizational relationship that is negotiated in an ongoing communicative process, and which relies on neither market nor hierarchical mechanisms of control” (p. 323). These collaborations are proliferating among universities, non-profit organizations, and so forth (Hara et al., 2003).

Baldwin and Austin (1995) reported that their review of the literature on collaborations among faculty in higher education research yielded a sparse collection of studies. Drawing from an array of studies on collaboration elsewhere, we found widespread agreement that the likelihood of success of collaborations

is enhanced by the degree of trust, mutual respect, interpersonal compatibility, prior working relationships, and established effective communication patterns among participating parties (Hara et al., 2003). The metaphor of *romances* and *marriages* comes up repeatedly in the literature (Baldwin & Austin, 1995; Hara et al., 2003; Kanter, 1994; Melin, 2000), as is the case, of course, in the present study. Elaborating on this metaphor, Kanter (1994) concluded that, similar to good marriages, successful collaborations between organizations require that all parties be aware of their own strengths and weaknesses, that there be successful personal relationships between the collaboration leaders, and that the organizations be compatible philosophically and strategically. Collaborations are enhanced as well when participants of similar professional standing are involved (Katz & Martin, 1997). The higher the status, the easier the collaboration, because the credibility of seasoned participants is well-established, and threats to advancement or reputation are diminished and few (Macduff & Netting, 2000).

The drawbacks of collaboration include the extensive time and additional resources that sometimes must be devoted to the collaborations (Baldwin & Austin, 1995; Cummings & Kiesler, 2005) and the hindrance of geographic distances between collaborating organizations (Katz & Martin, 1997; Melin, 2000). The greater the distance, the greater the need for conversations that are more structured than is necessary intra-organizationally (Baldwin & Austin, 1995; Cummings & Kiesler, 2005). Furthermore, recent research suggests that technological remedies have not successfully addressed difficulties due to distance among collaborating organizations (Cummings & Kiesler, 2005).

Hardy et al. (2003) reported that the greater the collaboration, the more knowledge is created (see also Powell, Koput, & Smith-Doer, 1996). Collaborations allow organizations to share knowledge and to develop new knowledge (Cummings & Kiesler, 2005; Melin, 2000), often due to the cross-fertilization that results from debates between parties with differing perspectives and ideas (Katz & Martin, 1997). Collaborations—especially between research organizations—have the strong potential to be constructivist enterprises; just as FAST students cooperating on scientific experiments collectively develop and share knowledge, inter-organizational collaborators share their existing knowledge, skills, and resources and develop new knowledge in their interaction.

## PROJECT PLANNING AND INITIAL ASSESSMENT DEVELOPMENT

### Initial Steps

The SEAL–CRDG embedded assessment project began with project planning and the initial steps of developing embedded formative assessments. The SEAL Principal Investigator (PI) met with the key CRDG project participants. The

group developed a common understanding of the knowledge-type framework (declarative, procedural, schematic, and strategic; see Shavelson et al., this issue) and the project goals and schedule. They decided that the assessments should address schematic knowledge, which the program's existing assessments did not already emphasize. They discussed the project timeline and decided to include a pilot test (which had not been included in the original National Science Foundation proposal), thereby lengthening the project by one year. They also discussed ADT member roles (described by Ayala et al. and Shavelson et al., this issue) and made the final decision about the FAST unit in which to embed assessments. CRDG recruited its members on the ADT, and the CRDG PI met the SEAL team at Stanford. Following the meeting, SEAL reviewed the literature on the development of the conception of buoyancy and density, developed a summary of the knowledge types that are addressed in the FAST program, and developed two ADT planning and training workshops.

These steps at the beginning of the first phase allowed the collaborators from the two universities to renew their previous working relationship. Rapport of this nature has been found to be a key feature of successful collaborations (Baldwin & Austin, 1995; Hara et al., 2003; Melin, 2000). "Trust, 'chemistry,' personal friendship, good communication, and mutual efforts to make things work" (Hara et al., 2003, p. 961) are essential characteristics of successful collaborations among team members and across disciplines. This proved true in the present study, in which team planning was expedited and enhanced because team members had a prior working relationship. Moreover meetings of representatives of both universities occurred and were essential at each site so that all participants shared the rapport.

