Designing a Performance Assessment to Measure Student Communication Skills in Multimedia-Supported Project-Based Learning

William R. Penuel

Barbara Means

Center for Technology in Learning

SRI International

Running Head: EVALUATING THE MULTIMEDIA PROJECT

Paper to be presented at the annual meeting of the American Educational Research Association, April 2000, New Orleans, LA. The research described in this paper was supported in part by U.S. Department of Education grant #303A50357 to the San Mateo County Office of Education. The opinions expressed are those of the authors and do not imply endorsement by the county office or the Department of Education.

Abstract

This paper describes the design, implementation, and results from a performance assessment aimed at measuring student communication and design skills in the context of constructing multimedia products for a specific purpose and audience. The assessment was developed and administered as part of the evaluation of the Challenge 2000 Multimedia Project, an effort to provide teachers with the professional development and other support they would need to implement classroom projects within which students use multimedia technology. The performance assessment was used to compare the performance of middle-school students in classrooms with Multimedia Project teachers with middle-school students in classrooms where teachers did little with technology or project-based learning. Results showed that Multimedia Project students evidenced greater mastery of content, more sensitivity to their audience, and better design skills than students in classes taught by teachers who did not participate in the project. Students in classes of teachers who had spent more time in the project produced better products than those in classes of teachers who had spent less time in the project.

Since 1995, SRI International has been conducting an evaluation of the Challenge 2000 Multimedia Project, a federal Technology Innovation Challenge Grant effort aimed at infusing Silicon Valley classrooms with a model of technology use in which students acquire curriculum knowledge and composition and presentation skills as they pursue interdisciplinary projects built around real-world problems. The evaluation has incorporated a variety of data collection methodologies—case studies, surveys, observations, and analyses of school-wide indicators of learning—to gauge how widely the multimedia-supported project-based learning model has been implemented in participating schools. (See Penuel, Golan, Means, & Korbak, 2000 for these data.)

The multimedia-supported project-based learning model calls for teachers to organize their students into small groups which pursue research, design, or communication learning goals over an extended period of time and express what they have learned and their ideas through some form of multimedia (e.g., a Web page or HyperStudio or PowerPoint presentation). Exhibit 1 contains a description of one such classroom project as designed and implemented by one of the Multimedia Project teachers. In addition to documenting the frequency, quality, and implementation issues connected with these classroom projects, the evaluation has examined the question of whether the typical teaching and learning experience in classrooms of Multimedia Project teachers has changed over the course of the project. Independent observations of classroom practice have found differences in the typical organization of instruction and in the nature of the assignments in classrooms of Multimedia Project teachers. Compared to other teachers in their districts, these teachers were more likely to be acting as a facilitator rather than lecturing or asking questions during the classroom observation periods; they were more likely to have their students involved in small-group activities; and, during the spring semester, more likely to have their students working on long-term projects (Means & Golan, 1998; Penuel, Golan, Means, & Korbak, 2000). While changes in teaching practices were an important goal for the Multimedia Project, by themselves they do not constitute proof that students are learning more in these classes. In the fourth year of the evaluation (1998-99), we addressed the question: Are students who are completing multimedia projects learning something that their peers in other classrooms are not?

The most obvious area in which differential skills could be expected concerned the use of multimedia software per se. Observations of classroom activities and of student work products demonstrated that the students of Multimedia Project teachers were acquiring skills with software programs such as KidPix, HyperStudio, PowerPoint, and various Web page building tools. Comparing students in Multimedia Project classrooms to students in classrooms of other teachers, who had less (and often no) access to these technologies was deemed superfluous. (The percentage of teachers in these districts overall who had students use multimedia presentation software was just 9%.)

A more potentially informative research question was the effect of multimedia project experiences on students' learning of academic subject matter. We began with hopes of comparing classrooms in terms of the students' acquisition of subject matter knowledge, on the assumption that students who spent many more hours researching a topic and presenting their views on it through a multimedia product would have better knowledge of that topic, as indeed found in experiments by Spoehr (1992) and Lehrer (1992). The Multimedia Project, however, involved K-12 teachers from nine different districts with very different student bodies and curriculum foci. Teachers were encouraged to develop projects that fit their particular curriculum and students, and the fact that they did so meant that there were few classroom multimedia projects dealing with the same topic, yet alone appropriate comparison classrooms addressing the same material. The norm-referenced standardized tests used by districts and the state of California for accountability purposes offered a poor match for the kinds of learning that take place in innovative, student-centered classrooms (Resnick, 1987; Sternberg, 1985). Accordingly, we were led to the decision to provide our own content to a set of classrooms as well matched as we could make them in order to observe what students in the classrooms of Multimedia Project and comparison teachers would do with a task similar to a classroom multimedia project in many respects except for the requirement to actually use technology.

