

**Professional Development for Authentic
Pedagogy in the Social Studies:
An Evaluation**



A Collaborative Project:

**LaCrescent-Hokah Public Schools
Minneapolis Public Schools
West Saint Paul-Mendota Heights-Eagan Public Schools**

**The Center for Applied Research in Educational Improvement
The College of Education and Human Development, University
of Minnesota**

Elisabeth Palmer, Evaluation Coordinator

October 1999

**UNIVERSITY
OF MINNESOTA**



**The College of Education
& Human Development**

carei

ACKNOWLEDGEMENTS

The evaluation of the *Authentic Pedagogy in the Social Studies* professional development project and preparation of this report were authorized by the Department of Children, Families and Learning.

Patricia Avery served as the Project Director and Elisabeth Palmer as the Evaluation Coordinator. Vera Kovacovich, Tricia Dieck, and Terry Johnson assisted with data collection. Amy Mayer assisted with data collection, analysis, and report preparation.

We gratefully acknowledge the assistance of the teachers and administrators from the LaCrescent-Hokah, Minneapolis, and West Saint Paul-Mendota Heights-Eagan school districts.

Comments and questions concerning any aspect of this report should be addressed to:

Elisabeth Palmer, Evaluation Coordinator
Center for Applied Research and Educational Improvement
College of Education and Human Development
University of Minnesota
275 Peik Hall, 159 Pillsbury Drive S.E.
Minneapolis, MN 55455-0208
Phone: 612-624-6359 Fax: 612-625-3086
E-mail: palme015@tc.umn.edu

CONTENTS

EXECUTIVE SUMMARY	i
INTRODUCTION	1
PROJECT FEATURES.....	3
PARTICIPANT CHARACTERISTICS	8
IMPACT ON TEACHING AND LEARNING.....	13
CONCLUSIONS	30
APPENDIX A: EVALUATION DESIGN AND WORK PLAN	40
APPENDIX B: AUTHENTIC INSTRUCTION SCORING RUBRIC	45
APPENDIX C: AUTHENTIC ASSESSMENT SCORING RUBRIC.....	49
APPENDIX D: AUTHENTIC PERFORMANCE SCORING RUBRIC.....	53

EXECUTIVE SUMMARY

BACKGROUND

Throughout the state of Minnesota, teachers have been attending brief staff development workshops to familiarize themselves with the Minnesota Profile of Learning and to help them use model performance packages in their classrooms. The movement toward more authentic standards-based performance assessment, however, requires a significant shift in thinking about teaching and learning--changes in assessment require corresponding changes in instruction. With funding from the Department of Children, Families and Learning, the *Authentic Pedagogy in the Social Studies* (APSS) project provided secondary social studies teachers with sustained professional development to assist them in their implementation of the graduation standards throughout the 1998-99 academic year.

PROJECT FEATURES

Research-Based

The APSS project is based primarily on the work of Dr. Fred Newmann and his associates at the University of Wisconsin. Newmann et al. offer an integrated conception of authentic achievement that encompasses *assessment tasks, instruction, and student work*. All three elements—the task, instruction and student work—must focus on the core principles of authenticity: the construction of knowledge, disciplined inquiry, and value beyond the classroom.

In a national study by Newmann, authentic pedagogy (assessment task and instruction) was a strong predictor of student performance.¹ A study by Avery of teachers in an area urban high school's social studies department supports Newmann's research.² Both studies found that *the higher the level of authentic instruction* (i.e., focus on higher level thinking, disciplined inquiry, substantive conversation and connections beyond the classroom), *the higher the level of student performance for all students*, regardless of achievement level or demographic characteristics.

The Professional Development Seminars

During the 1998-99 academic year, monthly seminars were held with teachers from middle and high school social studies departments in districts committed to long-term changes in pedagogy. These seminars focused on developing an in-depth understanding of “authentic pedagogy” and its significance for instruction and

¹ Newmann, F.M., & Associates. (1996). *Authentic Achievement: Restructuring Schools for Intellectual Quality*. San Francisco: Jossey-Bass Publishers.

² Avery, P.G. (1999). Authentic assessment and instruction, *Social Education* 65(6), pp. 366-373.

assessment in relation to the Profile of Learning. They also provided *extended opportunities for teachers to apply and reflect* on the process. This *peer-based* professional development model was successfully piloted the previous year with teachers from one urban high school who attended a seminar offered at the University of Minnesota.

In an effort to reach more teachers working in diverse educational settings, APSS was offered *on-site* in three districts that included *urban, suburban, and rural* schools. *Entire social studies departments* were encouraged to attend the seminars as teams to help institutionalize systems of on-going support for teachers engaging in authentic pedagogy. Dr. Patricia G. Avery, Associate Professor of Social Studies at the University of Minnesota, designed and facilitated seven of the nine seminars held in each of the school districts. Dr. Fred Newmann, Professor of Curriculum and Instruction at the University of Wisconsin-Madison, facilitated the two sessions attended by all three districts.

The purpose of the APSS seminars was to assist teachers in:

- *translating* the theoretical framework that links the Profile of Learning, authentic assessment tasks, and authentic instruction *into practice*;
- *creating* meaningful *assessments* and corresponding *rubrics* that address the Profile of Learning; and
- *evaluating* the Profile of Learning performance packages and teacher-designed assessment tasks, student work, and their own teaching in terms of *authenticity*.

PARTICIPANT CHARACTERISTICS

Districts

Three school districts—one each from urban, suburban, and rural locations—participated in the project. The middle schools and high schools involved were similar in size with the exception of the rural high school, which enrolled about half as many students as the other upper secondary schools. Students in the urban schools were three times more likely than students in suburban schools to be eligible for free or reduced price lunches and seven times more likely than rural students. The same trends were evident in the pattern of minority enrollment.

Teachers

A total of 37 social studies teachers attended the initial seminars of which 34 completed the project. Of those completing the project, 15 taught in urban schools, 13 in suburban, and 6 in rural. Nine of the teachers taught social studies in middle schools; the remaining 25 taught at the high school level. Male and female teachers were equally likely to attend the seminars. The typical participant, however, was white, 38 years of age, taught high school, and had been teaching an average of 10 years. Just under half of the teachers had advanced degrees, most of them from suburban and rural school districts.

Students

Teachers participating in the APSS project were asked to gather samples of student work from two different classes during the school year, once in the fall and once in the spring, as part of the evaluation. The typical student in the sample was white, born in the United States, about 16 years old, attended high school, and received mostly A's and B's. About half of the students said they enjoyed learning social studies. Boys and girls were equally represented in the samples. There were no significant differences between the fall and spring samples of students on key characteristics. Students' ethnicity and poverty status (i.e., eligibility for free or reduced price lunches) reflected the district trends.

IMPACT ON TEACHING AND LEARNING

Impact on Instruction and Assessment

Authentic pedagogy—instruction and assessment—requires students to construct knowledge, engage in disciplined inquiry, and make connections to the world beyond the classroom and the school. In looking at changes over time, we saw statistically *significant improvements in the authenticity of assessment tasks* that teachers created but no real change in the extent to which they engaged in authentic instruction. In particular, *teachers showed great improvements in their ability to create tasks that required students to organize information, consider alternatives, and express themselves through elaborated written communication.*

Overall, these findings were not surprising given that the APSS seminars were structured to meet the needs of the participants who were heavily involved in the early implementation of the Minnesota Profile of Learning which, at this stage involved selecting, modifying, and creating assessment tasks.

In comparing the assessment tasks created and submitted by project teachers at the end of the project to the “typical tasks” submitted by non-participating teachers, we again saw statistically significant differences in the level of task. The tasks created by project teachers represented much greater authenticity overall.

In sum, teachers who participated in the APSS Project significantly enhanced their ability to create more tasks that require students to construct knowledge, engage in disciplined inquiry, and make connections to the world beyond the classroom and the school.

Impact on Learning

Programs designed to improve instruction and assessment are ultimately seeking improvements in student achievement. The concept of authenticity used in the APSS Project highlights the “intellectual quality” of a student’s performance. Again, a more authentic student performance would demonstrate students’ ability to construct knowledge, engage in disciplined inquiry, and understand that the value of the problem or task extends beyond school.

In evaluating the project's impact on student performance, we looked at the effects of instruction, assessment, and student characteristics on students' performance using data from each of the project teachers' classrooms.

Given that the students in the fall and spring samples did not differ significantly in key characteristics, attitudes, or prior school performance, we first examined differences in their performance on the tasks assigned by their teachers. *The samples of student work gathered from the spring sample—near the conclusion of the project—demonstrated a higher level of authenticity or “intellectual quality.”* In particular, these students showed a *deeper level of analysis and greater understanding of disciplinary concepts.*

Next we looked at the impact of authentic instruction, authentic assessment, students' prior performance, interest in social studies, and key demographics on their performance on this particular task. The findings indicate that, of these factors, *the greatest predictors of more authentic student performance, for both the fall and spring samples, were the use of authentic assessment tasks and students' prior grades in social studies.* Students who were given more authentic tasks and those who typically did well in social studies turned in assignments that demonstrated a greater “intellectual quality.” In other words, *the work of these students was more likely to show evidence of knowledge being constructed, a better understanding of and articulation of the discipline, and a clearer understanding of the connections between the assignment and the larger world.*

Students' ethnicity, gender, socioeconomic status, attitude towards social studies, and overall grades in school had no real influence on their performance suggesting that, for the most part, all students are likely to benefit equally from the use of more authentic assessment tasks.

In sum, we can conclude that APSS Project did indeed have a positive impact on students' performance by way of improvements in the quality of assessment tasks: more authentic assessment tasks were associated with a greater “intellectual quality” in students' performance and for all students.

CONCLUSIONS

Making it Meaningful

Teachers commented that the *overarching framework* of authentic pedagogy and authentic achievement not only encompassed both state and national standards and thus helped them think about what they are asking their students to know and do, but it also helped them to see connections between their pedagogy and the intellectual quality of students' work. Being able to walk away with a solid *set of tools and a common language* for assessing the quality of their instruction, assessment tasks, and student work, and for developing quality tasks was also identified as an important feature of the project.

Teachers also liked the collegial nature of the seminars that provided them with numerous *opportunities to meet regularly with colleagues* from their own social studies department, as well as with colleagues from other grade levels and districts. Such meetings, which are typically quite rare given teachers' busy schedules, allowed teachers to give and receive feedback on their instruction and assessment and to reflect on how curriculum and instruction might be aligned across subjects and grade levels. Regular meetings also facilitated the development of ongoing relationships to support teaching that comes from having a shared language and a common pedagogical framework.

Seeing the Results

Not only was the APSS project effective in giving teachers the tools they need to create assessment tasks that ultimately raise the intellectual quality of student work, but, more generally, *it raised teachers' awareness of criteria for assessing the quality of their own practices*. As a result, both new and veteran teachers are *making conscious decisions to incorporate different types of assessments into their instruction on a more regular basis*. Many teachers are also *setting goals for changing their instruction and assessment practices to better align with the standards for authentic pedagogy*. The APSS seminars were also effective in improving teaching and learning in part because they kept the *focus on students*, what we ask of them, and how we can best meet their needs. The seminars also showed teachers that they could *improve their own practice a little at a time* and still make an impact on their students.

Why it Worked and What it Takes

Twenty-five of the thirty-one teachers interviewed said the APSS project was better at meeting their needs than other professional development workshops and cited these reasons: (1) it fit closely with the graduation standards, (2) was responsive to the needs of participants, (3) had an immediate practical application, and (4) provided sustained effort and support.

Teachers and administrators also agreed that it takes the following to achieve the benefits provided by this model of professional development: (1) a clear vision for implementing the graduation standards; (2) a commitment from individual teachers, departments, schools, and the district; (3) flexible staffing to accommodate release time for an entire department; (4) educational standards resource books and existing materials and resources; and (5) time to develop assessment tasks, observe colleagues' classes, and meet with colleagues.

Sustaining It

When asked if they would like another department from their school to participate in a professional development model like the APSS project, all of the administrators were in agreement: yes.

The APSS professional development model clearly demystified the process of implementing state and national standards for many teachers. It also showed the

positive results that may be gained by empowering individual teachers, whole departments, and, eventually, entire schools.

The APSS project reflects a philosophy that higher quality standards can be best achieved through giving the classroom teacher tools for both assessing the quality of tasks and assessing student achievement. Rather than having some kind of standardized process that is brought to the classroom teacher, this actually shows a great deal of confidence in the teacher's ability to do this. And yet it doesn't take away from the standardization of tasks and the standardization of quality.

Something like the seminars almost ought to be institutionalized, so that every few years every department could have a chance to rejuvenate itself in this way. It did a lot to strengthen our department.



INTRODUCTION

In an effort to improve teaching and learning and to assist teachers in implementing the graduation standards, the Department of Children, Families and Learning sponsored several projects during the 1998-99 school year. One such project, *Authentic Pedagogy in the Social Studies (APSS)*, provided social studies teachers with sustained professional development by way of daylong seminars held each month throughout the school year. Teachers from one middle and one high school in each of three school districts attended the seminars along with their colleagues in the social studies department. Dr. Patricia Avery, Associate Professor of Social Studies at the University of Minnesota, designed and coordinated the sessions.

The purpose of the seminars was to assist teachers in:

- *translating* the theoretical framework that links the Profile of Learning, authentic assessment tasks, and authentic instruction *into practice*;
- *creating* meaningful *assessments* and corresponding *rubrics* that address the Profile of Learning; and
- *evaluating* the Profile of Learning performance packages and teacher-designed assessment tasks, student work, and their own teaching in terms of *authenticity*.

The evaluation of the APSS professional development project was designed to address the following questions:

- What were the nature of the program activities and the extent of participation?
- How effective was the project in promoting authentic pedagogy in a variety of educational contexts?
- What were the effects of authentic pedagogy on learning and achievement of the graduation standards for students of different backgrounds?

To answer these questions, both formative and summative data were collected. Formative data collection involved participant's evaluations of individual seminars. Summative data was gathered during two major phases, one in the fall of 1998 and the other in the spring of 1999. The pre/post data collection included teacher and student surveys, classroom observations, assessment tasks, and student work samples. Other data collection activities included telephone interviews with participating teachers, key school contacts, and administrators; seminar attendance records, and review of project materials (see Appendix A for a complete description of the evaluation).

This report summarizes the results of this evaluation. It is organized by the major topics generated from the above guiding questions:

- *Project Features* describes the model of professional development;

- *Participant Characteristics* describes the school districts, teachers, and students involved in the project;
- *Impact on Teaching and Learning* examines the short- and long-term impact of the seminars on teaching, learning, and implementation of the graduation standards;
- *Conclusions* draws upon the words of participants to present conclusions about the effectiveness of this particular model of professional development;
- *Appendix A* describes the evaluation design and work plan; and
- *Appendices B-D* contain copies of the criteria used to score the authenticity of classroom instruction, assessment tasks, and student performance.



