Proc. Natl. Sci. Counc. ROC(D)
Vol. 10, No. 1, 2000. pp.15-23

The Role of Student Outcome Data in Teacher Certification

JAMES A. SHYMANSKY

E. Desmond Lee Professor of Science Education
University of Missouri-St. Louis
St. Louis, Mo. U.S.A.

(Received October 12, 1999; Accepted January 19, 2000)

Abstract

In this paper the need for using student outcome data in the certification of science teachers is argued. A brief review of the basic model underpinning current certification practices is presented to demonstrate that extant practices ignore student outcome data despite student, parent, public, and school officials' interest in these data. The need for a comprehensive program of research into the effects of teacher behaviors and practices on student outcomes is discussed. Finally, results of a teacher inservice project in which student outcome data were used to evaluate the effectiveness of the project and a description of one state's efforts to use student outcome data in school certification are presented to support arguments that student outcome data can and should be used in the ongoing certification of teacher cohorts and possibly of teachers individually.

Key Words: certification, licensure, accountability

I. The Issue

The supply of quality teachers is an ongoing topic of concern and debate in almost every industrialized country including Taiwan (Guo, 1995). Education issues are probably second only to the economy and are an especially favorite topic of debate during governmental elections. In the United States, potential presidential candidates for the 2000 election are making speeches that express strong views on what they see as the problems with education and what must be done to improve schools. For example, presidential hopeful Elizabeth Dole recently told attendees at a conference of the American Council of Education that "colleges and universities must do a better job of educating the teachers of the future." She went on to say that "in too many schools teachers don't have college majors or even minors in the subjects they teach" and that "new teachers should be subject to performance exams" (Walsh, 1999). On the same day, US Secretary of Education Riley promoted a national teacher licensing plan consisting of three levels: an initial three year license for new teachers who pass a written exam and a classroom performance test; a professional license, awarded to teachers who meet state standards and pass a panel review; and a voluntary advanced license issued through the National Board for Professional Teaching Standards. Riley further recommended that teachers' salaries be based on both license type and years of experience (Henry, 1999).

The statements by candidate Dole and Secretary Riley are significant in several ways. Riley's statement is evidence that more and more pressure is being applied for a national system of teacher certification and licensure in the US. Such a national system would reverse a long tradition of state control of education. His suggestion that salaries be tied to level of licensure would also require radical restructuring of teacher salary scales, the cornerstone of teacher unions in the US. Dole's statements on the other hand don't suggest any specific system of certification, but do imply what defines a quality teacher-a person who is steeped in content. Though this statement appears reasonable enough on the surface, it raises some serious questions, e.g., for science educators: do science teachers need the same education as science majors? Are science fac-

ulty interested in preparing teachers? Are science faculty the best models of good science teaching?

The Dole and Riley statements are more significant for what they don't say than for what they do say. They both skirt the fundamental issue of what a good teacher can do that a poor teacher cannot-and that is to produce students who have learned important lessons and who choose to learn even more on their own when they are not in school. What seems to be getting lost in current discussions of teacher competence and quality is the fact that student learning is the product of teaching. By definition, if there is no learning, there was no teaching. And if this is so, should student performance not be included as a primary measure of teacher competence and quality?

In this paper the link between teacher competence and student performance and how it relates to teacher certification is explored and suggestions for how student performance might be integrated into the teacher certification process will be presented.

II. The Basis of Teacher Evaluation

Teacher evaluation systems (and teacher education programs as well) assume that teaching traits, behaviors and practices are identifiable, stable and consistent in their effects on student learning (Medley, 1985). A simple translation of this statement appears in Fig. 1. The model implies that teachers can be certified on the basis of such things as their level of content expertise, their years of classroom experience, their skill at questioning, listening and responding, their ability to plan and organize instruction, etc. because these teacher factors are thought to impact such things as student attitudes, self concept, achievement, problems solving skills, creativity, etc. The model further implies that as new teacher factors are identified, they can be added to the model.