In setting out to design a way to measure the impact of the Multimedia Project on student learning, we sought to design an authentic task that would include many of the features of performance tasks like those found in the *Jasper* Challenge series (CTGV, 1997). Following the approach taken by researchers at the Center for the Study of Evaluation at CRESST in UCLA, we developed a model of student learning in the Multimedia Project that we used as a guiding framework in designing our assessment (Baker, 1998; Baker, Aschbacher, Nieme, & Sato, 1992; Baker, Aschbacher, Niemi, Sato, & Redfield, 1992). The model begins with a decision by designers concerning the

specific kinds of learning to be assessed (e.g., problem solving, communication, collaboration) and then embeds this learning within a specific task. Specifications for the task are then developed to define the content, task process, and scoring approach to be used.

Constructing a Performance Assessment for the Multimedia Project

We used two main sources of data to develop a model of learning to guide the development of the performance assessment task. First, we consulted project documents about the goals of the project. The project objectives articulated in most official documents were broad, emphasizing student learning in the areas of curricular content, collaboration, communication, and problem solving. Similarly, annual judging of student Multimedia Fair projects nominated by their schools as exemplary, Multimedia Project staff had been using a rubric organized into three broad areas: (1) content; (2) design; and (3) collaboration.

Given the lack of any single subject area with adequate numbers of student products to support a research design, we decided to concentrate on the process skills the project experience was expected to support. Prior research on technology supports for project-based learning provided guidance in identifying more specific, measurable skills that students can attain through these experiences.

Carver, Lehrer, Connell, and Erickson (1992) identified a set of design skills that can be expected to emerge when students work on a series of multimedia products:

- Allocating resources and time to different segments of a project
- Searching for information
- Analyzing and interpreting information
- Developing representations of information
- Developing a structure for a presentation
- Catching and maintaining audience interest
- Evaluating the process and reflecting on one's product.

Subsequent research by these and other researchers has borne out the potential of multimedia development activities to support the development of these design skills and the emergence of self-regulation of learning among students who participate in extended projects. For example, Allen and Pea (1992) traced the joint construction of a set of expectations for student learning by teachers and students over the course of an extended

student project involving the construction of multimedia presentations. They documented the ways that teachers and students negotiated a balance between a focus on design and one on content and came to see the two elements of constructing a presentation as interdependent. In a similar study of project-based learning with multimedia, Erickson and Lehrer (1998) researched the evolution of "critical standards" for judging the quality of student work. Over the course of 2 years' involvement in projects, students and teachers came to develop shared representations of what constitutes a good multimedia product.

The set of skills taken from Carver et al. provided a framework for designing a task for use in a performance assessment for the Multimedia Project. The task had to provide opportunities for students to exhibit each of these skills. Further, since we wanted to compare students of Multimedia Project teachers with those of similar teachers who did not take part in the project, the exhibition of these skills had to be possible without technology supports. That is, the skills we sought to measure were things that we believed students could learn well in a technology environment, but they are broadly useful skills that are used off-line as well as on-line and whose assessment does not depend on the availability of computers and familiarity with multimedia software. We were also constrained by the requirement to assess a large number of student groups simultaneously in order to meet the needs for statistical power and of data collection efficiency. This requirement limited the degree to which we could capture student collaboration skills for every group, pushing us toward use of student self-reporting with respect to this important objective.

Study Sample

Students in grades K-12 participate in the Challenge 2000 Multimedia Project, and one would expect markedly different levels of design and communication skills at the different developmental levels encompassed in this age range. Because our project resources would not support creation of separate assessments for every developmental level, we chose to develop our assessment task for middle-school-aged students in grades 6 and 7. There were a large number of classrooms with experienced Multimedia Project teachers in these grade levels, allowing for a large enough sample size for a study of the impact of the project. Furthermore, at these grade levels, many students are beginning to develop deeper metacognitive skills (Flavell et al., 1981) that are aids to the design of projects. As students are beginning to develop such skills, one might expect scaffolding

through project-based learning with multimedia to support more rapid development of design skills, resulting in a difference between project and comparison classrooms.

The 12 project classrooms selected for the performance assessment (and for classroom observations, as described in Penuel, Golan, Means & Korbak, 2000) included a combination of experienced and novice teachers within the Multimedia Project. All of these Multimedia Project teachers were funded with mini-grants ranging in size from $3,500-$7,000 to support their technology activities for the 1998-99 school year. For each participating Multimedia Project teacher, the principal from the teacher's schools was asked to identify as appropriate comparison teacher. Principals were given instructions to select a teacher in the same grade who was not part of the Multimedia Project but who taught in the same or a similar subject area and had similar students. Because the Multimedia Project encouraged teacher partnerships within schools, finding a comparison non-project teacher at the same grade level in the same school was not always possible. In two cases, classrooms from the same grade level at a comparable school in the same district were chosen. Even with this flexibility, we were able to recruit only 6 comparison classrooms for participation in the observations and performance assessment. Still, the resulting classrooms were similar in size and in demographic composition, as shown in Table 1.