PROJECT FEATURES

This section describes key features of the *Authentic Pedagogy in the Social Studies* (APSS) professional development project. Information was drawn from program documents.³

PROJECT FEATURES

In an effort to improve teaching and learning and to assist teachers in implementing the graduation standards, the Department of Children, Families and Learning sponsored several projects during the 1998-99 school year. The APSS Project provided secondary social studies teachers with sustained professional development throughout the academic year.

Rationale

Throughout the state of Minnesota, teachers have been attending brief staff development workshops to familiarize themselves with the Minnesota Profile of Learning and to help them use model performance packages in their classrooms. The movement toward more authentic standards-based performance assessment, however, requires a significant shift in thinking about teaching and learning; changes in assessment require corresponding changes in instruction. For this more integrated perspective of pedagogy, in-depth, sustained professional development models are critical. Using the concept of authentic pedagogy, standards from the Minnesota Profile of Learning, and best practices in professional development as a foundation, Dr. Avery developed a model for enhancing teachers' understanding and skills as they relate to authentic assessment *and* instruction. This professional development model also provided teachers with sustained opportunities to practice and reflect on their work over the course of the school year.

Research-Based

³ The source document for this section was the grant proposal prepared by the Principal Investigator, Patricia G. Avery, and the Evaluation Coordinator, Elisabeth Palmer.

The APSS Project is grounded in research on instruction, assessment and student achievement. Leading proponents of authentic, performance-based assessments contend that traditional assessments often focus on students' recall of decontextualized, isolated bits of information. Authentic, performance-based assessments, on the other hand, are characterized by demonstrations of achievement on challenging, "real-world" tasks. Students are required to "do" history or to "do" science in realistic simulations or actual use. Such tasks are multifaceted and non-routine, and require students to engage in problem solving and critical thinking similar to that of a professional or citizen⁴.

APSS is based primarily on the work of Fred Newmann and his associates at the University of Wisconsin and, as such, reflects a strong research base and a cohesive framework for thinking about the nature of "authenticity." Newmann et al. offer an integrated conception of authentic achievement that encompasses *assessment tasks, instruction, and student work*. All three elements—the task, instruction and student work—must focus on the core principles of authenticity: the construction of knowledge, disciplined inquiry, and value beyond the classroom.

In Newmann's national study, authentic pedagogy (assessment task and instruction) was a strong predictor of student performance.⁵ In addition, all students, regardless of achievement level or demographic characteristics, benefited from authentic pedagogy. A study by Avery of teachers in one urban high school's social studies department supports Newmann's research.⁶ When six U.S. history teachers used the same authentic assessment task, the level of authentic instruction demonstrated by the teachers accounted for over a third of the difference in students' performances. In other words, *the higher the level of authentic instruction* (i.e., focus on higher level thinking, disciplined inquiry, substantive conversation and connections beyond the classroom), *the higher the level of student performance*.

The Professional Development Seminar

During the 1998-99 academic year, monthly seminars were held with teachers from middle and high school social studies departments in districts committed to long-term changes in pedagogy. These seminars focused on developing an in-depth understanding of "authentic pedagogy" and its significance for instruction and assessment in relation to the Profile of Learning. They also provided *extended opportunities for teachers to apply and reflect* on the process. This *peer-based* professional development model was successfully piloted the previous year with

⁴ Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass Publishers.

⁵ Newmann, F.M., & Associates. (1996). *Authentic Achievement: Restructuring Schools for Intellectual Quality*. San Francisco: Jossey-Bass Publishers.

⁶ Avery, P.G. (1999). Authentic assessment and instruction, *Social Education* 65(6), pp. 366-373.

teachers from one urban high school who attended a seminar offered at the University of Minnesota.

In an effort to reach more teachers working in diverse educational settings, APSS was offered *on-site* in three districts that included *urban, suburban, and rural* schools. *Entire social studies departments* were encouraged to attend the seminars as teams to help institutionalize systems of on-going support for teachers engaging in authentic pedagogy. Dr. Patricia G. Avery, Associate Professor of Social Studies at the University of Minnesota, designed and facilitated seven of the nine seminars held in each of the school districts. Dr. Fred Newmann, Professor of Curriculum and Instruction at the University of Wisconsin-Madison, facilitated the two sessions attended by all three districts.

The project’s three goals were to assist teachers in:

1. *translating* the theoretical framework that links the Profile of Learning, authentic assessment tasks, and authentic instruction *into practice*;
2. *creating* meaningful *assessments* and corresponding *rubrics* that address the Profile of Learning; and
3. *evaluating* the Profile of Learning performance packages and teacher-designed assessment tasks, student work, and one’s own teaching in terms of *authenticity*.

The content of the seminars was organized as follows:⁷

Skills & Knowledge		
Session 1 Conceptualizing “Authenticity”	Session 2 Authentic Pedagogy (Assessment and Instruction)	Session 3 The Research on Authentic Pedagogy and Student Achievement (<i>Newmann</i>)
Application & Reflection		
Session 4 Integrating Minnesota’s High Standards and the National Content Standards	Session 5 Analyzing Assessment Tasks, Instruction, and Student Performance	Session 6 Developing Professional Community (<i>Newmann</i>)
Application & Reflection		
Session 7 Assessing Assessment Tasks, Instruction, and Student Performance	Session 8 Assessing Assessment Tasks, Instruction, and Student Performance	Session 9 Analyzing Assessment Tasks

During the first four seminars, teachers learned about Newmann’s concept of authenticity and critiqued its utility as a framework for implementing the graduation standards, became familiar with the Profile of Learning and national content standards documents, and designed rubrics. The remaining seminars gave teachers extended

⁷ For further details on how this program was implemented, see the project handbook entitled *Toward Authentic Assessment AND Instruction: A Framework for Educators*.

opportunities to apply and reflect on their activities, develop authentic assessment tasks aligned with the Profile of Learning standards, analyze their own instruction, and critique student work.

Each of these processes—assessment, instruction, and student performance—were analyzed and critiqued according to the scoring criteria developed by Newmann. For example, teachers developed authentic assessment tasks and scored them according to the criteria for *authentic tasks* (e.g., to what degree does the task resemble one that students have encountered, or are likely to encounter, in life beyond school?). Teachers also videotaped their own instruction and scored it according to the criteria for *authentic instruction* (e.g., to what degree were students engaged in higher order thinking?). Within departments, teachers observed one another's instruction and analyzed it according the same criteria. Teachers also brought in samples of their *students' work* and scored it together (e.g., to what degree does the work reflect a substantial level of analysis?). Throughout the entire process, teachers shared their work with one another and were invited to critique the tasks, instruction, and student work produced by the other districts.

Finally, teachers from all three districts were brought together as a group on two occasions to share their experiences and reflect on their activities. Dr. Fred Newmann, who served as an outside consultant to the project, led these sessions and shared his expertise. Readings for the seminar included relevant texts, such as *Authentic Achievement*⁸ and *Educative Assessment*.⁹

Relationship to the Profile of Learning

The model performance packages developed by the Minnesota Department of Children, Families and Learning are illustrative of the type of authentic assessment tasks advocated by Newmann and Wiggins. They require students to grapple with complex problems or issues similar to those confronted in everyday and professional situations, require exhibitions of deep understanding, and represent a view of learning that is dynamic and constructivist.

Scholars such as Newmann and Wiggins also provide a conceptual framework for understanding and working with the Profile of Learning standards. By engaging teachers in an ongoing and deeply reflective dialogue around the Profile of Learning standards within the context of authentic pedagogy, teachers were able to move far beyond simply knowing how to implement, embed or adapt the standards. Through the teachers' own demonstrations of more authentic instruction and assessment of student knowledge they came to understand the potential power of the standards to enhance teaching and increase student learning. *Rather than viewing the standards*

⁸ Newmann, F.M., & Associates. (1996). *Authentic achievement: Restructuring schools for intellectual quality*. San Francisco: Jossey-Bass Publishers.

⁹ Wiggins, G. (1998). *Educative Assessment: Designing Assessments to Inform and Improve Student Performance*. San Francisco: Jossey-Bass Publishers.

as an “add-on” or an “activity,” teachers came to view the standards as part of a larger vision of authentic learning environments.

Why Social Studies?

Social studies presents special challenges in terms of the implementation of state and national standards, as well as authentic pedagogy. First, unlike other subject areas, social studies *encompasses seven different disciplines*, four of which have issued national standards (history, civics, geography, economics). Teachers wanting to incorporate the national standards to support their work with the Profile of Learning needed time to become familiar with the different standards documents. Second, there is a *lower level of consensus about the goals and content* of social studies than in any other core subject area (Stodolsky & Grossman, 1995). Third, given the scope of material presented in social studies, teachers are particularly likely to be *concerned about being able to “cover” all of the course content* in light of the Profile of Learning and goals of creating more authentic assessment tasks. *In view of these challenges, a successful model of professional development with social studies departments was quite likely to be successful with other subject areas.*

Best Practices in Professional Development

The structure and format of the APSS seminars exemplifies many of the characteristics of high-quality professional development: adequate time for inquiry, reflection and mentoring; a focus on improving student learning; and collaboration among colleagues.¹⁰ By involving entire social studies departments, the project encouraged *collegiality and systems of on-going support* among teachers.¹¹ By involving middle *and* high school teachers, it *increased communication across grade levels* about goals, content and processes.

Effectiveness

Although this model of professional development had been implemented successfully as a pilot program by teachers in one urban high school, the APSS Project included a formal evaluation to document its effectiveness in *changing teachers’ instructional practices in a variety of educational contexts* (i.e., at the middle and high school levels in urban, suburban, and rural settings). It also examined the *effects of authentic pedagogy on the performance of students from different backgrounds* (race, socioeconomic status, and prior achievement). Findings related to the project’s effectiveness are documented in the sections that follow.

¹⁰ National Foundation for the Improvement of Education (NFIE). (1996). *Teachers Take Charge of Their Learning: Transforming Professional Development for Student Success*. Washington, DC: NFIE.

¹¹ Louis, K.S. , Marks, H.M., & Kruse, S. (1996). Teachers’ professional community in restructuring schools. *American Educational Research Journal*, 33 (4), 757-798.

Dissemination for Immediate Application

To move Minnesota teachers toward full implementation of the state graduation standards, effective models of ongoing peer-based professional development must be documented and disseminated. Already, the APSS Project has reached 34 secondary teachers through *direct delivery of professional development*. Two other resources from the project will be made available to the more than 800 educators in school districts and post-secondary institutions throughout the state and over 50 faculty from the Minnesota Association of College Teacher Educators (MACTE):

- *Toward Authentic Assessment AND Instruction: A Framework for Educators*, a handbook that provides step-by-step instructions for implementing this school-based model of teacher professional development, and
- *Authentic Assessment Tasks for Secondary Social Studies Teachers*, a compilation of tasks created by teachers in the project.

In addition, the project staff consulted with each of the sites on their efforts to extend this model to other subject areas within the school by incorporating it into their staff development plan.

PARTICIPANT CHARACTERISTICS

This section describes characteristics of the districts, schools, and teachers participating in the *Authentic Pedagogy in the Social Studies* (APSS) project. Also included are profiles of the students from these teachers' classrooms. Information was drawn from surveys of participating teachers and their students, and the Minnesota Department of Children, Families and Learning's district and school online databases.

SCHOOLS AND DISTRICTS

A key component of the APSS design was the inclusion of a wide variety of students, teachers, and schools to assess how well this approach to professional development worked for diverse populations.

Three school districts – one each from urban, suburban, and rural locations – participated in the project. The middle schools and high schools involved were similar in size with the exception of the rural high school, which enrolled about half as many students as the other upper secondary schools. Students in the urban schools were three times more likely than students in suburban schools to be eligible for free or reduced price lunches and seven times more likely than rural students. The same trends were evident in the pattern of minority enrollment.

Table 1
Demographic characteristics of districts and schools participating in the APSS project, fall 1998-spring 1999.

District/School (location)	% Minority	% eligible for free or reduced lunch	Total K-12 Enrollment
Minneapolis (urban)	70	67	48,797
Folwell Middle	77	76	827
Roosevelt High School	77	70	1,565
West Saint Paul/Mendota-Heights/Eagan (suburban)	16	20	4,915
Heritage Middle	19	25	766
Friendly Hills Middle	15	17	777
Sibley High School	13	13	1,465
LaCrescent-Hokah (rural)	7	10	1,674
LaCrescent Elem.	4	13	665
LaCrescent-Hokah Secondary	2	8	868

Source: Department of Children, Families and Learning (fall 1998 district data).

PARTICIPATING SOCIAL STUDIES TEACHERS

A total of 37 social studies teachers attended the initial seminars of which 34 completed the project. Of those completing the project, 15 taught in urban schools, 13 in suburban, and 6 in rural. Nine of the teachers taught social studies in middle schools; the remaining 25 taught at the high school level.

Male and female teachers were equally likely to attend the seminars. The typical participant, however, was white, 38 years of age, taught high school, and had been teaching an average of 10 years (see Table 2). Just under half of the teachers had advanced degrees, most of them from suburban and rural school districts.

Table 2
Demographic characteristics of social studies teachers participating in the APSS project by location, fall 1998-spring 1999.

	Urban (n=15)	Suburban (n=13)	Rural (n=6)	Total (N=34)
Percent female	46.2	50.0	50.0	48.4
Percent white	69.2	100.0	100.0	87.1
Percent with advanced degree ¹	30.8	50.0	66.7	45.2
Percent Middle School (6-8)	21.4	30.8	33.3	27.3
Percent High School (9-12)	78.6	69.2	66.7	72.7
Average age	38.5	35.4	42.0	38.0
Average years teaching	7.2	11.6	13.2	10.0

Notes: ¹ Advanced degrees include specialist certificate, master's degree, Ph.D. or Ed.D.
Source: AP Teacher Survey (fall 1998).

COMPARISON SITE TEACHERS

A total of 32 secondary social studies teachers from comparable urban, suburban and rural schools provided additional data with which to examine the effectiveness of the APSS seminars. Of these, 7 taught in urban schools, 16 in suburban, and 9 in rural. Five of the teachers taught social studies in middle schools, 18 in high schools, and 9 in schools that served both lower and upper secondary students.

The typical comparison teacher was male, white, 44 years of age, taught high school, and had been teaching an average of 17 years (see Table 3). Overall, 70% of these teachers had advanced degrees, most of them from suburban and rural school districts. Thus, the comparison teachers were, on average, somewhat older and had been teaching a little longer than the teachers participating in the APSS project.