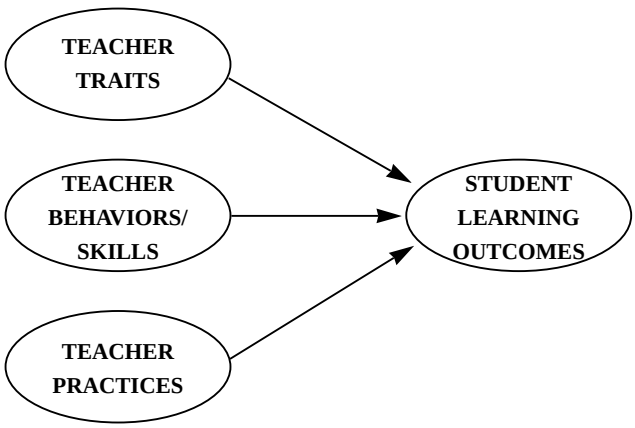

Fig. 1. A model of teaching-learning connections.

In the United States the National Board of Professional Teaching Standards (NBPTS) has been aggressively pursuing the development of teacher evaluation systems that focus on what Shulman describes as "pedagogical content knowledge" (PCK) (1988). PCK is concerned with situation-specific teacher traits, behaviors and practices (.e.g., "what a teacher need to know and be able to do to teach simple electrical circuits to fourth graders in Taipei"). In the NBPTS system, for example, candidates seeking certification as a master teacher of "adolescent and young adulthood science" (AYA/Science) must prepare a portfolio and complete specific assessment center activities. In the portfolio they must provide evidence of both classroom-based practices and work outside the classroom with families, community, and colleagues. Evidence of classroom competence includes: (a) teaching a major idea over time, (b) assessing student work in a content area, (c) facilitating student inquiry of specific concepts, and (d) conducting whole-class discussions of important science ideas. Candidates must also complete exercises in four assessment centers. At one of the centers they are asked to design instruction or describe an instructional strategy to teach students an important piece of science content. Elements in the teaching portfolio as well as responses to exercises at each of the assessment centers are scored using a modified four-point rating scale. Certification decisions are based on the total scaled score that candidates earn on the overall assessment-the weighted portfolio and center scores combined (Gitomer, 1999).

The NBPTS master teacher certification process is clearly the most sophisticated, time-intensive and expensive of the systems that the NBPTS is promoting. To date, only a handful of teachers have gone through the NBPTS process and been awarded a master teacher certificate. Few states or school districts hold their teachers to anything that even closely approximates anything like the NBPTS "board certification." Currently only two states recognize and reward teachers who earn NBPTS certification. But the pressure on schools of education and on practicing teachers continues to mount.

III. Missing Ingredient

My purpose in reviewing and modeling the 1985 Medley statement and providing a brief discussion of the NBPTS process is to show that, despite the rigor and technical merit of the NBPTS, it requires no evidence of student performance for any level of certification. Half of the teaching-learning model in Fig. 1 is

not being considered in the NBPTS or in most other certification systems being promoted. Is this a problem? I think that it is. And others agree. Andrews and Barnes in their 1990 review of the teacher evaluation literature put it this way: "Teacher assessment systems that do not include some measure of student growth will always be suspect by policymakers (p. 595)." If you don't believe this is true, just think about what kinds of education news stories get the most media coverage. It is not those that focus on teacher accomplishments; it is those that report on student performances, particularly those that compare students over time or between schools, cities, states, and countries.

A case in point is the "Third International Mathematics and Science Study" (TIMSS). About two years ago TIMSS results appeared in almost every newspaper in the United States and became a major topic of discussion on radio and television talk shows and at education meetings at all levels. The data from the TIMSS project haven't even been fully analyzed yet and there are already plans underway to conduct a "fourth" (FIMSS) study in the new millenium. States and individual school systems in the United States are lining-up to be included in the new test samples so that they can get results specific to them. Why? Because when all is said and done, everyone--parents, politicians, teachers, administrators, even students themselves--want to know how well or poorly the students have performed. So keen is the interest in student performance that more than a billion dollars is spent every year on formally testing students in the US!