Table 1

COMPOSITION OF CLASSROOMS IN THE STUDY

	MM Project Classrooms	**Comparison Classrooms**
Average attendance	27.5	28.4
Ethnic composition		
White	56%	61%
Asian/Pacific Islander	20%	17%
Latino	15%	16%
African American	2%	4%
Other	7%	2%
Average Age	11.9 years	11.7 years
Average Assigned Grade Level	6.27	6.18
% Groups with more than 50% of ELL students	31.0%	30.3%
Mean Achievement Levels on SAT-9 (in NCEs)		
Total Reading 1998	59.46	60.83
Total Reading 1999	62.66	63.60
Total Language 1998	61.94	64.75
Total Language 1999	63.01	64.08

Although similar in terms of class size (average attendance) and ethnic composition, Multimedia Project and comparison classrooms did differ on one significant measure, namely, the number of computers that were in their classrooms. On average, Multimedia Project classrooms had six computers, whereas comparison classrooms had only two.

Assessment Task Specifications

Because the study included classrooms from multiple subject areas, we needed to provide students with the content they would require to complete the assessment task (as well as related but nonessential content). To permit assessment of the skills identified by Carver, Lehrer, Connell, and Erickson, we had to design a task that would require students to analyze and synthesize this content as part of their design process.

We decided to develop a performance task similar in some respects to the kind used in social studies, especially history instruction, called a document-based question (DBQ). Document-based questions are used in advanced placement examinations and in student-centered history classes to encourage deep student thinking about historical problems. They require students to review primary documents, interpret them in the context of a problem or question, and frame a response using the content provided.

Although our task did not involve a historical question for students to answer, we did present students with social studies content for which they had to solve a problem in design. We chose to engage students in addressing a social problem that was experienced by students somewhat younger than themselves, namely, school difficulties faced by elementary-aged students who are homeless. Our aim was to choose a topic that we believed would be engaging for students and would require them to use information provided to them from both personalized accounts and evaluative data to develop strategies for addressing problems faced by homeless students. This problem context required students to engage in literacy practices, including reading for understanding, and in mathematical interpretation and data analysis. The documents provided to students for use in the task are available on the Multimedia Project's Web site (http://pblmm.k12.ca.us/evaluation).

The performance assessment task required students to work in small groups of four or five to design a brochure aimed at principals and teachers of elementary schools to help them become aware of the problems faced by homeless students and of strategies

they might use to help address those problems. Students were read an introduction to the problem and were advised to use the documents provided to them to guide their design process. Students were told that there were a large number of documents and were advised to work with their team to develop a strategy to read through the documents and complete a draft brochure in the short 1-hour time frame they were given. After one-half-hour, students were given a piece of paper folded in three and told to use this paper to develop their brochure. At the end of the hour, students completed a reflection survey, which invited them to reflect on their experiences in designing the brochure and to assess the quality of their collaboration process and finished product.

The model of design developed by Carver et al. was used to guide development of the performance assessment task:

1. *Allocating resources and time to different segments of a project.* Students had a short time to complete the project and had to decide among themselves a process that would enable them to complete a draft brochure by themselves.
2. *Searching for information*. Several types of information were provided to students. Not all of it was of immediate use in constructing the brochure, and no document was labeled in such a way to indicate how it should be used.
3. *Analyzing and interpreting information.* Students had to infer what would be the most useful information for their brochure. For example, they had to analyze graphs showing evaluative data from homeless education programs and infer that this type of program might be an example of an effective strategy for addressing the problems of homeless students.
4. *Developing representations of information.* Students had to transform the text they received in the documents into a brochure aimed at a particular audience. Students had to integrate text, graphics and images into their brochures.
5. *Developing a structure for a presentation.* Students were told to incorporate at least three elements into their brochure: a description of the problems homeless students face; a description of some proposed solutions or strategies for addressing those problems; and an argument about why their proposed solutions would work. It was up to students to determine how to address these specific concerns, using the content provided to generate a coherent brochure.
6. *Catching and maintaining audience interest.* A key feature of the task was developing a presentation that would be eye-catching and convincing to teachers and administrators. In addition, we had secured agreements from a homeless education program to take the best ideas from students' work and develop a brochure for actual use in schools, and we told students this to encourage them to work on the project as if their brochures might eventually be used in schools.
7. *Evaluating the process and reflecting on one's product.* As students were completing their brochures, each group was invited to look at the brochure checklist, a list of features their brochure should include. These features

reflected the original directions, plus some design features not explicitly addressed in the original instructions. These features would be incorporated into the scoring rubric developed to judge the quality of the students' brochures. In addition, students were asked to complete a reflection survey. The reflection survey included questions about their interest in the task, the perceived difficulty of the task, the quality of collaboration within their group, and the anticipated response of their audience to the brochure. Students were also asked how they might revise their brochure if given additional time.

Limitations to the Task Design

There were, to be sure, some important limitations to the task design. First, the task did not require students to engage in a project over an extended period of time (weeks or even months) in the way that multimedia projects typically do. Teachers range in the amount of time they allocate to projects, but most multimedia projects take up to a month or more to complete. Instead, students were given roughly one period and a half to read the documents and complete a draft brochure. Time constraints for teachers necessitated the creation of a shorter task; teachers in some comparison classrooms especially were reluctant to administer additional time-consuming assessments when their district was already taking 2 to 3 weeks to administer the state-mandated SAT-9 test. Moreover, students were not able to develop final products or receive feedback from the homeless education program staff who had agreed to use the students' ideas in a real brochure.