Table 3
Demographic characteristics of social studies teachers at comparison sites
by location, spring 1999.

	Urban (n=7)	Suburban (n=16)	Rural (n=9)	Total (N=32)
Percent female	50.0	25.0	44.4	35.5
Percent white	100.0	100.0	100.0	100.0
Percent with advanced degree ¹	16.7	86.7	77.8	70.0
Percent Middle School (6-8)	14.3	25.0	--	15.6
Percent High School (9-12)	85.7	75.0	--	56.3
Percent Secondary (7-12)	--	--	100.0	28.1
Average age	46.5	43.6	42.8	44.0
Average years teaching	13.5	17.9	17.0	16.7

Notes: ¹ Advanced degrees include specialist certificate, master's degree, Ph.D. or Ed.D.
Source: AP Comparison Teacher Survey (spring 1999).

STUDENTS

Teachers participating in the APSS project were asked to gather samples of student work from two different classes during the school year, once in the fall and once in the spring, as part of the evaluation.

The typical student in the sample was white, born in the United States, about 16 years old, attended high school, and received mostly A's and B's (see Tables 4, 5, and 6). About half of the students said they enjoyed learning social studies. Boys and girls were equally represented in the samples. As Table 6 indicates, there were no significant differences between the fall and spring samples of students on key characteristics. Students' ethnicity and poverty status (i.e., eligibility for free or reduced price lunches) reflected the district trends noted earlier.

Table 4
Demographic characteristics of participating students by location, fall 1998.

	Urban (n=124)	Suburban (n=134)	Rural (n=90)	Total (N=348)
Average age	16.1	15.7	15.5	15.8
	Percent			
Female	43.9	48.0	60.7	49.9
White	29.6	83.7	94.3	67.7
Eligible free/reduced lunch	66.1	10.8	9.8	30.6
Middle School (6-8)	26.8	30.6	34.4	30.3
High School (9-12)	73.2	69.4	65.6	69.7
Enjoy learning social studies most or all of the time	58.5	41.1	50.0	49.6
Typical Grades in School				
A's	20.2	16.4	14.4	17.2
A's and B's	36.3	39.6	54.4	42.2
B's and C's	31.5	30.6	26.7	29.9
Typical Social Studies Grades				
A's	27.5	28.4	21.1	26.2
A's and B's	30.8	32.8	48.9	36.3
B's and C's	25.8	25.4	22.2	24.7

Source: AP Student Survey (fall 1999).

Table 5
Demographic characteristics of participating students, spring 1999.

	Urban (n=82)	Suburban (n=111)	Rural (n=60)	Total (N=253)
Average age	16.5	16.6	16.1	16.5
	Percent			
Female	56.8	43.7	50.0	49.6
White	27.6	81.8	96.5	67.7
Eligible free/reduced lunch	55.3	12.4	8.8	25.7
Middle School (6-8)	17.1	25.5	48.3	28.1
High School (9-12)	82.9	74.5	51.7	71.9
Enjoy learning social studies most or all of the time	61.3	51.0	58.3	56.2
Typical Grades in School				
A's	13.8	7.9	19.0	12.6
A's and B's	41.3	42.6	58.6	46.0
B's and C's	35.0	37.6	15.5	31.4
Typical Grades in Social Studies				
A's	22.7	21.4	31.7	24.4
A's and B's	36.0	41.7	41.7	39.9
B's and C's	30.7	27.2	18.3	9.2

Source: AP Student Survey (spring 1999).

Table 6
Comparison of demographic characteristics of students in the fall 1998 and
spring 1999 populations.

	Fall (n=348)	Spring (n=253)
	Percent	Percent
Female	49.9	49.6
White	67.7	67.7
Born in the U.S.	82.6	84.5
Eligible free/reduced lunch	30.6	25.7
Location		
Urban	35.6	32.4
Suburban	38.5	43.9
Rural	25.9	23.7
Grade Level		
Middle School	30.3	28.1
High School	69.7	71.9
Grades in School		
Mostly A's	17.2	12.6
Mostly A's and B's	42.2	46.0
Mostly B's and C's	29.9	31.4
Mostly C's and D's	9.2	10.0
Mostly D's and F's	1.4	0.0
Grades in Social Studies		
Mostly A's	26.2	24.4
Mostly A's and B's	36.3	39.9
Mostly B's and C's	24.7	26.1
Mostly C's and D's	9.6	9.2
Mostly D's and F's	3.2	0.4
Enjoys learning social studies most/all of the time	49.6	56.2

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$
Source: AP Student Surveys (fall 1998, spring 1999).



IMPACT ON TEACHING AND LEARNING

This section draws upon all of the available data to examine the impact of the *Authentic Pedagogy in the Social Studies* (APSS) professional development project on the professional climate for teachers, teachers' classroom practices and student learning. It begins with a description of the instructional contexts—schools and classrooms—from which the teachers were drawn and in which they would test their skills.

INSTRUCTIONAL CONTEXT

At the beginning of the seminar, teachers were asked to describe the instructional settings in which they were currently teaching. While we often think of the classroom as the place where teachers spend most of their time, the broader context of the school and teachers interactions with students and staff outside of the classroom also affect their performance and satisfaction.¹² Key areas to be addressed in this section therefore include teachers' views on:

- their school's general climate;
- their instructional goals for students;
- the impact they as teachers can have on students' academic success;
- typical classroom activities and assignments;
- the degree of collaboration and sense of professional community among their teaching staff; and
- their familiarity with and use of state and national standards.

The findings presented in this section include breakdowns by the schools' locations—urban, suburban, rural—as well as the instructional context for all project teachers. The reader is cautioned against making comparisons by location given the small and unequal sample sizes (urban=15, suburban=13, rural=6). Rather, these displays are intended as a way to quickly illustrate differences in the populations of teachers who participated in the APSS project.

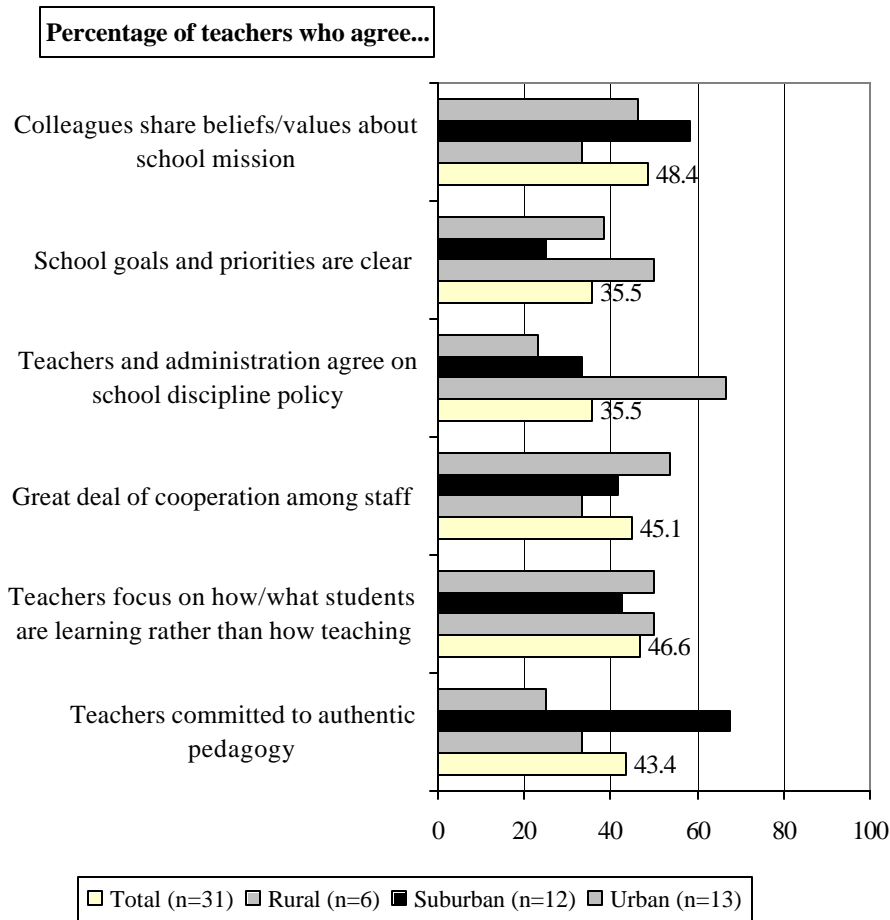
School Climate

Teachers were asked to rate their school's climate on several items (see Figure 1). Overall, about half of the teachers taught in schools where their colleagues shared their beliefs and values of what the central mission of the school should be. About 2 out of 5 teachers also felt that there was a great deal of cooperation among the staff at their school and that the focus at their school was on how and what students were learning rather than how teachers were teaching. Still, just as many teachers said that they and their colleagues were committed to the framework of authentic pedagogy.

¹² Louis, K.S. Marks, H.M., and Kruse, S. (1996). "Teachers' Professional Community in Restructuring Schools," *American Educational Research Journal*, 33(4): 757-798.

Fewer teachers, 1 out of 3, felt that their school’s goals and priorities were clear or that teachers and administrators were in close agreement on the school’s discipline policy.

Figure 1
Teachers' perceptions of school climate by district, fall.



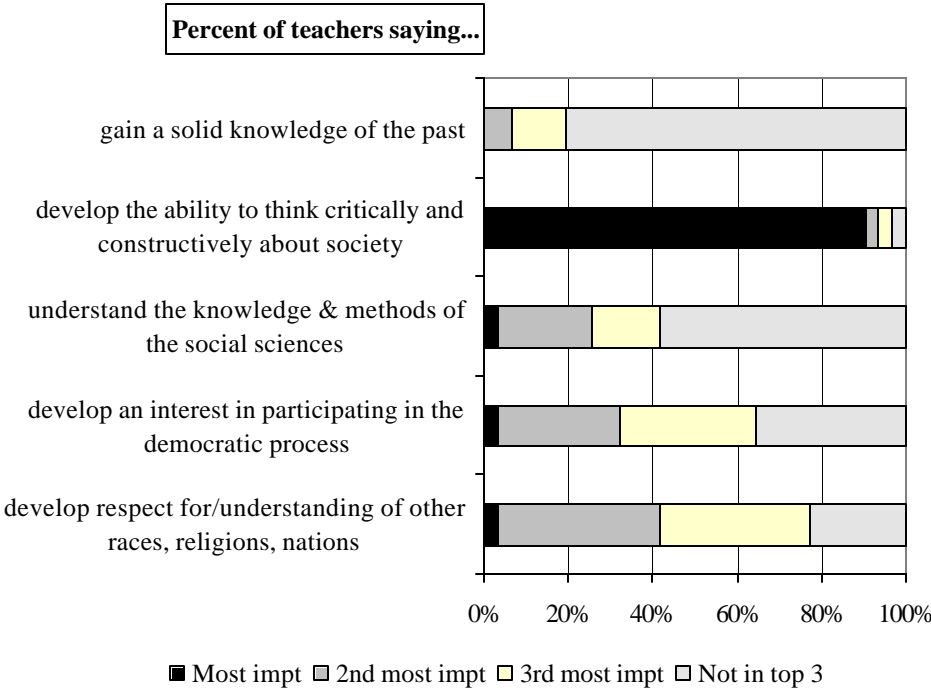
Instruction and Assessment

Instructional Goals

When asked to choose from five common goals teachers have for their students, the social studies teachers in the APSS project overwhelmingly selected “to develop the ability to think critically and constructively about society” as the most important goal (see Figure 2). The second most important goal was for students to develop a respect for and understanding diversity of other races, religions, and nations. This was followed by developing students’ interest in participating in the democratic process and for students to understand the knowledge and methods of the social sciences.

Wanting students to gain a solid knowledge of the past was least likely to be among teachers' top three goals for their students.

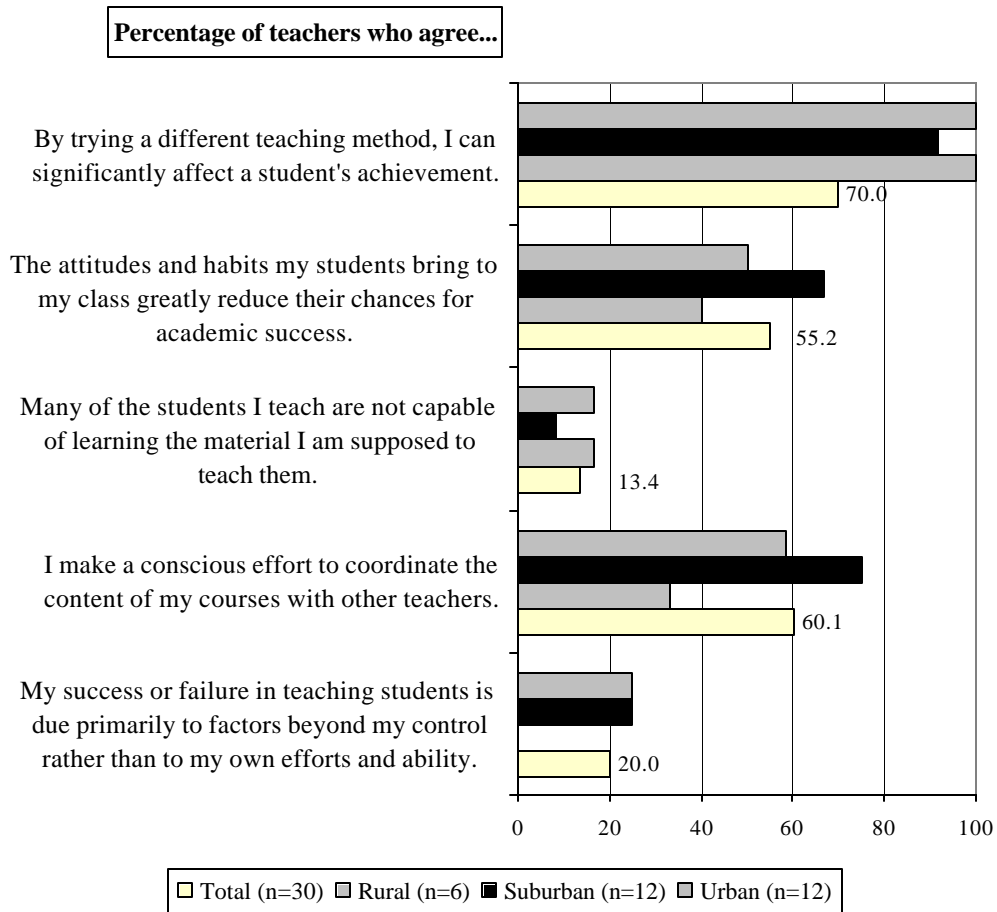
Figure 2
Teachers' initial top three goals for their students, fall.