The manner in which the sharing of expertise leads to new knowledge (Hardy, Phillips, & Lawrence, 2003; Katz & Martin, 1997) was manifested in our project by the realization that embedded assessments should address aspects of science knowledge—schematic knowledge—that the FAST curriculum assessments did not already address directly (see Ayala et al., this issue). The team also collaboratively realized that, contrary to their original plans, pilot tests of embedded assessments must be planned and administered before conducting full field tests.

## ADT Workshops

### *Workshop 1*

During the next two months, the ADT conducted two- and three-day workshops. The purpose of the first workshop was to introduce the project to the full ADT and to begin the first steps of developing formative embedded assessments. The SEAL PI described the purpose and methods of the project to the ADT members and trained them in how to identify assessment goals and select items.

The CRDG PI conducted hands-on demonstrations of the student investigations in the FAST unit for which embedded assessments were to be developed, and an ADT member from SEAL showed storyboards giving the types of knowledge addressed in each investigation. The ADT members then divided into two sub-teams, developed statements of the goal of the FAST unit, and reconvened as a group to develop a single statement. This served to develop a common vision. The workshop concluded by brainstorming ideas about the summative end-unit assessment. The SEAL members of the ADT made the most suggestions about assessment possibilities, but the CRDG members also had some novel ideas.

Both the group of CRDG staff members and the group of SEAL staff members serving on the ADT had steep learning curves. In particular, the ADT members found it more difficult than expected to develop the end-unit goal statement. Some of the ADT members suggested that it would have been helpful to present a sample goal statement to the groups before they began their discussions. This would have helped inform the curriculum developers about the form and nature of good goal statements and thereby might have helped focus the two sub-teams' efforts. The ADT members also suggested that they should have been instructed to attend closely to the knowledge types, which they ignored when developing the goal statement. These oversights might have been due, in part, to our familiarity and prior collaboration.

### *Workshop 2*

The purpose of the second workshop was to plan collaboratively the formative embedded assessments. The ADT identified the "joints" in the unit at which to embed the formative assessments, selected the assessment types for each joint, discussed concept maps and when to use them in the assessments, and discussed the next steps in the project. When selecting where to embed the assessments, the curriculum developers contributed their knowledge of the unit goals, and the assessment developers made it clear that the assessments should be embedded at points where understanding student knowledge would be most useful to teachers and students.

The contributions of the two organizations to the process reinforced the appropriateness of the plan to constitute an ADT with members who had expertise in the appropriate domains. For example, the FAST teacher and one of the trainers on the team tended to be concerned about the length of time the assessments would add to each lesson. One ADT member from CRDG felt that student progress might be slowed by the embedded assessments. In response to this concern, a SEAL team member noted that the assessment should help teachers decrease their teaching time because they know the students' deficiencies. Discussions such as these helped develop the ADT members' understanding of formative assessment.

The ADT members' contributions to the two workshops most likely were affected by the other members' sources of influence (Bacharach & Lawler, 1980) on workshop discussions and decisions. Participation is equitable to the extent that members' sources of influence do not interfere with other members' full and forthright participation. One of these sources is job position. The team members' positions potentially could have interfered because some members had more prestigious positions (both inside and outside of the workshop) than others. This has been a finding of some of the research on collaboration among disciplines and organizations (Macduff & Netting, 2000). The differences in the effects of position were perceived the most strongly by the FAST trainer and FAST teacher members of the ADT. In post-workshop interviews, they noted that they sometimes had to express concerns about some assessments (e.g., that the assessments seemed too lengthy or too difficult) rather forcefully. The trainer and the teacher said that on occasion that they felt somewhat uncomfortable expressing their concerns. However, they were satisfied that their concerns were heard, and other ADT members expressed gratitude that the teachers contributed their expertise.

## ASSESSMENT DEVELOPMENT, PILOT-TESTING, AND REVISION

### Assessment Development and Initial Review

The purpose of the next phase of the project was to develop the embedded assessments, pilot test them, and revise them in preparation for the randomized experiment. The process began with SEAL staff developing draft assessments and trying them out with middle-school students. They then posted the drafts on a Yahoo Groups website for the CRDG team members' review. The purpose of these steps was for the SEAL team members to apply their assessment expertise to item development and the CRDG members to verify alignment of items with FAST and to judge the appropriateness and quality of the items.