An additional, perhaps more significant limitation was the absence of multimedia from the task. Given that we were attempting to compare student products from multimedia classrooms with products from classrooms with little or no technology, we believed we needed to design a task fair to both, but one that would allow students to use the design skills they might have developed in their computer-based multimedia work. The design process, however, cannot be assumed to be the same in both instances. Although most students in Multimedia Project classrooms do in fact compose storyboards of their presentations using pencil and paper before they use the computer, ultimately, the use of multimedia technology shapes the design process itself. We believe that in certain respects, Multimedia Project students may have been at somewhat of a disadvantage compared with students in comparison classrooms, in that a key tool in their accustomed design activity was taken away from them. Nonetheless, we believe it is important to be able to show that students in project classrooms have developed skills that transfer to new design settings, regardless of whether or not their usual technology tools are accessible to them.

Designing the Rubric and Scoring the Brochures

Exhibit 2 shows the rubric used to assess student brochure drafts on each of three dimensions: content, audience, and design. Each dimension had five levels, with the highest level being best. The *content* dimension focused on students' effective use of documents to identify problems, solutions, and reasons why their solutions would adequately address the problems faced by homeless students in school. The *audience* dimension focused on students' skill in constructing solutions that their audience—teachers and principals—could address and in appealing to the specific concerns and values of school staff. Finally, the *design* dimension focused on students' effective integration of text, pictures, and graphics into an eye-catching format, where all elements contribute to a coherent message to the audience. This rubric was developed by the research team through revising a student-developed rubric developed as a Multimedia Project collaborative activity. Exhibit 3 describes the way in which students participated in the rubric development process.

To score the brochures with the rubric, four SRI researchers, a project staff representative, and the teacher from the classroom where students had worked with researchers on generating the criteria for a good brochure were invited to participate in a day of reviewing each student brochure from both project and comparison classrooms. All reviews were blind: Judges did not know whether they were examining a brochure from a project or a comparison classroom. Each reviewer was trained in the use of the rubric by anchoring raters' judgments through reviews of sample brochures before actual student brochures from the study were judged. Raters then worked in pairs and came to agreement on scores for each brochure on each of the three dimensions. If raters could not come to agreement, a third rater served as a tie-breaker for scoring a particular brochure.

Our analysis of results focuses on students' scores in each of the three domains for which we developed a scoring rubric: content, audience, and design. In the *content* area, we measured the extent to which students were skillful in using factual and narrative information to back their claims about what were the most difficult problems faced by homeless students in school, proposed solutions to address those problems, and the reasons why students believed their solutions would work. In contrast to actual Multimedia Project activities, time studying the content was quite limited and held constant for purposes of our performance assessment. For this reason, we did not

necessarily expect students of Multimedia Project teachers to do better than their peers in comparison classes on dimension of content. If the project has been successful, however, participating students should have gained skills in selecting from available content to tailor a presentation or argument for a specific *audience.* We gave the highest audience scores to brochures in which students presented solutions that could be implemented by school officials and showed evidence that they were aware of some of the specific concerns and needs of teachers and principals (e.g., limited time to focus on specific problems). Finally, we expected Multimedia Project students to be more skilled in *design* than their comparison classroom counterparts. Although they were not using multimedia software in our performance assessment, their experience with that medium should have produced greater attention to combining graphic and text elements, to creating a unified visual whole, and to using stylistic elements to highlight or show relationships among elements. We expected project students to be able to integrate text, images, and graphics into a coherent, eye-catching whole.

Results

Overall Differences in Scores

In scoring completed by pairs of scorers blind to students' membership in project or classroom groups, there were significant differences between the mean scores given to brochures from Multimedia Project classrooms and scores given to brochures from comparison classrooms (Table 2).

Table 2. Overall Mean Scores on Performance Task: Multimedia Project versus Comparison Students

	Multimedia Project Students (n=79)	*Comparison Students* (n=40)	*t*	*p*
Mean Content*	2.15	1.78	2.33	.021
Mean Audience	2.14	1.65	2.44	.016
Mean Design	1.92	1.63	2.07	.041

* Average score from 1-5 on Rubric (Exhibit 2)

The two groups differed significantly on all three dimensions ($p<.05$) with the Multimedia Project student groups out performing the other students in each case. Thus, the kinds of activities Multimedia Project teachers have instituted with their students appear to have contributed to the development of skills that get applied to similar tasks even when technology is not used.