Locus of Control

Only a few teachers felt that their students were not capable of learning the material being presented to them (see Figure 3). Rather, most of the teachers in the APSS project felt that they could significantly impact students' achievement through their own teaching, and, to that end, 2 out of 3 teachers said they made a conscious effort to coordinate the content of their courses with other teachers. Over half of the teachers also highlighted the importance of students' attitudes and habits in predicting whether a student would be successful academically.

Figure 3
Teachers' perceptions of locus of control by district, fall.



Classroom Activities

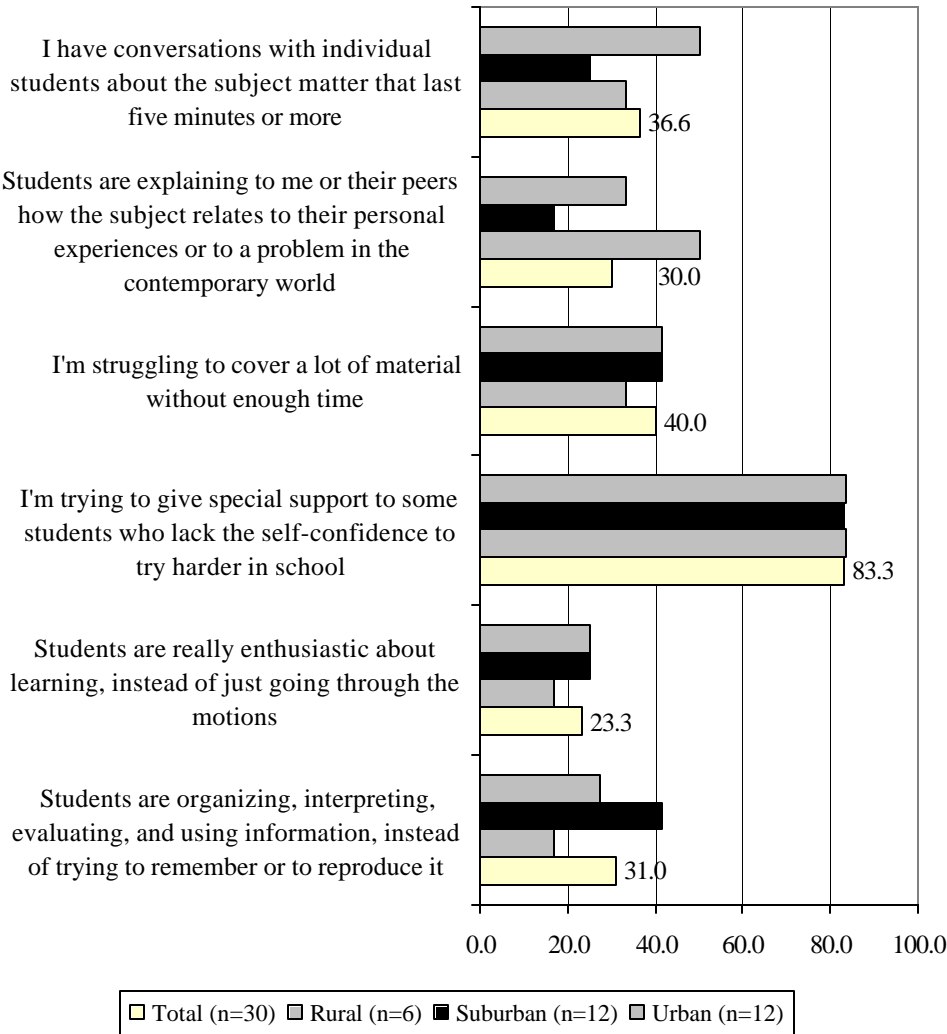
Teachers were asked to describe how they and their students spend their time during a typical class (see Figure 4). Most of the teachers (4 out of 5) said they were frequently or always giving special attention or support to students who lacked the self-confidence to try harder in school. Moreover, only 1 out of 5 teachers felt that their students were truly engaged in the learning process on a regular basis and 2 out of 5 teachers were clearly struggling to cover specified material in the time allotted.

Given these needs, it was not surprising that just 1 out of 3 teachers were engaging in activities that could be considered more authentic in nature when the project began: (1) holding sustained conversations with individual students, (2) having students discuss connections between the class material and the real world, and (3) engaging students in activities that required them to do more than memorize and reproduce information. (Note: Although the frequency of these activities increased significantly in the spring, it was also evident that teachers felt that the students in these classes

were more engaged, that fewer students required special attention, and that they, themselves, were not feeling as pressed to cover a lot of material. For more discussion of how teachers' instructional practices changed as a result of the project, see the "Impact on Instruction" in this section and Conclusions)

Figure 4
Teachers' report of typical class activities by district, fall.

Percentage of teachers saying the following frequently or always occur...

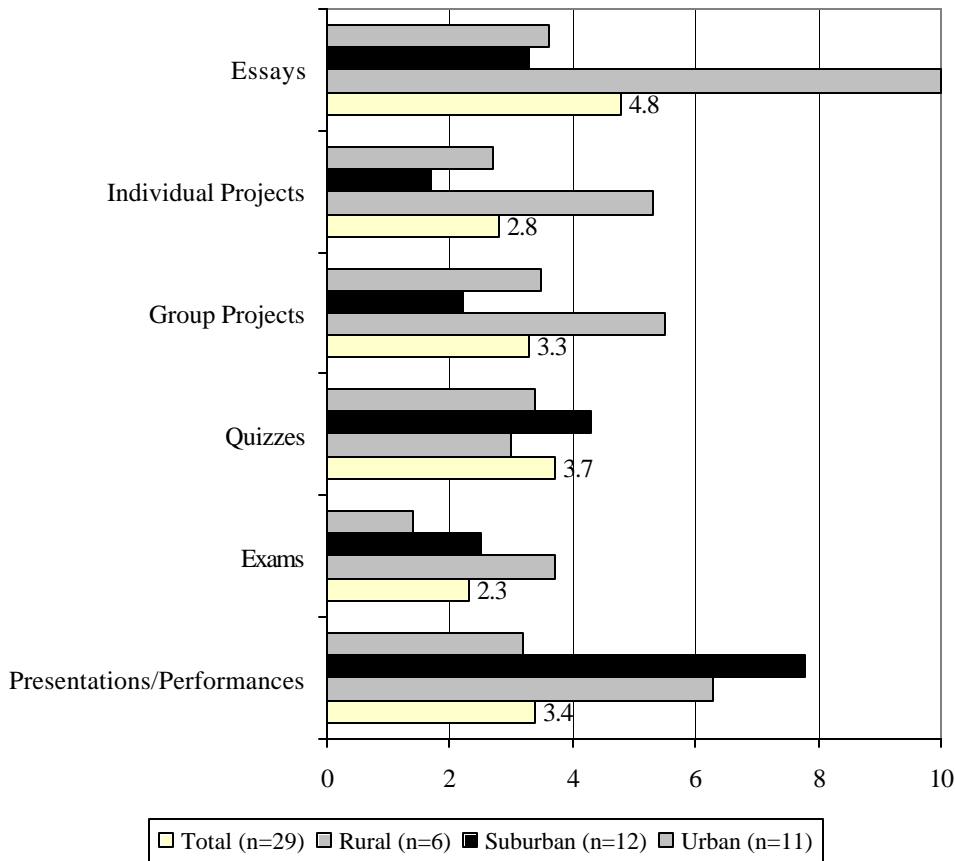


Types of Assessments

To get a sense of the different ways in which the teachers assessed what their students were learning, teachers were asked how many times they usually administered or assigned different kinds of tasks during a report card or grading period (see Figure 5). In general, teachers offered their students a variety of opportunities to demonstrate their knowledge and skills. In a typical class, students would likely be asked to write multiple essays, take quizzes, participate in a few group projects or presentations, as well as, work on a few of their own, and take exams.

Figure 5
Teachers' report of typical class assignments by district, fall.

Ave. # times teachers administer the following assignments during a grading period...

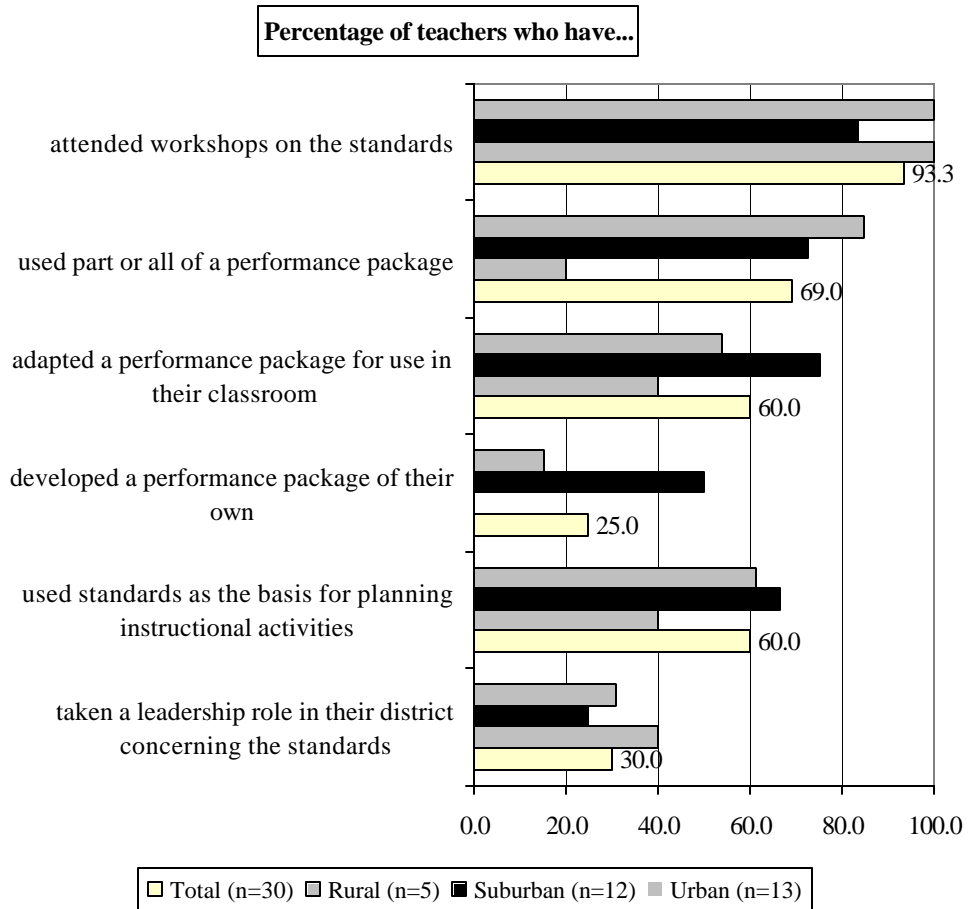


Standards-Based Instruction

Involvement with the MN High Standards

Teachers participating in the APSS seminars varied in their level of involvement with the MN High Standards (see Figure 6). Almost all of the teachers had attended workshops on the standards and 2 out of 3 had used one of the state's performance packages, either in part or in its entirety. Many teachers had also adapted a performance package for use in their classroom; few, however, had developed their own. Of the teachers participating in the seminars, about 1 in 3 had taken a leadership role with regard to standards implementation in their school district.

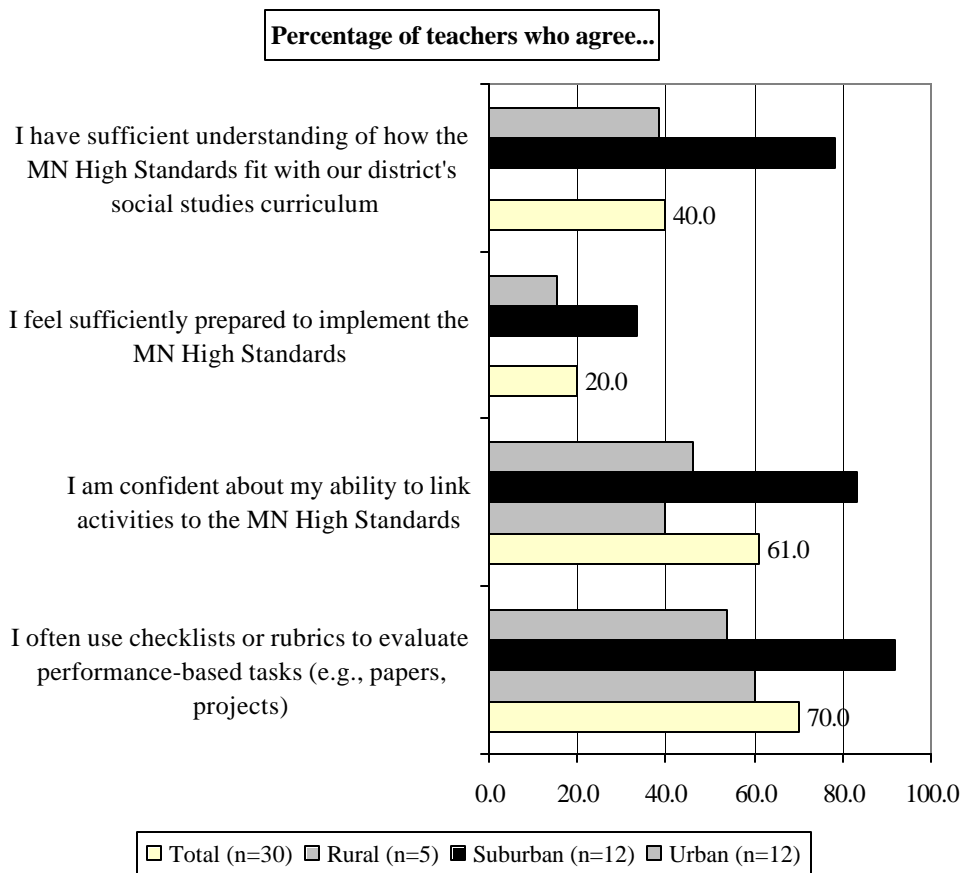
Figure 6
Teachers' involvement with the MN High Standards by district, fall.



Use of the MN High Standards

Almost 2 out of 3 teachers in the project were confident in their ability to link instructional activities to the MN High Standards (see Figure 7) with just as many saying they had indeed done so (see Figure 6). Moreover, the majority of teachers were already using checklists and rubrics to evaluate students' performance in class. Still, despite their facility in using the standards for planning instruction and their familiarity with scoring schemes, only 2 out of 5 teachers felt that they understood how the MN High Standards fit with their district's social studies curriculum and only 1 in 5 felt prepared to implement the standards (see Figure 7).