The use of Internet technology to review drafts was much less successful than anticipated. It reflected the caution expressed in the literature about the complications caused by distance between the locations of collaborators. SEAL members were less clear about the purpose of the review than desirable, and CRDG members were distracted by the demands of other activities and did not monitor the drafts as they were posted. SEAL members revised the instruments several times, and CRDG members lost track of the revisions. Also, SEAL members did not describe the intended context for the items, so some of the assessments, especially the concept maps, surprised and worried CRDG staff. When SEAL and CRDG staff finally met to review the draft assessments, they found that they could not identify the most recent versions of the assessments or connect the

FAST content targets with the appropriate assessments. The purpose of the meeting was to discuss glitches in the review of assessments and to brief CRDG on the link between assessments and FAST. From that point on, the two groups endeavored to meet face-to-face more frequently.

Posting draft assessments on the Yahoo Groups site was not a viable method to deal with the distance—not an uncommon occurrence in collaboration among institutions. As Cummings and Kiesler (2005, p. 704) noted about inter-organizational collaborations, “Recent research suggests that, even with some signs of progress, technology has not yet conquered distance.”

## Pilot Testing

### *Pilot of the training in how to use embedded assessments*

The next steps were to conduct a pilot test and to use the results of the pilot test to revise the assessments. The pilot test included training volunteer FAST teachers at CRDG’s University Laboratory School during the summer and having the teachers use the embedded assessments in the classroom during the following school year. CRDG staff recruited four FAST teachers (one a teacher at the Laboratory School and three Hawaii public school teachers) for the pilot test. The Laboratory School FAST teacher, a member of the ADT, participated in the preparation for training. Her suggestions about the use and the implementation of the assessments guided the SEAL assessment specialists, especially as to feasibility and usefulness.

In the teacher training, the ADT members introduced the project and the knowledge types to the teachers. They reviewed the assessments at each “natural joint” of FAST and had the three public school teachers (a) carry out the assessments as students, (b) review actual student work, (c) observe the Laboratory School FAST teacher providing feedback to summer school students in her classroom using the assessments, and (d) review different kinds of feedback. The teachers then administered each assessment with the summer school students and interviewed the students about their opinions about the assessments.

The ADT members commented on the importance of having a laboratory school to introduce teachers to the embedded assessments through observation and through implementing the assessments. The availability of and access to the Laboratory School was one of the key advantages of the SEAL–CRDG collaboration, because the iterative development and testing of assessments requires easy access to classrooms. It is a prime example of the sharing of resources that can occur in an inter-organizational collaboration.

The teachers expressed concerns about the time requirements for discussing students’ performance on assessment items in class, the high reading levels of the assessments, and the appropriateness of the assessments for mainstreamed

special education students and lower-achieving students. In response, the ADT stated that the teachers could decide how much time to allot to the assessments. The team stressed that the assessments were teaching tools and that their purpose was to give the teachers snapshots of student learning.

### *Piloting of the use of embedded assessments in the classroom*

The four teachers piloted the assessments in their classrooms during the school year following the summer training. The experience of the three public-school teachers showed serious shortcomings in the conceptualization and form of the assessments, as described in depth by Ayala et al. (this issue). The three public school teachers did not use the embedded formative assessment results to produce “teachable moments.” The ADT members concluded that their expectations about how the assessments would be used to provide student feedback were probably unrealistic, in part because they were based on the pre-pilot-test experiences of an exemplary FAST ADT member. It was clear to the ADT that it needed to revise the methods for having teachers use embedded assessment results to achieve desired teaching practices.

Assessment and curriculum developers should heed the advice of practicing teachers about feasibility. The FAST teachers serving on the ADT tried to keep the pilot testing practical, but others did not heed their advice. This was an instance of the hindrance to knowledge sharing that can occur between collaborating organizations.

A three-day meeting to assess the results of the pilot test and to plan the field experiment demonstrated the collaboration of the ADT. This was a pivotal point in the direction of the project where both SEAL and CRDG members provided important feedback and played a significant role in shaping the future decisions of the project. It was at this meeting that the assessments were restructured and given a new focus as a result of the pilot teachers’ experiences.

Based on the discussions among the ADT members, SEAL developed “reflective lessons” and the methods and materials for training the teachers in how to implement them (see Ayala et al., this issue.) SEAL developed the implementation framework; CRDG provided insights about the extent to which the implementation model adequately addressed the intended curriculum.