At the same time it is important to note that scores from both groups of students were low, on average. The vast majority of students from both groups failed to back their ideas about what were the biggest problems homeless students face with evidence from the documents provided. Many students failed to identify solutions that could be implemented by teachers and principals, instead focusing on solutions that could be implemented only by other powerful adults in the community (e.g., providing housing to the students). Third, most students failed to use graphical elements meaningfully throughout their brochure to support their basic message. Unrelated elements, such as peace symbols and moving vans, appeared frequently in student brochures. So although the Multimedia Project does appear to stimulate classroom activities that have some impact on students' skills in communication and design, there is clearly still much room for growth as students gain more experience in designing projects and as teachers become more skilled in drawing students' attention to content mastery, audience, and design.[1]

Differences between Scores from Classrooms with Experienced and Novice Multimedia Teachers

Of the 12 Multimedia Project classrooms that participated in the performance assessment, 8 had teachers with 2 or more years' experience in the project. The remaining four teachers were relative novices, having spent only 1 year in the project. Prior research on the implementation of complex instructional practices supported by technology suggest that teachers believe that they require multiple years working at these practices in order to acquire proficiency (Sandholtz, Ringstaff, & Dwyer, 1997; Means & Olson, 1995). Contrasting the scores of student groups in classes with veteran Multimedia Project teachers with those of first-year Multimedia Project teachers revealed trends (though nonsignificant) in which students with more experienced project teachers produced products that received higher ratings for attention to audience and design skill (see Table 3). These data tend to support observers' and teachers' perceptions that first-

[1] In contrast to this performance assessment, in which student groups worked entirely on their own, students developing multimedia projects as part of their regular classroom activities would typically receive feedback and support from their teachers with respect to these aspects of their products.

year implementations of technology-supported projects are not as effective as those of later years.

Table 3. Mean Scores on the Performance Assessment: Veteran vs. Novice MMP Teachers

	Students in Classrooms with Veteran MMP teachers (n=54)	*Students in Classrooms with Novice MMP teachers* (n=23)	*t*	*p*
Mean Content*	2.22	2.04	0.81	.418
Mean Audience	2.30	1.83	1.70	.093
Mean Design	2.01	1.74	1.48	.411

* Average score from 1-5 on Rubric (Exhibit 2)

Effects of Other Student Characteristics on Performance

Although overall the mean age, grade, and English language status of students across the project and comparison groups were not significantly different (see Table 1), we conducted a series of analyses to assess whether any of these factors was a potential alternative explanation for the difference between Multimedia Project and comparison classroom students' scores on our assessment.

Home language was not a significant predictor of scores in any of the three areas. Student groups with larger proportions of students (50% or more of group members) whose home language was not English performed just as well as those from groups with smaller proportions of these students.

While the average proportion of non-fluent English speakers in a group did not relate significantly to the quality of the group's brochure, the average grade assignment level did make a difference on all three quality dimensions. There were some mixed-grade groups in each sample, so while the overall average grade levels between samples were equivalent, the breakdown by group was not.

Table 4. Mean Scores by Home Language for MMP students

	Groups with 50% or more English Language Learners (n=48)	*Groups with <50% English Language Learners* (n=22)	*t*	*p*
Mean Content*	2.23	2.15	-.349	.728
Mean Audience	2.18	2.10	-.269	.789
Mean Design	1.98	1.86	.574	.568

* Average score from 1-5 on Rubric (Exhibit 2)

To analyze the effects of grade and classroom type (project versus comparison), we performed a multinomial logistic regression, treating student scores as a categorical dependent variable, with classroom type (Multimedia Project versus comparison) as a dichotomous independent variable and average grade level of student as a covariate. The results we report are the significance levels from the chi-square likelihood-ratio tests for the individual effects of a model that includes classroom type and average grade.

When controlled for grade level, the differences between Multimedia Project and comparison classrooms on the content and audience dimensions were no longer statistically significant (Table 5). The difference in quality of design did remain significant even after stastistically removing the contribution of average grade level (p=.045). On the measure of content quality, the advantage of MMP classrooms with grade controlled statistically approached but did not attain significance (p=.126); the advantage for audience was not significant (p=.351).

We then repeated this analysis, contrasting the students of veteran Multimedia Project teachers with those of teachers in the comparison classrooms (Table 6). When grade level is used as a covariate and veteran teachers' and comparison teachers' students are contrasted, the difference in design quality remained significant (p=.047) with Multimedia Project students outscoring comparison students, but Multimedia Project students also outscored comparison students in the content domain (p=.033). Differences between the groups in scores on the audience remained statistically nonsignificant (p=.186).

Table 5. Multimedia Project vs. Comparison Students' Scores: Multinomial Logistic Regression Results

Dimension	Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
CONTENT	Intercept*	84.134	.000	0	
	Average Grade	102.571	18.44	3	<.0005
	Home Language	89.864	5.73	3	.126
	GROUP (MMP vs. Comparison)	89.760	5.63	3	.131
AUDIENCE	Intercept**	93.467	.000	0	
	Average Grade	111.302	17.84	4	.001
	Home Language	96.54	3.07	4	.546
	GROUP (MMP vs. Comparison)	97.82	4.35	4	.360
DESIGN	Intercept***	91.838	.000	0	
	Average Grade	78.419	8.27	3	.041
	Home Language	74.293	4.15	3	.246
	GROUP (MMP vs. Comparison)	78.207	8.06	3	.045

*Model Fit: Chi-Square = 26.61 (df=9, p=.001) **Model Fit: Chi-Square = 26.12 (df=12, p=.010)
***Model Fit: Chi-Square = 21.69 (df=9, p=.010)

Table 6. Veteran Multimedia Project vs. Comparison Students' Scores: Multinomial Logistic Regression Results