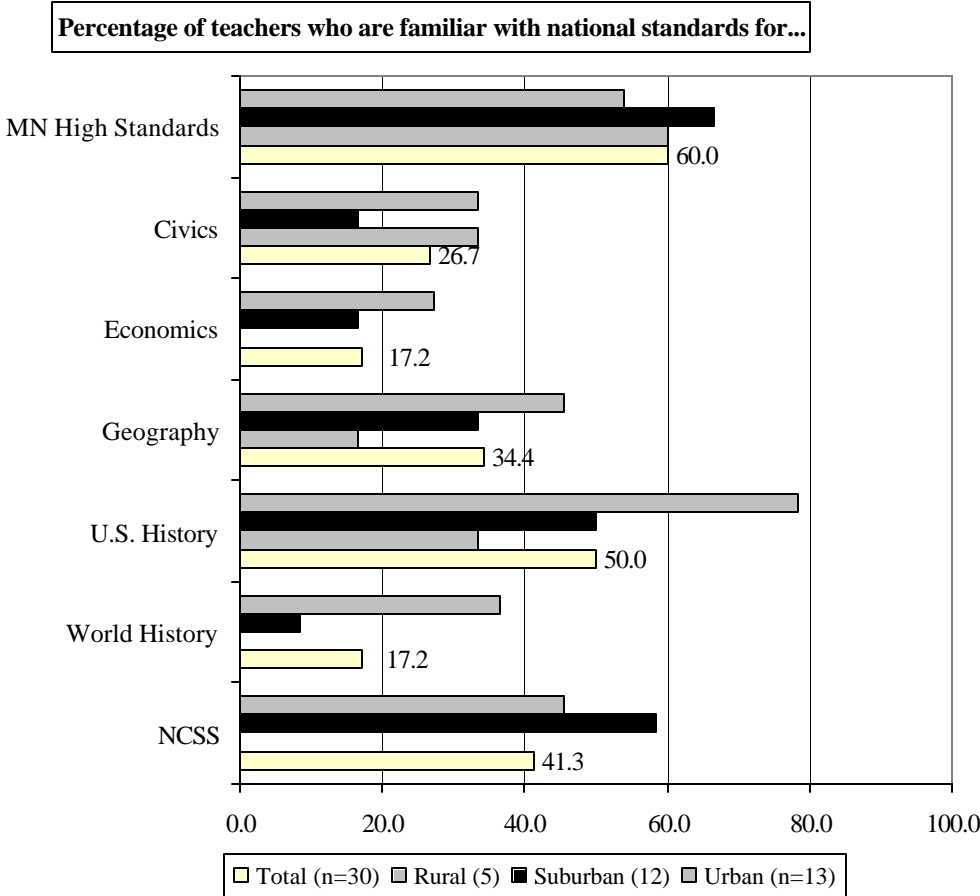
Figure 7
Participating teachers' views of the MN High Standards by district, fall.



Knowledge of State and National Standards

Overall, the teachers in the APSS project felt they were more familiar with the state graduation standards than they were with national social studies standards (see Figure 8). Still, half of the teachers were familiar with the national standards for U.S. History and many were well aware of the standards proposed by the National Council for the Social Studies (NCSS). Teachers were much less likely to say they were familiar with the other national content standards.

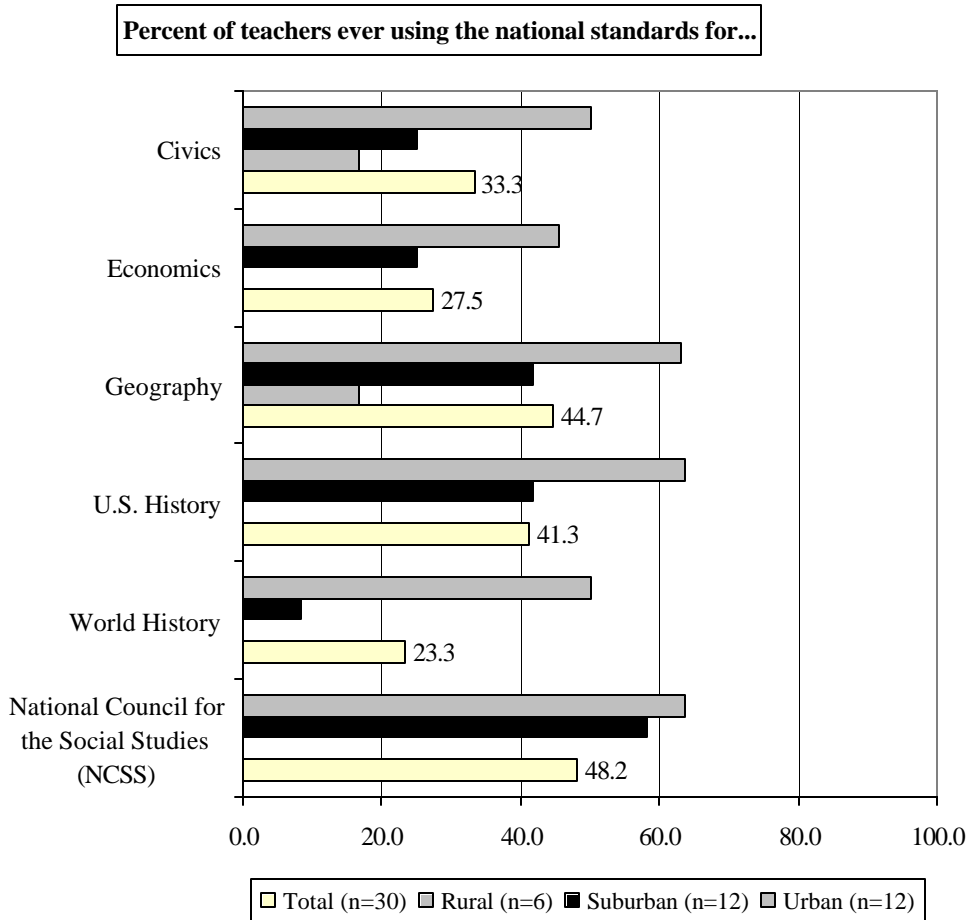
Figure 8
Teachers' familiarity with the State and National Social Studies Standards by district, fall.



Use of the National Social Studies Standards

Although the extent of teachers' familiarity with the national standards for social studies varied, nonetheless, they were attempting to incorporate them into their instruction (see Figure 9). About half of the teachers said they had used the NCSS standards for planning instruction. The national geography and U.S. History standards were also commonly used.

Figure 9
Teachers' use of the National Social Studies Standards by district, fall.



IMPACT ON PEDAGOGY

One goal of the APSS project was to assist teachers in *translating into practice* the framework that links the Minnesota Profile of Learning, authentic assessment tasks, and authentic instruction. This included teaching them how to create meaningful assessment tasks and to evaluate tasks, student work and their own instruction in terms of authenticity. To assess the effectiveness of the project in helping teachers to become more authentic in their instruction and assessment we conducted two types of comparisons. The first examined changes over the course of the project and looked for an increase in the authenticity of teachers' instruction and the assessment tasks. The second compared the authenticity of assessment tasks created and submitted by project teachers to those submitted by non-participating teachers.¹³

Instruction and Assessment¹⁴

Authentic pedagogy—instruction and assessment—requires students to construct knowledge, engage in disciplined inquiry, and make connections to the world beyond the classroom and the school. In looking at changes over time, we saw statistically significant improvements in the authenticity of assessment tasks that teachers created (see Table 8) but no real change in the extent to which they engaged in authentic instruction (see Table 9). In particular, teachers showed great improvements in their ability to create tasks that required students to organize information, consider alternatives, and express themselves through elaborated written communication. The tasks teachers created at the end of the project were also more likely to require that students draw upon their disciplinary knowledge and skills, and to consider an audience beyond the school. The component of authenticity that was generally missing from teachers' tasks was the need for students to work on a problem with connections to the world beyond the classroom. Making it “real” in the classroom, however, was one area of instruction in which the teachers showed some improvement.

¹³ Three comparison districts were selected to reflect those participating in the APSS project. They included secondary social studies teachers from comparable urban, suburban, and rural districts.

¹⁴ See Appendix A for a description of the overall evaluation plan and Appendices B and C for the rubrics used to score instruction and assessment tasks.

Table 8
Change in authenticity of assessment tasks, fall 1998 to spring 1999.

	Fall	Spring	Effect Size
Standard 1: Organization of Information	1.96	2.46	0.75**
Standard 2: Consideration of Alternatives	1.54	2.18	1.02***
Standard 3: Disciplinary Content	1.86	2.18	0.46
Standard 4: Disciplinary Process	1.89	2.07	0.26
Standard 5: Elaborated Written Communication	2.64	3.25	0.75**
Standard 6: Problem Connected to the World Beyond the Classroom	1.75	1.82	0.10
Standard 7: Audience Beyond the School	1.46	1.64	0.20
Total Authentic Task Score	13.11	15.61	0.78**

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$

Scores on Standards 1, 2, 3, 4 and 6 could range from 1-3; scores on Standards 5 and 7 from 1-4; and the total score from 7-23.

Source: Task ratings (fall 1998, spring 1999, $n=28$).

Overall, these findings were not surprising given that the APSS seminars were structured to meet the needs of the participants who were heavily involved in the early implementation of the Minnesota Profile of Learning which, at this stage involved selecting, modifying, and creating assessment tasks. In response to this need, the nine-month APSS project focused more on the development of authentic assessment tasks and less on nuts and bolts of planning and delivering more authentic instruction. In addition, the reliability of ratings on the classroom instruction was relatively low.¹⁵ With a less accurate measure of instruction, the results are more difficult to interpret.

Table 9
Change in authenticity of instruction, fall 1998 to spring 1999.

	Fall	Spring	Effect Size
Standard 1: Higher Order Thinking	2.31	2.22	-0.10
Standard 2: Deep Knowledge	2.29	2.29	0.00
Standard 3: Substantive Conversation	2.25	2.03	-0.26
Standard 4: Connections to the World Beyond the Classroom	2.50	2.81	0.35
Total Authentic Instruction Score	9.35	9.35	0.00

¹⁵ The inter-rater reliability on the total instruction score was estimated as a correlation coefficient of .45. Future analyses will include examination of the extent to which the total instruction score was influenced by individual raters (i.e., analyses will include generalizability to partition out the error variance contributed by a particular rater). Inter-rater reliability on the total task score was estimated as a correlation coefficient of .86, an acceptable level. See Table A2 in Appendix A for further information on score reliability.

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$
 Scores on individual standards could range from 1-5 and on the total score from 4-20.
 Source: Observation Ratings (fall 1998, spring 1999, $n=34$).

In one final analysis of task authenticity, we examined improvements by grade level (middle/high school) and by district location (urban/non-urban). As Table 10 indicates, high school teachers in the APSS project were much more likely to show improvements in the quality of task they produced over time. No differences were noted by location.

Table 10
Change in authenticity of assessment tasks by grade level and location, fall 1998 to spring 1999.

Grade Level			Location		
Middle ($n=8$)	High School ($n=20$)	Effect Size	Urban ($n=15$)	Non- Urban ¹ ($n=13$)	Effect Size
-.63	3.75**	1.50	2.54	2.47	.02

Notes: ¹ Suburban and rural schools were combined.

* $p < .05$, ** $p < .01$, *** $p < .001$

Scores on Standards 1, 2, 3, 4 and 6 could range from 1-3; scores on Standards 5 and 7 from 1-4; and the total score from 7-23.

Source: Task ratings (fall 1998, spring 1999).

In comparing the assessment tasks created and submitted by project teachers at the end of the project to the “typical tasks” submitted by non-participating teachers, we again saw statistically significant differences in the level of task authenticity (see Table 11). The tasks created by project teachers represented much greater authenticity both overall and on every component but one: asking students to address a problem with connections beyond the classroom. In general, both the project and comparison teachers were less likely to create and/or use tasks that required students to address problems or audiences beyond the school.

In sum, teachers who participated in the APSS Project significantly enhanced their ability to create more authentic assessment tasks; that is, tasks that require students to construct knowledge, engage in disciplined inquiry, and make connections to the world beyond the classroom and the school. No real differences were noted in their ability to deliver more authentic instruction. Moreover, these findings clearly reflect the emphasis of the project on tasks rather than instruction.

Table 11
Authenticity of assessment tasks for project and comparison teachers, spring 1999.

	Project Teachers (n=34)	Comparison Teachers (n=32)	Effect Size
Standard 1: Organization of Information	2.50	1.80	1.00**
Standard 2: Consideration of Alternatives	2.20	1.60	0.89**
Standard 3: Disciplinary Content	2.20	1.53	1.03**
Standard 4: Disciplinary Process	2.10	1.53	0.81*
Standard 5: Elaborated Written Communication	3.23	2.07	1.36***
Standard 6: Problem Connected to the World Beyond the Classroom	1.87	1.67	0.25
Standard 7: Audience Beyond the School	1.60	1.07	0.95*
Total Authentic Task Score	15.70	11.27	1.23***

Notes: * $p < .05$, ** $p < .01$, *** $p < .001$

Scores on Standards 1, 2, 3, 4 and 6 could range from 1-3; scores on Standards 5 and 7 from 1-4; and the total score from 7-23.

Source: Task ratings for participating and comparison teachers (spring 1999).

IMPACT ON LEARNING¹⁶

Programs that are designed to improve instruction and assessment are ultimately seeking improvements in student achievement. The concept of authenticity used in the APSS Project highlights the “intellectual quality” of a student’s performance. Again, a more authentic student performance would demonstrate students’ ability to construct knowledge, engage in disciplined inquiry, and understanding that the value of the problem or task extends beyond school.

In evaluating the project’s impact on student performance, we looked at the effects of instruction, assessment, and student characteristics on students’ performance using data from each of the project teachers’ classrooms.

¹⁶ See Appendix A for a description of the overall evaluation plan and Appendix D for the rubric used to score student performance. Inter-rater reliability on the total student performance score was estimated as a correlation coefficient of .63. See Table A2 in Appendix A for further information on score reliability.

Table 12
Differences in the authenticity of student performance for fall and spring populations of students¹, fall 1998-spring 1999.

	Fall (n=348)	Spring (n=253)	Effect Size
Standard 1: Analysis	2.05	2.19	0.19*
Standard 2: Disciplinary Concepts	1.97	2.12	0.20*
Standard 3: Elaborated Written Communication	1.96	2.09	0.15
Total Authentic Performance Score	5.98	6.41	0.21*

Notes: * p < .05, ** p < .01, *** p < .001

¹ Fall and spring samples represent different populations of students.