## SELECTING A SAMPLE FOR THE RANDOMIZED EXPERIMENT

In preparing for the randomized experiment, CRDG staff identified a pool of FAST teachers from which the treatment and control groups could be sampled. The CRDG members of the ADT reviewed past FAST teacher training records to

identify possible teachers to participate in the forthcoming field test. The records were much less complete than CRDG had previously described to SEAL, in large part because records from a computerized teacher database had been lost. CRDG staff identified potential participants for the field test by reviewing paper records and consulting the FAST trainers. They had difficulty contacting teachers because some were no longer teaching, some who were teaching were no longer teaching FAST, and some had gone to unknown schools; also, many teachers' only point of contact was via the school, but they were unavailable by phone during the school day. Many fewer Hawaii teachers than anticipated were reached, and, contrary to previous expectations, few California teachers could be identified and none could be recruited.

When CRDG identified a small pool of interested teachers, they obtained information on the teachers' schools' student ethnic distributions, percentages of students receiving free and reduced lunch, and achievement test scores from a national school-demographics website (see Shavelson et al., this issue, for details). Using the school data, SEAL instructed CRDG about stratifying and matching pairs of the sample of 12 teachers on school characteristics, and then randomly assigning 6 per group. Unfortunately, the school data base, unbeknownst to CRDG, was out of date. Moreover, the achievement test scores were for varying years, grades, and instruments, resulting in poor statistics with which to stratify teachers and randomly assign them to treatment and control groups. Furthermore, some characteristics of the selected teachers (such as two teachers being located in different schools at the same physical plant and one teacher primarily teaching university faculty's children) could not have been learned from demographic data available on the Internet. The project would have been better served had some ADT members visited the sites before settling on the teachers for the study, but due to cost, this was not considered.

### SUMMARY OF LESSONS LEARNED

The often-mentioned metaphors of *romances* and *marriages* in the literature on inter-organizational collaborations serve to emphasize the aspects of collaborations that are necessary if they are to succeed. Collaborating organizations are most likely to work together successfully when they are compatible philosophically; show self-awareness of the resources, experience, knowledge, and skills that they contribute to mutual endeavors; and are led by people with similar professional standing and a prior successful working relationship that exhibited trust and "chemistry." They succeed to the extent to which they respect each other's contributions to the arrangement, irrespective of the formal employment positions that they hold in their respective organizations. The closer the geographic distance,

the greater the likelihood of success; like romances and marriages, long distances can be a hindrance. Furthermore, they are more likely to succeed when they each can contribute the knowledge, skills, and material resources that the other lacks. These contributions often lead to new knowledge.

### Lessons Learned About the Strengths of the Collaboration

These aspects of collaborations aptly depict the SEAL-CRDG collaboration and help frame the lessons that we learned about the strengths and weaknesses of the collaboration. One of the strengths was that the SEAL and CRDG PIs had worked together previously, and other CRDG staff members had worked with the SEAL PI. There was considerable trust and respect between the participating parties. Another strength was that the two PIs were both well known and respected within their respective professional fields. A third strength was characterized by the unique sets of expertise that each group as a whole and the members individually brought to the project. The SEAL staff participating on the ADT contributed knowledge and skills about developing assessments, and the CRDG staff contributed knowledge and skills about curriculum development and the FAST program in particular. The groups were motivated to work together because each group had this expertise.

Beyond these generalities, there were numerous specific instances, apparent throughout all the steps of the project, of the ADT members' intellectual and material resource contributions. These included (a) the teachers' cautions about the feasibility of embedding assessments, (b) the value of having a laboratory school for training teachers in how to use the assessments, and (c) the contributions that curriculum developers made when brainstorming about types of formative assessments to use. Furthermore, sharing knowledge and skills led to new knowledge—for example, that embedded assessments should contribute new measures to a program's existing array of assessments, that embedded assessment projects should include pilot tests, and that teachers do not use embedded assessments in the manner prescribed and are more likely to incorporate them into teaching when the assessments are described and used as additional pedagogical tools ("reflective lessons").

### Lessons Learned About the Weaknesses of the Collaboration

The ADT learned a number of lessons about the glitches that can occur in inter-organizational "romances." These primarily had to do with (a) giving insufficient attention to the suggestions of the FAST teachers and trainers about the feasibility of the assessments, (b) hindrances caused by the geographic distance and the failure of technological means for addressing the distance, and (c) the difficulties in identifying the population of FAST teachers to sample for the randomized field experiment.