Dimension	Effect	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
CONTENT	Intercept*	67.285	.000	0	
	Average Grade	98.807	31.52	3	<.0005
	Home Language	76.554	9.27	3	.027
	GROUP (MMP vs. Comparison)	75.997	8.71	3	.033
AUDIENCE	Intercept**	84.118	.000	0	
	Average Grade	111.883	27.77	4	<.0005
	Home Language	86.287	2.169	4	.705
	GROUP (MMP vs. Comparison)	90.303	6.85	4	.186
DESIGN	Intercept***	65.087	.000	0	
	Average Grade	77.793	12.71	3	.005
	Home Language	67.342	2.26	3	.521
	GROUP (MMP vs. Comparison)	73.046	7.96	3	.047

*Model Fit: Chi-Square = 42.68 (df=9, p=<.0005) **Model Fit: Chi-Square = 35.77 (df=12. p=<.0005)
***Model Fit: Chi-Square = 23.59 (df=9, p=.005)

Reflection Survey

Consistent with what observers found in the Multimedia Project classrooms, individual students in these same classrooms who completed the reflection survey reported using technology more often than their peers in comparison classrooms. There were no significant differences between the classrooms in students' self-reported prior knowledge of homelessness or in their experience with designing brochures. In addition, both groups reported similar levels of interest in the task and had equivalent perceptions of the task's difficulty (Table 7).

Table 7. Selected Results from Reflection Survey: Multimedia Project Students vs. Comparison Students

	MMP Students (307≤n≤314)	*Comparison Students* (130≤n≤133)	*t*	*p*
Prior study of homelessness	0.17	0.26	-2.13	.034
Prior knowledge of homelessness	2.06	2.07	-.178	.859
Prior experience in design	1.88	1.89	-.104	.917
Average tech use in class	3.84	2.92	7.57	<.0005
Interest in task	3.34	3.19	1.16	.247
Perceived difficulty level	2.61	2.74	-1.27	.206
Success in finishing in time	0.64	0.50	2.79	.006
Average number of documents used	2.66	2.33	2.18	.030

Differences between the two groups were limited to a few areas identified in the reflection survey. First, students on the whole in the comparison group said they had studied homelessness to a greater extent than students in the Multimedia Project. This prior study, if accurately reported, could have been expected to turn into an advantage for the comparison students on the content dimension of the rubric, but it does not appear to have done so. Second, students in the Multimedia Project reported they were more likely to finish the task in the time allotted. Given the short period of time students had to read the documents and complete the brochure, finishing the whole task in one hour required students to be efficient in collaborating with other members of their groups. The

Multimedia Project students also reporting using more of the documents provided in formulating the content for their brochures than comparison students used.

Results comparing students from classrooms with more veteran Multimedia Project teachers with students in comparison classrooms yielded similar results. There were only two areas where students in Multimedia Project classrooms with veteran teachers differed from comparison students: in the self-reported uses of technology ($t=4.68$, $df=85$, $p<.0005$) and in their skill in finishing on time (85% versus 50%, $t=2.56$, $df=85$, $p=.012$).

Discussion and Implications

The results of the performance assessment designed to measure the impact of the Multimedia Project on student learning suggest that through the project, students acquire design skills that are critical to many of today's jobs in the new economy. Students in the Multimedia Project were better able to create complex products that exhibited their skill in integrating textual and graphical elements into a coherent whole. While few students scored high on the rubric in the design area, significant differences were found between project and comparison students that remained even after age differences were considered. When student group scores on the assessment from veteran Multimedia Project teachers' classrooms are compared to scores of groups in the comparison classrooms, the Multimedia Project students outscore comparison students on both content and design. When specific student characteristics are accounted for in the analysis of differences in scores, these results suggest the possible impact that teachers who are experienced in conducting projects with multimedia technologies can have on students' design skills and their mastery of content.

The performance assessment described in this paper was explicitly developed to match the content of the project-based learning with multimedia model promoted by the Multimedia Project. Evaluation resources were devoted to developing a new assessment precisely because the Multimedia Project's pedagogical model was so radically different from the kinds of performances captured on California's multiple-choice standardized achievement tests. Nevertheless, when we have presented evaluation findings to Silicon Valley business and education groups supporting Challenge 2000, one or more audience members always raises the question, "So what impact did the Multimedia Project have on SAT-9 test scores?" Students in the Multimedia Project classrooms participating in our

study gained 3.20 Normal Curve Equivalents (NCEs) in Total Reading during the school year in which the performance assessment was administered compared to 2.77 for students in comparison classrooms (a nonsignificant difference). For Total Language, the mean gains were similarly equivalent (a mean gain of 1.07 NCEs for Multimedia Project students versus –0.67 NCEs for comparison students). Thus, the communication and design skills acquired in Multimedia Project classrooms appear to be experienced without sacrificing students' progress in basic skills—a concern often cited by teachers afraid to adopt a project approach in the face of pressure to teach to tests that measure breadth, rather than depth, of student learning (see Penuel, Means, Golan, & Korbak, 2000). The lack of a significant difference between the two sets of classrooms on standardized test score gains suggests also that the advantage of Multimedia Project students on performance assessment does not stem from their having teachers who are simply more effective across the board. Rather, the Multimedia Project support for teachers appears to help them stimulate the kinds of learning that are closely related to the project-based learning with multimedia conceptual model.