Scores on individual standards could range from 1-4 and on the total score from 3-12.

Source: Ratings of student work samples (fall 1998, spring 1999).

Given that the students in the fall and spring samples did not differ significantly in key characteristics, attitudes, or prior school performance (see Table 6, page 13), we first examined differences in their performance on the tasks assigned by their teachers. As Table 12 illustrates, the samples of student work gathered from the spring sample—near the conclusion of the project—demonstrated a higher level of authenticity or “intellectual quality.” In particular, these students showed a deeper level of analysis and greater understanding of disciplinary concepts.

Next we looked at the impact of authentic instruction, authentic assessment, students’ prior performance, interest in social studies, and key demographics on their performance on this particular task (see Tables 13 and 14). The findings indicate that, of these factors, the greatest predictors of more authentic student performance, for both the fall and spring samples, were the use of authentic assessment tasks and students’ prior grades in social studies. Students who were given more authentic tasks and those who typically did well in social studies turned in assignments that demonstrated a greater “intellectual quality.” In other words, the work of these students was more likely to show evidence of knowledge being constructed, a better understanding of and articulation of the discipline, and a clearer understanding of the connections between the assignment and the larger world.

Tables 13 and 14 also show the authenticity of instruction as having a significant influence on students’ performance. However, due to questions raised earlier about the accuracy of measuring the authenticity of instruction, the effects of this factor, though statistically significant, are not subject to easy interpretation. For example, one possible explanation for the negative relationship between authentic instruction and students’ performance could be that teachers were spending more time on activities that engage students in higher-order thinking *processes* (e.g., learning how to ask questions and solve problems through sustained interactions with their peers) and less time on mastering the *content* or central ideas of the discipline (see Table 9). Thus, while students might demonstrate excellent thinking skills, their overall performance might be diminished because of a superficial knowledge of the discipline.

Students' ethnicity, gender, socioeconomic status, attitude towards social studies, and overall grades in school had no real influence on their performance suggesting that, for the most part, all students are likely to benefit equally from the use of more authentic assessment tasks. In both the fall and spring samples, the performance of urban students was of a lesser "intellectual quality" than that of their peers in rural and suburban schools. Similarly, high school students in the spring sample demonstrated a greater "intellectual quality" in their work than their counterparts in the middle school.

In sum, we can conclude that APSS Project did indeed have a positive impact on students' performance by way of improvements in the quality of assessment tasks: more authentic assessment tasks were associated with a greater "intellectual quality" in students' performance and for all students.

Table 13
Effects of authentic instruction, authentic assessment, and student characteristics
on the authenticity of student performance, fall 1998.

Variables	B¹	Beta²	Squared Semipartial Correlation³
Constant	1.388	--	.00
Authentic instruction	-.237	-.312***	.06
Authentic assessment	.275	.386***	.10
Grades in school	.288	.133	.01
Grades in social studies	.432	.218**	.02
White	.035	.008	.00
Female	.142	.034	.00
Born in the U.S.	.333	.061	.00
Eligible for free/reduced lunch	-.213	-.048	.00
Engagement ⁴	.022	.021	.00
High School	.498	.102	.01
Urban	-.652	-.151*	.01
Adjusted R-Squared	.22		
N	316		

Notes: ¹ B = unstandardized coefficient; ² Beta = standardized coefficient;

³ The squared semipartial correlation is an indication of the unique contribution of each predictor variable, or the proportion of the total explained variance (R^2) accounted for by an individual variable. The formula is $[(t\text{-value squared})/df \text{ residual}] (1-R^2)$. See Tabachnik & Fidell, 1989.⁴ Engagement is represented by a composite score as four items, each response on a scale of 1 (low) to 5 (high).

(Example: How often would you say you try as hard as you can in this class?)

* $p < .05$, ** $p < .01$, *** $p < .001$

Source: Ratings of student work samples (fall 1998, spring 1999).

Table 14
Effects of authentic instruction, authentic assessment, and student characteristics
on the authenticity of student performance, spring 1998.

Variables	B ¹	Beta ²	Squared Semipartial Correlation ³
Constant	2.114	--	.01
Authentic instruction	-.223	-.273***	.06
Authentic assessment	.190	.301***	.06
Grades in school	.216	.099	.00
Grades in social studies	.224	.114	.01
White	.032	.008	.00
Female	.030	.008	.00
Born in the U.S.	.361	.071	.00
Eligible for free/reduced lunch	.200	.048	.00
Engagement ⁴	.046	.050	.00
High School	1.495	.350***	.09
Urban	-.730	-.179*	.01
Adjusted R-Squared	.33		
N	229		

Notes: ¹ B = unstandardized coefficient; ² Beta = standardized coefficient;

³ squared semipartial correlation is an indication of the unique contribution of each predictor variable, or the proportion of the total explained variance (R^2) accounted for by an individual variable. The formula is $[(t\text{-value squared})/df \text{ residual}] (1-R^2)$. See Tabachnik & Fidell, 1989.

⁴ Engagement is represented by the question "I enjoy learning social studies...all of the time, most of the time, some of the time, rarely, or never."

* $p < .05$, ** $p < .01$, *** $p < .001$

Source: Ratings of student work samples (fall 1998, spring 1999).



CONCLUSIONS

This section draws conclusions about the effectiveness of the *Authentic Pedagogy in the Social Studies* (APSS) professional development model and presents them in the words of the participants.

MAKING IT MEANINGFUL

When asked in what ways the APSS seminars had been meaningful to them, teachers said they were most pleased about the opportunity to put theory into practice.

Teachers commented that the **overarching framework** of authentic pedagogy and authentic achievement not only encompassed both state and national standards and thus helped them think about what they are asking their students to know and do, but it also helped them to see connections between their pedagogy and the intellectual quality of students' work.

The planning piece [was particularly meaningful]...aligning instruction with standards--national and local--and designing assessment tasks and rubrics.

I think the thing that stands out the most is when we went over individual tasks, especially as a department. It made me think not what I was getting from students, but what I was asking of them...that it's so important the way you test them and the individual assignments you give...that the work you will receive will be directly correlated to what you ask of students. Now, every time I design a quiz, test, or individual assignment I remember the tasks that we assessed together.

What was most meaningful to me was the idea of structuring an entire curriculum with [authentic achievement] in mind, the idea that you always have authenticity in instruction or assessment in the back of your mind. It's a paradigm shift, something to strive for.

Teachers also appreciated having **time to practice and reflect** upon what they were learning.

I liked reviewing lessons and looking at the authenticity of materials [my colleagues] use. It was good to actually take pieces from all of our teaching and assess them. It helped me get a feel for how I would grade them, how colleagues would grade them, how other schools would grade them, and look at them in terms of what we were really trying to accomplish.

I taught a class for the first time this semester and was able to use this framework to conceptualize how to build the course. I made some changes I think were successful because of it. I was able to make something pretty

theoretical much more practical and tangible. [Authenticity] can be used as a benchmark, rather than existing at just a theoretical level.

Being able to walk away with a solid **set of tools and a common language** for assessing the quality of their instruction, assessment tasks, and student work, and for developing quality tasks was also identified as an important feature of the project.

I've become a more enriched teacher. Before I took the [seminar] I was doing short-answer, multiple-choice questions. Through the seminars I've learned that [those types of assessments] will not give my students deep knowledge if they're not asking them to think critically. So, I've been having my students write essays where they have to compare and contrast. In doing so, I believe that I've taken my students to a higher level of thinking.

[The framework of authentic pedagogy and authentic achievement] goes way beyond what I have learned in social studies in the past. It gives a specific set of standards to reach for in instruction, the development of tasks, and student achievement. These rubrics make a lot of sense to me and they're an excellent tool for assessing quality.

[The seminars] provided everyone with good tools to walk away with. [They] put everyone on the same foundation and everyone had a starting point with the tools.

Teachers also liked the collegial nature of the seminars that provided them with numerous **opportunities to meet regularly with colleagues** from their own social studies department, as well as, with colleagues from other grade levels and districts. Such meetings, which are typically quite rare given teachers' busy schedules, allowed teachers to give and receive feedback on their instruction and assessment and to reflect on how curriculum and instruction might be aligned across subjects and grade levels. Regular meetings also facilitated the development of ongoing relationships to support teaching that comes from having a shared language and a common pedagogical framework.

We're no longer isolated doing our own little thing in our own little world.

It gave us a common language to use when talking about teaching and assessment. When asked to review a new teacher's lesson, I could use the language of authentic pedagogy in discussing and thinking about how to make some changes in the lesson. In the past, we would have run an idea past a peer, who might have said, "Oh, cool." Now, our peers ask questions that get us to think about an idea in light of [authenticity].

While it's hard to be out of the classroom once a month, it was nice to be together, to talk about where we should be moving as a department, and as a district, as well. When students move from middle to high school, they will see a common language, rubrics, and other things that carry over from 7th,

8th, and 9th grades all the way into high school. I think they'll be more familiar with it and more comfortable with the things we ask them to do.

SEEING THE RESULTS

Not only was the APSS project effective in giving teachers the tools they need to create assessment tasks that ultimately raise the intellectual quality of student work, but, more generally, it **raised teachers' awareness of criteria for assessing the quality of their own practices.**

My teaching has changed drastically. Now, I'm writing authentic tasks. This is how I've always wanted to teach, but didn't know how. Now I continually have the authentic pedagogy standards in mind when planning.

I don't ever hand out an assignment now without considering Newmann's [authentic pedagogy framework] and how I can make the assignment a little bit better. Even though not every assignment is going to require higher-level thinking, I feel that at least I'm aware of what I'm doing. I know why I'm doing something.

I'm more aware of conversation, dialogue, and trying to ask students more questions to make them think, giving them assignments that are more authentic or meaningful. Tests have changed in that I'm trying to make them more authentic, with more writing involved. The rubrics have been beneficial to students because they know ahead of time what they need to do for a certain grade.

I think that in everything I'm doing now, in the back of my mind, is a recognition of whether or not this task is at a high level. Specifically, it's given me tools to develop better rubrics which help me feel more confident that the instruction, the tasks, and the achievement is at a higher level.

As a result, both new and veteran teachers are **making conscious decisions to incorporate different types of assessments into their instruction on a more regular basis.**

I use more projects and essays and assessment tasks that are scattered throughout the course instead of relying on objective testing. I also use [authenticity] as a benchmark. I've really implemented the real world connection and kids seemed to respond; they felt that it was real, useful information and something they might actually implement or reflect on in the future.

This is my first year of teaching, so in a lot of ways, I feel like this has kind of helped me get on a good path right away and not have to break old habits. I've been using a lot of rubrics in my grading. I also feel like the idea of giving students a lot of choices about what they do, or how they choose to demonstrate knowledge learned, that's something that I have really tried to put into my lessons and tasks.

Many teachers are also **setting goals** for changing their instruction and assessment practices **to better align with the standards for authentic pedagogy**.

Now I'm scoring discussions, which I didn't do before. This has greatly improved students' discussions of issues. I've also tried to work on the questions that I ask students this year. When starting a new unit, I try to think about interesting controversies, issues, and topics within that unit and formulate question(s) that are pretty comprehensive, where to answer it students would really have to learn something. I think of several questions that guide them through the unit.

Midway through the year I was so focused on assessment tasks and good rubrics that I was missing some on instruction. So, during the second semester I focused more on cohesive instruction and scaffolding students up to where I wanted, so they could do the tasks I wanted. I became more focused on the day-to-day instruction as it related to cumulative assessment tasks.

When I compare what I did in Civics and World Studies, where I focused on more depth and intellectual work, I made a very conscious decision to go much more in depth next year and focus on more intellectual work instead of cramming facts and concepts in.

I try to have one well-designed authentic instruction or assessment per unit that I'm doing. The whole idea is overwhelming for a whole year, so I've created a starting point to do a couple, or one per unit. I also try to do little things to make things more authentic. I'm trying to create habits that include some of the tools I learned in the seminars.

I've not had time to go back and redo all of my lesson plans...I think this will be more long-term, over a period of years.

The APSS seminars were effective in improving teaching and learning in part because they kept the **focus on students**, what we ask of them, and how we can best meet their needs. The seminars also showed teachers that they could **improve their own practice a little at a time** and still make an impact on their students.

I feel I'm a work in progress. I'm trying harder to make work more practical, something students will actually be doing. I won't say I'm good at it or doing it all the time, but I'm trying to look more at students' needs. I'm less driven by the syllabus and more focused on what is needed for students to improve their achievement.

If anything, I let students show me more, instead of me trying to dig knowledge out of their heads. Instead of giving a final test, I give them information and they pull it together in a way that shows me what they've learned.

I find myself consciously thinking about being more authentic in my teaching style. Through the seminars I realized a lot of what I do is not authentic. Even if I'm not always teaching authentically, I'm at least changing to make an attempt to do it, whether it's once or twice a week. If it's not authentic, I'm at least thinking about it.

WHY IT WORKED

Twenty-five of the thirty-one teachers interviewed said the APSS project was better at meeting their needs than other professional development workshops and cited these reasons:

Close Fit with the Standards

The whole concept of authentic assessment matched well with the graduation standards. It was helpful to work through it until we figured out the correct way.

It fit perfectly. It really demystified the standards. We participated because of our work the previous year with the state's performance packages. We weren't satisfied with the packages and this project helped us to see that we can do the packages or write our own in our classrooms and that they can be useful, valuable. It also helped teachers see that the standards can be integrated, that the standards are about assessment, not just ugly politics. (Teachers at other buildings felt that integration of the standards was an "add-on". At our school, we've been able to handle it more easily). Because of our work with the project, we were able to quickly pound out an assessment for use in all 19th grade World History classes in the district.

The APSS project was a very, very close fit with our department's plans for implementing the graduation standards. This was another way of looking at the standards. The standards emphasis is on meeting the standards-- Newmann's parallels that but is not quite the same. His gives ideas on how to do that, especially on developing assessment tasks.

Meeting the Needs of Participants

In the APSS seminars we were allowed to express what we needed and then it was tailored to meet those needs. If we needed time to work on one aspect or another, we were given that time. Other seminars have a schedule and that's what they are going to do.

It didn't just address one part of being a teacher. It didn't just focus on "here's a way to write assignments...or a way to grade student work...or a way to create better tasks." It was all of it. It was not only instruction, not only the student work, not only designing tasks, but all three.

Practical Application

It had practical application to my everyday job as a teacher. Some [professional development programs] have great theory, but are not practical, or are very difficult to put into use. This is stuff that can actually be used and applied.

It was practical. We could actually assess our own material. Even if we were not assessing our own work, we could still assess student work and compare how we rated it with how others rated it. It was very valuable because we don't have the leisure to look at other people's work and reexamine our own.