Some of the problems that arose when piloting the formative assessments with the group of four teachers could have been avoided had the ADT paid closer attention to the cautions of the FAST teacher and FAST trainer members of the ADT. When the FAST teacher and trainer expressed concerns about the time requirements for implementing the assessments, the ADT responded that the teachers could decide how much time to spend on them. This resulted in less effective implementation of the assessments than expected and less use of the assessment results. It became clear that assessment developers need to attend to issues of assessment practicality and feasibility.

The second major lesson that we learned was about the difficulties of conducting a collaboration over a lengthy distance. In particular, we found that the use of the Internet was not successful. It took a face-to-face meeting to address issues of the alignment of assessment items with the program and to plan future steps. Indeed, the ADT endeavored to meet more in person from that point forward.

The third major lesson was that identifying a pool of teachers to sample for the randomized experiment was more difficult than anticipated. Records of teachers using any school program are likely to be quickly out of date because teachers stop using the program, change teaching locations, or leave the profession. Furthermore, valid and timely school demographic and achievement data are often difficult to obtain for matching teachers between treatment and control groups. The ADT assumed that this information was readily available, but most organizations would find it difficult to muster the resources to obtain it.

## CONCLUSION

Romances often begin with blind infatuation that lessens after the two parties come to know each other's strengths and weaknesses, find the need to respond to each other's exigencies, are constrained by context, and become absorbed in themselves. Our "romance" by and large followed this familiar pattern. After the initial rush of enthusiasm of the ADT members, we learned more about what each side needed to achieve our goals, addressed our separate needs as they arose, faced environmental distractions, and sometimes listened less than desirable. For example, long-distance collaboration via the Internet was less successful than necessary, CRDG was unable to provide a sufficient number of sites for the randomized experiment, and SEAL failed to heed the FAST trainers' and teacher's concerns about the practicality of the assessments. Furthermore, both the SEAL and CRDG team members were somewhat blinded by their initial zeal to implement formative assessments and experienced a steep learning curve when developing what eventually became known as reflective lessons. Eventually, however, the mutual respect of each team member for the others moved the project forward. Both the curriculum developers and the assessment developers

profited immensely from the learning about embedded formative assessments that was gained during the project.

## ACKNOWLEDGMENTS

This research was supported, in part, by a grant from the National Science Foundation (NSF/Award ESI-0095520) and, in part, by the National Center for Research on Evaluation, Standards, and Testing (CRESST/Award 0070 G CC908-A-10).

## REFERENCES

- Bacharach, S. B., & Lawler, E. J. (1980). *Power and politics in organizations*. San Francisco: Jossey-Bass.
- Baldwin, R. G., & Austin, A. E. (1995). Toward greater understanding of faculty research collaboration. *The Review of Higher Education*, 19(2), 45–70.
- Cummings, J. N., & Kiesler, S. (2005). Collaborative research across disciplinary and organizational boundaries. *Social Studies of Science*, 35, 703–722.
- Hara, N., Solomon, P., Kim, S. L., & Sonnenwald, D. H. (2003). An emerging view of scientific collaboration: Scientists' perspectives on collaboration and factors that impact collaboration. *Journal of the American Society for Information Science and Technology*, 54, 952–965.
- Hardy, C., Phillips, N., & Lawrence, T. B. (2003). Resources, knowledge and influence: The organizational effects of interorganizational collaboration. *Journal of Management Studies*, 40, 321–347.
- Kanter, R. M. (1994, July–August). Collaborative advantage: Successful partnerships manage the relationship, not just the deal. *Harvard Business Review*, 21(4), 96–108.
- Katz, J. S., & Martin, B. R. (1997). What is research collaboration? *Research Policy*, 26, 1–18.
- Macduff, N., & Netting, F. E. (2000). Lessons learned from a practitioner-academician collaboration. *Nonprofit and Voluntary Sector Quarterly*, 29(1), 46–60.
- Mattessich, P. W., & Monsey, B. R. (1992). *Collaboration: What makes it work*. St. Paul, MN: Amherst H. Wilder Foundation.
- Melin, G. (2000). Pragmatism and self-organization research on collaboration on the individual level. *Research Policy*, 29, 31–40.
- Powell, W. W., Koput, K. W., & Smith-Doer, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41, 116–145.