As evaluators of innovative reform initiatives like the Challenge 2000 Multimedia Project continue to design performance tasks to measure student learning, we will have to pay careful attention to the wider assessment context in which teachers and policy makers live. Student scores on large-scale, high-stakes assessments remain the yardstick by which many initiatives' success is judged, whether or not those assessments are well-aligned with the goals of the initiative or even well-suited to measure the kinds of skills that the initiative seeks to foster. We will need convincing evidence that students are indeed learning skills that are valued by their families and communities, and this will no doubt require that researchers become more skilled in developing alternative assessments with face as well as construct validity and in communicating their research and evaluation findings to diverse audiences.

References

Allen, C., & Pea, R. (1992). *The social construction of genre in multimedia learning environments.* Menlo Park, CA: Institute for Research on Learning.

Baker, E. L., Aschbacher, P. R., Nieme, D., & Sato, E. (1992). *CRESST performance assessment models: Assessing content area explanations.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Baker, E. L., Aschbacher, P., Niemi, D., Sato, E., & Redfield, D. (1992). *Issues in large-scale performance assessment.* Invited minicourse presentation at the annual meeting of the American Educational Research Association/National Council on Measurement in Education, San Francisco.

Baker, E. L. (1998). *Model-based Performance Assessment.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Carver, S. M., Lehrer, R., Connell, T., & Erickson, J. (1992). Learning by hypermedia design: Issues of assessment and implementation. *Educational Psychologist, 27,* 385-404.

Cognition & Technology Group at Vanderbilt. (1997). *The* Jasper *Project: Lessons in curriculum, instruction, assessment, and professional development.* Mahwah, New Jersey: Erlbaum.

Darling-Hammond, L., Ancess, J., & Falk, B. (1995). *Authentic assessment in action: Studies of schools and students at work.* New York: Teachers College Press.

Erickson, J., & Lehrer, R. (1998). The evolution of critical standards as students design hypermedia documents. *Journal of the Learning Sciences, 7,* 351-386.

Flavell, J., et al. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development, 46,* 1-57.

Lehrer, R. (1992). Authors of knowledge: Patterns of hypermedia design. In S. Lajoie & S. Derry (Eds.), *Computers as cognitive tools*. Hillsdale, NJ: Erlbaum.

Means, B., & Olson, K. (1995). *Technology's role in education reform: Findings from a national study of innovating schools.* Menlo Park, CA: SRI International.

Penuel, W. R., & Means, B. (1999, July). *Observing teaching practice in project-based learning using multimedia: a tool for evaluators.* Presented at the Secretary of Education's National Conference on Educational Technology. Washington, D.C.

Penuel, W. R., Means, B., Golan, S., & Korbak, C. (2000). *Challenge 2000 Year 4 Evaluation Report.* Menlo Park: SRI International.

Perrone, V. (Ed.) (1991). *Expanding student assessment.* Alexandria, VA: Association for Supervision and Curriculum Development.

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher, 16,* 13-20.

Sandholtz, J.H., Ringstaff, C., & Dwyer, D.C. (1997). *Teaching with technology: Creating student-centered classrooms.* New York: Teachers College Press.

Spoehr, K. (1992). Using hypermedia to clarify conceptual structures: Illustrations from history and literature. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Sternberg, R. J. (1985). *Beyond IQ.* Cambridge: Cambridge University Press.

Exhibit 1

A Classroom Multimedia Project: American Discrimination

Kathleen Cecchi's middle school studies class participated for the first time in the Multimedia Project last year. Cecchi taught social studies for 27 years before joining the Challenge 2000 Multimedia Project. She had typically used project work at the end of each unit in her curriculum. Cecchi's participation in the Multimedia Project introduced two new dimensions—collaboration between two classrooms of students and multimedia technology—to her classes' project work. Her classes' project, "The Evolution of the American Struggle Against Discrimination" won the Best Overall Secondary Multimedia prize at the 33rd Annual California Student Media and Multimedia Festival in June 1999. It is an impressive project that incorporates sound clips, photographs, video, music, graphics, and animations to convey both factual content from reference materials and personal content.

Students did their research for the project working in cooperative groups within their own classes. They interviewed friends and family and wrote poems and essays expressing their thoughts about the effects of discrimination. When the research was complete, Web pages were composed from that research. Each student had responsibility for a Web page, which he or she presented to the class for an individual grade. Finally, the two classes collaborated to create one Web site for the project.