Generally, a lot of staff development tends to be infomercials—they give a lot of information and think that's sufficient. They leave the teacher to figure out how to apply it. I don't think until someone's done it they have really processed or owned the information, so if you don't give people the change to start it, it just hangs on the theoretical level.

Sustained Effort and Support

Usually with staff development, you go, you sit through it, you get a bunch of handouts, you come back, put them in a pile, and that's about it. No one ever checks to see if you are making an effort. No one ever cares. With this, with the [required] observation of colleagues and the development of tasks we had to try.

Because it was a continuous, extended project it required more loyalty and commitment. It wasn't just viewed as a two-hour thing where you go in and get the idea. The content was studied in greater depth.

It had a clear focus. Often times, the purpose of staff development in the district is unclear. I have spent a lot of time on the Minnesota graduation standards, but there was always a big gap in what different teachers know so the programs were always very simplistic. In the APSS seminars, I felt that at least 85% of the time was very, very productive.

It was ongoing and other staff members were already involved. One person was not bringing it back to the other teachers and trying to convince them of the merit of something presented at the workshop.

WHAT IT TAKES

A Clear Vision

Key administrators from the participating schools were asked how the APSS project fit into their plans for implementing the Minnesota Graduation Standards. In general, they viewed it as a structured way to give more of their teachers the resources and time they needed to incorporate the standards into their social studies curriculum.

Three or four of our teachers were really good about the graduation standards. This project moved more of our teachers into using them. It fit where we wanted to go: to pick up different kinds of assessments and move out of regurgitation towards more standards-based instruction.

We feel like we have an innovative new curriculum and it was an opportunity for teachers to go to the next level, to get more information about the standards and resources for implementing them into the curriculum.

It advanced our work on the graduation standards and it was good for individual teachers. It was a guided, structured program with some time built in to pursue improvement.

We needed a paradigm shift with regard to how we assess student achievement. This type of program needs to be in social studies and in all departments.

Other Resources

In addition to having a clear vision of where they are going with regard to the graduation standards, school and districts considering a professional development model like the APSS project also need to secure the following resources:

Commitment at all levels

Because of the time and other resources needed to effectively carry out the project, it is not enough to have a commitment to the project at the district-level. It is essential that the individual teachers, departments, and schools demonstrate a clear commitment to the process.

Flexible staffing

Finding enough substitute teachers—the actual people—to provide release time for an entire department, once day a month, for nine months was the greatest obstacle participating schools faced. With enough advance notice, funding for substitutes was generally available through professional development or graduation standards accounts.

Materials and resources

Each department that participated in the APSS project received a set of the national content standards books and a set of articles on authentic pedagogy and authentic achievement. Each teacher received a copy of Newmann's *Authentic Achievement* and *A Guide to Authentic Instruction and Assessment: Visions, Standards and Scoring*. Other than that, as one administrator noted, this model of professional development, unlike many others, did *not* require any special resources to implement (e.g., specific textbooks, materials, etc.) but made use of existing materials.

Time to do the work

As part of the seminars, teachers were asked to write assessment tasks and observe colleagues' classes. The time required for these activities had to be arranged by individual teachers and departments. Teachers also wanted more time to work with their colleagues. To support these developmental and supportive activities, schools may want to restructure their schedule to provide common prep hours for a department. Additional sources of funding may also need to be tapped to support teachers as they develop assessment tasks.

Advice to Other Districts

With this in mind, the administrators had this advice for other schools or districts considering professional development models like this one:

Make sure they understand the intended outcomes of the program (e.g., if they are a teach-and-test district, it won't work as well) and that they will work on one thing over time. Therefore, they will need to be flexible in how they allocate their time and staff.

The key is making it fit with where you're going. That will help them find the time, money, and means of supporting this type of program in the long run.

Have districts closely assess their culture. If some teachers are high flyers and others are less inclined towards innovation (e.g., "don't do that, you'll make us look bad") the school/district will need to consider how to support growth and quell antigrowth attitudes.

It is difficult to accomplish logistically, but all the research says that it's the way to do it--working with whole departments so the teachers have continued support because someone else is around who went to the sessions and who they can ask questions of.

Teachers need to keep an open mind going into the program. It requires a commitment on the part of teachers--they're given the model, but they have to do the work.

[Keep in mind that] one outcome is increased dialogue among department members. Your other departments will become excited.

This just isn't another seminar you go to...it's active, interactive...and that's what learning is all about. That's what we want our students to do.

Teachers' advice to other teachers who might participate in a similar project highlighted the level of commitment it requires and the role the project can play in implementing the graduation standards:

This is not a one-day meeting to just go and sit around, there's work involved. Be prepared to take a risk and work; that is, to leave the session, prepare a task, meet again, and be observed. You must be willing to take criticism and be willing to change how you do things. It's about "revitalizing lessons" and creating something new.

A majority of teachers need to participate and willingly. Teachers have to be willing to put in the time, it's not just a day at a workshop. Our contract pays for 37 hours a week and I work 45 hours as a veteran teacher. But, this workshop energized me so that I got more done more quickly...the flow of ideas, editing my own work, seeing problems. I spent about 2.5 hours a week or 30 minutes a day talking with other teachers as resources.

You should either have the MN graduation standards in place--written and in use--or use this as preparation to do that, but, not at the same time.

SUSTAINING IT

When asked if they would like another department from their school to participate in a professional development model like the APSS project, all of the administrators were in agreement: yes.

I'd like to do it across the entire school, but spread out over different years to stagger. That way we can concentrate on one group and help them move forward. It's short and intense and it has a beginning, an end, and clear outcomes.

...if it's a choice. It offers unity and builds a common dialogue and language over time.

The APSS project was outstanding, impressive, and has long-lasting effects on teachers. I'd like to do this in all of my departments.

It got such a positive response from teachers. They were frustrated when they couldn't go and it wasn't easy for them to leave their classes. But they wanted to be there. They said it was a "valuable experience" and that they "learned a lot." They also felt valued because they had the opportunity and the school really supported them. They liked having different experiences over time and the opportunity to meet with Dr. Newmann.

Whole department is the way to go. It would also be nice to take it to the lower grade levels. We have other “exemplary” professional development programs going on at our school. This research-based model of “coaching” is just heads-and-shoulders above the rest.

The APSS professional development model clearly demystified the process of implementing state and national standards for many teachers. It also showed the positive results that may be gained by empowering individual teachers, whole departments, and, eventually, entire schools.

The APSS project reflects a philosophy that higher quality standards can be best achieved through giving the classroom teacher tools for both assessing the quality of tasks and assessing student achievement. Rather than having some kind of standardized process that is brought to the classroom teacher, this actually shows a great deal of confidence in the teacher’s ability to do this. And yet it doesn’t take away from the standardization of tasks and the standardization of quality.

I think this is the direction we should go. It seems to eliminate a lot of the problems that have developed with the controversy over standards and the performance packages in Minnesota. Instead of developing packages for teachers all over the state to use, I think it’s more appealing to and better for teachers to learn how to evaluate their own tasks. I think the framework of authenticity gives teachers much greater tools for implementing the graduation standards.

This is cutting-edge, working with the whole department instead of a cross section of teachers who normally choose to do things like this. What you’re doing is critical and can really impact student learning.

Would like to see how we can get the whole school on board. Then we’d really have something going!

Our social studies department made a presentation to the school board about their participation in the APSS project and the impact. The board was really impressed. During the 1999-2000 school year our social studies teachers will help start “Phase II” which will include other grade levels and other teams and focus on building professional community. Dr. Newmann will be coming in as well this summer to meet with the new teachers.

Something like the seminars almost ought to be institutionalized, so that every few years every department could have a chance to rejuvenate itself in this way. It did a lot to strengthen our department.



APPENDIX A EVALUATION DESIGN & WORK PLAN

This section describes the questions guiding the evaluation of the *Authentic Pedagogy in the Social Sciences* (APSS) project and the evaluation work and staffing plans.

KEY EVALUATION QUESTIONS

This evaluation was designed to address the following questions:

- What were the nature of the program activities and extent of participation?
- How effective was the project in promoting authentic pedagogy in a variety of educational contexts?
- What were the effects of authentic pedagogy on learning and achievement of the graduation standards for students of different backgrounds?

WORK PLAN

The evaluation plan included both formative and summative data collection (see Table A1 for a summary of the final samples).

Formative Data Collection

Individual Seminar Evaluations. (Ongoing.) Teachers attending the seminars were asked to complete an evaluation of the day's activities for every other session.

Summative Data Collection

Teacher Surveys. (Fall 1998 and Spring 1999.) All participating teachers completed pre/post surveys asking about their school climate, instructional goals and activities, professional activities, experience and familiarity with the Minnesota High Standards and the National Social Studies Content Standards, and basic demographic information.

Classroom Data Collection. (Fall 1998 and Spring 1999.) Data on the authenticity of instruction, assessment tasks, students' performance, and the students themselves was collected once in the fall and again in the spring. Teachers were asked to identify a target class that represented their "typical" social studies class. Given that classes change at the quarter or semester break, it was not feasible to track the impact of the project on the same students over time. Consequently, two different samples of students were included in this evaluation.

- **Assessment Tasks.** (Fall 1998 and Spring 1999.) As part of the classroom data collection, teachers were asked to select an assessment task, assignment, or test they had used since the beginning of the term in the target class. They were further instructed that the task should be one they think gave students an opportunity to demonstrate the students' understanding of the subject matter. The task was to include all directions given to students. Teachers were not allowed to submit a performance task from the Profile of Learning. All tasks were coded using the criteria for authentic assessment (see section on Reliability below and Appendix B).
- **Student Work.** (Fall 1998 and Spring 1999.) Teachers were also asked to provide copies of all student work for the assessment task they handed in. A stratified random sample of student work was selected from all received to include works from every school and every classroom. This sample was coded using the criteria for authentic student performance (see section on reliability below and Appendix D).
- **Student Surveys.** (Fall 1998 and Spring 1999.) All students in the target class completed surveys about the nature of instruction and assessment in that class, their attitudes toward social studies and learning, and their demographic characteristics.
- **Classroom observations.** (Fall 1998 and Spring 1999.) As part of the classroom data collection, trained observers visited the classrooms of participating teachers' twice in the fall and twice in the spring to score the authenticity of instruction (see Reliability section below and Appendix B). The classes in which these observations were conducted did not have to be the same as the target class from which the tasks, student work, and student surveys were drawn, but another "typical" class (i.e., no Advanced Placement or Honors courses). This decision was based on the assumption made by Dr. Newmann that a teacher's ability to teach in an authentic manner will be relatively stable across classes.

Teacher Telephone Interviews. (Spring 1999.) An attempt was made to contact each of the teachers attending the seminar by telephone for a brief interview. The purpose of the interview was to gather information on the impact of the APSS seminars on teachers' instructional and assessment practices, the impact on their department, and their assessment of the quality of these professional development seminars.

Administrator Telephone Interviews. (Spring 1999.) Key administrative contacts from each school and/or district were interviewed by telephone to obtain information on the reasons why they wanted their secondary social studies teachers to participate in the project, how the project fit with their plans for implementing the graduation standards, the resources needed to allow their social studies departments to participate in the project, and their perceptions of the overall quality of and satisfaction with the APSS professional development project.

Key Teacher Contact Telephone Interviews. (Spring 1999.) Key teachers—those identified as the lead teachers--from each of the districts were interviewed by telephone to obtain information on how the project fit with their plans for implementing the graduation standards and the resources needed to allow their social studies departments to participate in the project.

Table A1
Sample status for summative data collection by source.

Source	Fall n	Spring n
Teacher Surveys	34	34
Assessment Tasks	34	34
Student Work	359	261
Student Surveys	736	586
Classroom observations	68	68
Teacher Telephone Interview	--	31
Administrator Telephone Interview	--	6
Key Teacher Contact Telephone Interview	--	3
Source: APSS Project Records, 1999.		

STAFFING

Elisabeth Palmer served as the Evaluation Coordinator and worked closely with the Project Director, Patricia G. Avery, in designing and conducting the evaluation. Vera Kovacovich, Tricia Dieck, and Terry Johnson assisted with data collection. Amy Mayer assisted with data collection, analysis, and report preparation.

RELIABILITY

Classroom observations

Each teacher was observed four times during the year, twice in the fall and twice in the spring. A trained research assistant observed and scored the lesson according to the standards for authentic instruction (see Appendix B). The observer completed descriptive notes to document evidence of ratings.

About 25 percent of the lessons were observed by a second research assistant. Ratings were recorded independently. The overall level of agreement between the two raters on each standard is indicated in Table A2. When ratings on the four standards were added, the overall level of agreement between pairs of raters on the total score was estimated as a correlation coefficient of .45. (Note: Because three trained observers were used to gather this data, further analyses of inter-rater reliability need to be conducted to determine the effects of individual raters.)

Assessment Tasks

Each teacher submitted two typical assessment tasks, one from the fall and one from the spring. Assessment tasks were scored in the spring after both fall and spring tasks had been shuffled and distinguishing marks removed. The principal investigator scored each assessment task according to the standards for authentic assessment (see Appendix C).

About 25 percent of the tasks were also scored by a research assistant. Ratings were recorded independently. The overall level of agreement between the two raters on each standard is indicated in Table A2. When ratings on the seven standards were added, the overall level of agreement between the two raters on the total score was estimated as a correlation coefficient of .86.

Student Performance

Samples of student work on the submitted tasks were obtained in the fall and in the spring. A stratified random sample of student work was selected from all received to include works from every school and every classroom. All student work was scored in the spring after fall and spring tasks had been shuffled and distinguishing marks removed. Each sample of student work was scored by a research assistant according to the standards for authentic student performance (see Appendix D).

About 25 percent of the student works were scored independently by another research assistant. The overall level of agreement between the two raters on each standard is indicated in Table A2. When ratings on the three standards were added, the overall level of agreement between the two raters on the total score was estimated as a correlation coefficient of .63.

Table A2
Interrater Agreement for Instruction, Assessment Tasks, and Student Performance on Authentic Standards

	Exact Agreement (%)	Exact or Off by 1 (%)
Instruction (N = 31 lessons)		
Standard 1: Higher-Order Thinking	58	100
Standard 2: Deep Knowledge	45	81
Standard 3: Substantive Conversation	35	84
Standard 4: Connection to the World Beyond the Classroom	57	100
Assessment Tasks (N= 22 tasks)		
Standard 1: Organization of Information	59	96
Standard 2: Consideration of Alternatives	46	86
Standard 3: Disciplinary Content	50	100
Standard 4: Disciplinary Processes	77	100
Standard 5: Elaborated Written Communication	64	91
Standard 6: Problem Connected to the World Beyond the Classroom	82	100
Standard 7: Audience Beyond the School	73	86
Student Performance (N = 167 papers)		
Standard 1: Analysis	61	98
Standard 2: Disciplinary Concepts	62	97
Standard 3: Elaborated Written Communication	62	97
Source: Ratings of classroom observations, assessment tasks, and student work samples, fall 1998 and spring 1999.		