Exhibit 2 Scoring Rubric for Performance Assessment

	Content	Audience	Design
5	The brochure integrates stories and factual content into an accurate, coherent message about the problems homeless students face and presents credible solutions based on the documents provided. Problems are explained with data as well as stories. Students' claims about the effectiveness of proposed solutions are backed by warrants (i.e., supporting data or logic). Where appropriate, sources of information are identified.	The brochure is likely to appeal directly to the values and concerns of teachers and principals, and there is evidence that students have identified some of those values and concerns. The cover or title identifies the issue as education of homeless students. The brochure lends itself to quick comprehension by busy readers (e.g., use of fully explanatory headings and graphic elements). All of the solutions proposed are viable and could be implemented by teachers and/or principals.	The cover is eye-catching and contributes to the coherence of the brochure. All six surfaces of the brochure are used. The layout is integrated and ties together the content and objects used to illustrate that content. There is an inventiveness in the use and layout of objects in the brochure, such as overlapping images, text, or graphical* enhancements that result in a coherent collage and/or go beyond block style (by spanning more than one page or using creative layout such as putting a title on its side).
4	The brochure describes the biggest problems homeless students face in school, ideas about what can be done to help, and reasons why the solutions will work. Supporting information is used in at least one section to back claims and arguments. In most cases, students use information from documents. All the information included is correct.	Brochure is appropriate for teachers and principals. For the most part, the brochure flows and is easy to read. At least one of the proposed solutions could be implemented by teachers and principals. Adults in powerful institutions could implement any of the solutions. A clear, explicit introductory context is provided for the reader to understand the message.	The cover page includes a title and a related graphic identifying the topic as helping homeless students. Graphics are used throughout. All of the inside panels of the brochure are used and at least one panel on the back is used. Brochure uses stylistic methods to highlight or show relationships between elements.
3	The brochure identifies the problems homeless students face in school and ideas about what can be done to help. One of these sections is well-developed, containing at least 3-4 ideas backed by supporting information, or students also indicate reasons why the solutions will work. Factual errors are absent.	Brochure appears to be targeted to an audience beyond student peers. At least one of the proposed solutions could be implemented by adults in powerful institutions, though not necessarily teachers and/or principals. Enough explicit introductory context is provided for the reader to understand the message.	The cover page includes a title and a graphic related to the title. Graphics are used throughout most of the brochure to illustrate the text. Graphics are only laid out in a block style.** Overall layout has a coherent organization. Graphics directly relate to the core problem of the assignment.
2	The brochure identifies some of the problems homeless students face in school and some ideas about what can be done to help. One or more of these elements is unclear (e.g., failure to clarify the focus on homeless students as opposed to homeless people generally). Little or no concrete information is provided to back claims and arguments.	Brochure appears targeted to an audience beyond student peers, but is not clearly directed toward teachers and principals. Readers themselves may have to construct an introductory context for the brochure to understand the message.	There is a recognizable title and cover graphic for the brochure. Graphics are only on the cover panel of the brochure or lack a connection to the text with which they appear. Graphics relate to the topic of the brochure but lack a direct tie to the text. Brochure is either too cluttered, lacking a coherence of objects, or too sparse.
1	The brochure fails either to describe problems homeless students face or to present solutions to those problems. There are a number of factual errors, or the absence of much information at all. Extraneous facts and details may be present.	Target audience is other students, vague or unclear. No evidence of consideration that audience may have knowledge, values, and concerns different from those of the authors. There are no proposed solutions, or the proposed solutions are addressed to the wrong audience (e.g., other students).	Cover lacks either a recognizable title or a graphic. Placement of objects lacks a coherent organization. There are no images or graphics used in the layout or those that are used are not related to the topic of the brochure. Brochure is either too cluttered, lacking a coherence of objects, or too sparse.

* "graphic" includes pictures, graphs, symbols, and lines used for design

****** Meaning the element layout does not break the Cartesian grid or a directional imperative. An example to break this would to place the title of the brochure on its side.

Exhibit 3

Student Participation in Developing a Rubric for Scoring Brochures

A Multimedia Project classroom that was not part of the assessment study worked with SRI and the Institute for Research on Learning (IRL) to develop a rubric for scoring brochure designs. IRL had been working in this classroom over the course of the semester to help students and teachers develop checklists that students could use to judge the quality of their multimedia projects. Because these students were familiar with developing checklists and knowledgeable already about rubrics and scoring of student-designed products, we included them in the process of generating a checklist to define what makes a "good" brochure within our performance assessment task.

To aid the students, we created a series of short videos that included adult experts critiquing two brochures created by students from another classroom. One expert focused on the content of the brochure, outlining the strengths and weaknesses of two brochures that the students in the classroom also had before them. A second set of experts focused on the design of the brochure, describing the strengths and weaknesses of these same brochures on this dimension. Finally, the class's own teacher described the strengths and weaknesses of the brochure in convincing her as a representative of the target audience for the brochure.

Students made two revisions to their initial checklists as part of the process. After students generated some initial criteria for deciding what would constitute a good brochure, students were given a new brochure to score according to their criteria. They were then asked to revise their checklist to revise the language of criteria that were difficult to use to score or to add dimensions that they felt they had left out. Students were then given either the instructions to the task or the original task checklist and asked to revise their own checklists again, to ensure that all elements had been incorporated into the final checklist. Students then scored a second brochure using a revised checklist.

SRI took the student-generated checklist as a starting point for creating the final rubric for use in scoring the brochures students produced in the performance assessment.