APPENDIX B

AUTHENTIC INSTRUCTION

SCORING RUBRIC

Scoring Criteria for Authentic Classroom Instruction

(from Newmann, F.M., Secada, W.G. & Wehlage, G.G. (1995). *A guide to authentic instruction and assessment: Vision, standards and scoring*. Madison, WI: Wisconsin Center for Education Research, pp. 86 - 93.)

Note: For an in-depth description of the standards, see Chapter 3 in *A Guide to Authentic Instruction and Assessment: Vision, Standards and Scoring*, Newmann, Secada, & Wehlage.

Standard 1. Higher Order Thinking: Instruction involves students in manipulating information and ideas by synthesizing, generalizing, explaining, hypothesizing, or arriving at conclusions that produce new meaning and understandings for them.

_____5 = Almost all students, almost all of the time, are performing HOT.

_____4 = Students are engaged in at least one major activity during the lesson in which they perform HOT operations. This activity occupies a substantial portion of the lesson and many students are performing HOT.

_____3 = Students are primarily engaged in routine LOT operations a good share of the lesson. There is at least one significant question or activity in which some students perform some HOT operations.

_____2 = Students are primarily engaged in LOT, but at some point they perform HOT as a minor diversion within the lesson.

_____1 = Students are engaged only in LOT operations; i.e., they either receive, or recite, or participate in routine practice and in no activities during the lesson do students go beyond LOT.

Standard 2. Deep Knowledge: Instruction addresses central ideas of a topic or discipline with enough thoroughness to explore connections and relationships and to produce relatively complex understandings.

_____5 = Knowledge is very deep because during the lesson almost all students do at least one of the following: sustain a focus on a significant topic; or demonstrate their understanding of the problematic nature of information and/or ideas; or demonstrate complex understanding by arriving at a reasoned, supported conclusion; or explain how they solved a complex problem. In general, students' reasoning, explanations and arguments demonstrate fullness and complexity of understanding.

_____4 = Knowledge is relatively deep because either the teacher or the students provide information, arguments or reasoning that demonstrate the complexity of an important idea. During the lesson many students do at least one of the following: sustain a focus on a significant topic for a period of time; or demonstrate their understanding of the problematic nature of information and/or ideas; or demonstrate understanding by arriving at a reasoned, supported conclusion; or explain how they solved a relatively complex problem.

_____3 = Knowledge is treated unevenly during instruction; i.e., deep understanding of something is countered by superficial understanding of other ideas. At least one significant idea may be presented in depth and its significance grasped, but in general the focus is not sustained.

_____2 = Knowledge remains superficial and fragmented; while some key concepts and ideas are mentioned or covered, only a superficial acquaintance or trivialized understanding of these complex ideas is evident.

_____1 = Knowledge is very thin because it does not deal with significant topics or ideas; teacher and students are involved in the coverage of simple information which they are to remember.

Standard 3. Substantive Conversation: Students engage in extended conversational exchanges with the teacher and/or their peers about subject matter in a way that builds an improved and shared understanding of ideas or topics.

Substantive conversation has three features:

- (1) The talk is about subject matter in the discipline and includes higher order thinking, such as making distinctions, applying ideas, forming generalizations, or raising questions; not just reporting of experiences, facts, definitions, or procedures.
- (2) The conversation involves sharing of ideas and is not completely scripted or controlled by one party (as in teacher-led recitation). Sharing is best illustrated when participants explain themselves or ask questions in complete sentences, and when they respond directly to comments of previous speakers.
- (3) The dialogue builds coherently on participants' ideas to promote improved collective understanding of a theme or topic (which does not necessarily require an explicit summary statement).

_____5 = All three features of substantive conversation occur, with at least one example of sustained conversation, and almost all students participate.

_____4 = All three features of substantive conversation occur, with at least one example of sustained conversation, and many students participate.

_____3 = Features 2 (sharing) and/or 3 (coherent promotion of collective understanding) occur and involve at least one example of sustained conversation i.e., at least 3 consecutive interchanges).

_____2 = Features 2 and/or 3 occur briefly and involve at least one example of two consecutive interchanges.

_____1 = Virtually no features or substantive conversation occur during the lesson.

Standard 4. Connections to the World Beyond the Classroom: Students make connections between substantive knowledge and either public problems or personal experiences.

_____5 = Students study or work on a topic, problem or issue that the teacher and students see as connected to their personal experiences or actual contemporary public situations. Students recognize the connection between classroom knowledge and situations outside the classroom. They explore these connections in ways that create personal meaning and significance for the knowledge. This meaning and significance is strong enough to lead students to become involved in an effort to affect or influence a large audience beyond their classroom in one of the following ways: by communicating knowledge to others (including within the school), advocating solutions to social problems, providing assistance to people, creating performances of products with utilitarian or aesthetic value.

_____4 = Student study or work on a topic, problem or issue that the teacher and students see as connected to their personal experiences or actual contemporary public situations. Students recognize the connection between the classroom knowledge and situations outside the classroom. They explore these connections in ways that create personal meaning and significance for the knowledge. However, there is no effort to use the knowledge in ways that go beyond the classroom to actually influence a large audience.

_____3 = Students study a topic, problem or issue that the teacher and students see as connected to their personal experiences or actual contemporary public situations. Students recognize the connection between classroom knowledge and situations outside the classroom, but they do not explore the implications of these connections which remain abstract or hypothetical. There is no effort to actually influence a large audience.

_____2 = Students encounter a topic, problem or issue that the teacher tries to connect to students' experiences or to contemporary public situations; i.e., the teacher informs students that there is potential value in the knowledge being studied because it relates to the world beyond the classroom. For example, students are told that understanding Middle Eastern history is important for politicians trying to bring peace to the region; however, the connection is weak and there is no evidence that the students make the connection.

_____1 = Lesson topic and activities have no clear connection to anything beyond itself; the teacher offers no justification beyond the need to perform well in class.

APPENDIX C

AUTHENTIC ASSESSMENT

SCORING RUBRIC

Scoring Criteria for Assessment Tasks

(from Newmann, F.M., Secada, W.G. & Wehlage, G.G. (1995). *A guide to authentic instruction and assessment: Vision, standards and scoring*. Madison, WI: Wisconsin Center for Education Research, pp. 81-85)

Note: For an in-depth description of the standards, see Chapter 2 in Newmann, Secada, & Wehlage (1995).

Standard 1. Organization of Information: The task asks students to organize, synthesize, interpret, explain, or evaluate complex information in addressing a concept, problem, or issue.

Consider the extent to which the task asks the student to organize, interpret, evaluate, or synthesize complex information, rather than to retrieve or to reproduce isolated fragments of knowledge or to repeatedly apply previously learned procedures. To score high, the task should call for interpretation of nuances of a topic that goes deeper than surface exposure or familiarity.

When students are asked to gather information for reports that indicates some selectivity and organizing beyond mechanical copying, but are not asked for interpretation, evaluation, or synthesis, give a score of 2.

_____3 = high
_____2 = moderate
_____1 = low

Standard 2. Consideration of Alternatives: The task asks students to consider alternative solutions, strategies, perspectives, or points of view as they address a concept, problem, or issue.

To what extent does success in the task require consideration of alternative solutions, strategies, perspectives and points of view? To score high, the task should clearly involve students in considering alternatives, either through explicit presentation of the alternatives or through an activity that cannot be successfully complete without examination of alternatives implicit in the work. It is not necessary that students' final conclusions include listing or weighing of alternatives, but this could be an impressive indicator that is was an expectation of the task.

_____3 = high

_____2 = moderate

_____1 = low

Standard 3: Disciplinary Content: The task asks students to show understanding and/or use of ideas, theories, or perspectives considered central to an academic or professional discipline.

To what extent does the task promote students' understanding of and thinking about ideas, theories or perspectives considered seminal or critical within an academic or professional discipline, or in interdisciplinary fields recognized in authoritative scholarship? Examples in social studies could include democracy, social class, market economy, or theories of revolution. Reference to isolated factual claims, definitions, algorithms—though necessary to inquiry within a discipline—will not be considered indicators of significant disciplinary content unless the task requires students to apply powerful disciplinary ideas that organize and interpret the information.

_____3 = Success in the task clearly requires understanding of concepts, ideas, or theories central in a discipline.

_____2 = Success in the task seems to require understanding of concepts, ideas or theories central in a discipline, but the task does not make these very explicit.

_____1 = Success in the task can be achieved with a very superficial (or even without any) understanding of concepts, ideas, or theories central to any specific discipline.

Standard 4: Disciplinary Process: The task asks students to use methods of inquiry, research, or communication characteristic of an academic or professional discipline.

To what extent does the task lead students to use methods of inquiry, research, communication, and discourse characteristic of an academic or professional discipline? Some powerful processes of inquiry may not be linked uniquely to any specific discipline (e.g., interpreting graphs), but they will be valued here if the task calls for their use in ways similar to important uses within the discipline.

_____3 = Success in the task requires the use of methods of inquiry or discourse important to the conduct of a discipline. Examples of methods of disciplinary inquiry would include interpreting primary sources.

_____2 = Success in the task requires use of methods of inquiry or discourse not central to the conduct of a discipline.

_____1 = Success in the task can be achieved without use of any specific methods of inquiry or discourse.

Standard 5. Elaborated Written Communication: The task asks students to elaborate on their understanding, explanations, or conclusions through extended writing.

_____4 = Analysis/Persuasion/Theory. The task requires explanations of generalizations, classifications and relationships relevant to a situation, problem, or theme. Examples include attempts to argue, convince, or persuade and to develop or test hypotheses.

_____3 = Report/Summary. The task calls for an account of particular events or series of events (“This is what happened.”), a generalized narrative, or a description of a recurrent pattern of events or steps in a procedure (“This is what happens.” “This is the way it is done.”)

_____2 = Short-answer exercises. Only one or two brief sentences per question are expected.

_____1 = Multiple Choice Exercises; Fill-in-the-blank Exercises (answered with less than a sentence)

Standard 6. Problem Connected to the World Beyond the Classroom:

The task asks students to address a concept, problem, or issue that is similar to one that they have encountered, or are likely to encounter, in their professional, personal, or civic lives beyond the classroom.

To what extent does the task present students with a question, issue, or problem that they have actually encountered, or are likely to encounter, in their lives beyond school? In social studies, defending one's position on compulsory community service for students could qualify as a real world problem, but describing the origins of World War II generally would not.

_____3 = The question, issue, or problem clearly resembles one that students have encountered, or are likely to encounter, in their professional, personal, or civic lives beyond the classroom. The resemblance is so clear that teacher explanation is not necessary for most students to grasp it.

_____2 = The question, issue, or problem bears some resemblance to real world experiences of the students, but the connections are not immediately apparent. The connections would be reasonably clear if explained by the teacher, but the task need not include such explanations to be rated 2.

_____1 = The problem has virtually no resemblance to questions, issues, or problems that students have encountered, or are likely to encounter, beyond school. Even if the teacher tried to show the connections, it would be difficult to make a persuasive argument.

Standard 7: Audience Beyond the School: The task asks students to communicate their knowledge, present a product or performance, or take some action for an audience beyond the teacher, classroom, and school building.

_____4 = Final product is presented to an audience beyond the school.

_____3 = Final product is presented to an audience beyond the classroom, but within the school.

_____2 = Final product is presented to peers within the classroom.

_____1 = Final product is presented only to the teacher.

APPENDIX D

AUTHENTIC PERFORMANCE SCORING RUBRIC

Scoring Criteria for Student Performance

(from Newmann, F.M., Secada, W.G. & Wehlage, G.G. (1995). *A guide to authentic instruction and assessment: Vision, standards and scoring*. Madison, WI: Wisconsin Center for Education Research, pp. 94-102.)

Note: For an in-depth description of the standards, see Chapter 4 in *A Guide to Authentic Instruction and Assessment: Vision, Standards and Scoring*, Newmann, Secada, & Wehlage.

Standard 1. Analysis: Student performance demonstrates higher order thinking with social studies content by organizing, synthesizing, interpreting, evaluating, and hypothesizing to produce comparisons/contrasts, arguments, application of information to new contexts, and consideration of different ideas or points of view.

____4 = Substantial evidence of analysis. Most of the student's work includes analysis.

____3 = Moderate evidence of analysis. A central portion of the student's work includes analysis.

____2 = Some evidence of analysis. A small but not central, portion of the student's work includes analysis.

____1 = No evidence of analysis. Almost all statements consist of recording, or reporting specific information, without evidence of the student's organizing it or reflecting upon it; OR virtually all analysis offered is unsuccessful or in error.

Standard 2. Disciplinary Concepts: Student performance demonstrates an understanding of ideas, concepts, theories, and principles from the social disciplines and civic life by using them to interpret and explain specific, concrete information or events.

____4 = The student has used social studies concepts to organize, explain, interpret, summarize, and extend the meaning and significance of otherwise discrete pieces of information. The use of the ideas illustrates exemplary understanding.

____3 = The student has included social studies concepts to organize, explain, interpret, summarize, and extend the meaning and significance of otherwise discrete pieces of information. The use of the ideas is somewhat limited and/or shows some flaws in understanding.

____2 = Social studies concepts are included, but their use is significantly limited and/or shows significant flaws in understanding.

____1 = The work includes virtually no social studies concepts, or the use of any that are included shows almost no understanding.

Standard 3. Elaborated Written Communication: Student performance demonstrates an elaborated account that is clear, coherent, and provides richness in details, qualifications and argument. The standard could be met by elaborated consideration of alternative points of view.

____4 = Exceptional. The writer provides substantial and accurate elaboration for two or more important statements. The details, qualifications, and nuances are expressed within an overall coherent framework intended for the reader, and relevant to the topic. The response is so rich as to be worthy of display as an outstanding example of writing in social studies.

____3 = Elaborated. The writer provides some elaboration for two or three important statements OR provides substantial elaboration for one important statement. In either case, the details, qualifications, and nuances are expressed within a coherent overall framework intended for the reader, relevant to the topic, and without major inaccuracies.

____2 = Minimal. The writer provides reasonably accurate elaboration for at least one important statement.

____1 = Unsatisfactory. The writer provides virtually no information or provides only disjointed details. OR, the writer provides discrete claims, broad generalizations, slogans, or conclusions, but none are elaborated.