Clearly, everyone is interested in student test results like the TIMSS. But educators generally oppose, or at least find dozens of reasons against using student performance as a measure of teacher competence on the grounds that there are "too many factors other than the teacher that influence student performance" and that "most tests don't measure what is really important." Establishing causality in the teaching-learning context is certainly challenging and tests can always be improved. However, these arguments of "too many factors" and "invalid tests" are used even when the connection between teacher competence and student performance are clear and non-controversial.

So is it fair to use student performance to measure teacher effectiveness? I say yes! In what other profession is one's competence not in some way measured by the product of the professional? Physicians are ultimately judged by their success in curing illness. Attorneys are judged by their success in winning cases. Investment counselors are judged by how much money they make for their investors. And the list goes on. Obviously, the point about linking competence to outcomes is an over-simplification. Physicians don't always cure every patient, nor do attorneys win every case. And investment councilors surely don't make money for every investor. But would you go to a physician who routinely loses patients? To an attorney who rarely wins cases? To an investment councilor who continually loses clients' money? I think not. So why should we not consider that which defines a teacher's product-student learning-when it comes to evaluating teachers? I say we should.

IV. A Case Study

As an example of how student test results relate to issues of teacher certification and licensure, let's examine the case of one school district's efforts to improve their elementary science curriculum by enhancing their teachers' PCK through a long-term inservice program. In 1994, leaders from a school district partnered with science educators from the University of Iowa and launched a reform project. At the inservice summer workshops and school year meetings, teachers learned important science content embedded in a teaching strategy that utilized children's literature as a tool to get children to reflect about their science ideas and parents as partners to help the teachers promote science at home. The reform targeted the elementary school teachers who, for the most part, were very weak in their own science preparation and taught little or no science at all. The project started with a small cadre of teachers and expanded each year to include additional teachers in the district. Within four years, approximately 80% of the district's teachers had received approximately 120 hours of inservice training.

As the project unfolded, everyone in the district was feeling very positive about its success. Teachers, parents and administrators spoke enthusiastically about the use of the literature and the involvement of the parents. In their hearts, they knew the project was good for the teachers, the community, and most of all for the students. But while such "cardiac" data made people feel good, they said nothing about the impact on teacher or student performance. Thus, project staff faced the decision of how to assess the project's impact: examine the changes in teachers; examine the changes in students; or both? The decision was made to do some of both.

In the Spring, 1998, teachers were rated by the district's science supervisor on the pedagogical and content factors stressed in the inservice activities. Project staff also gathered data on all students' attitudes and perceptions and on the Grade 3 and 4 stu-

Table 1. Relationship Between Supervisor Ratings and Years of Inservice

	Inservice	General Reform Practices	Project-Specific Practices
Inservice	1.000		
General Reform	0.753	1.000	
Project Specific	0.724	0.890	1.000

dents' science achievement using six forms of the TIMSS. The staggered entry of teachers into the project and the fact that some 20% of the teachers had not yet participated by the end of the fourth year made it possible for project staff to conduct various analyses of the impact of the training program on both the teachers and the students.

V. Impact on Teachers

The science supervisor's ratings of all the teachers in the district provided the opportunity to study the difference between participating and non-participating teachers and the effects of different levels of inservice across the four years. To facilitate the examination of this question, every teacher in the school district was rated by the district's science supervisor on two dimensions: (1) their understanding and use of principles and practices prescribed in the "science reform" documents (i.e., the National Science Education Standards (1996) and (2) their understanding and use of principles and practices specifically stressed in the inservice sessions (the use of children's literature, parents as partners, and concept-specific challenge activities). It is important to note that participating and non-participating teachers did not differ in mean age, classroom experience or levels of previous preparation. Table 1 contains a summary of the statistics correlating supervisor ratings to years of inservice involvement.

The data in Table 1 show that the supervisor clearly viewed teachers with greater project involvement as having improved understanding and facility with principles and practices of science reform (r=0.753) and of project-specific practices (r=0.724). These data are not surprising since the project promoted both. We could easily envision examining the data in other ways or gathering additional data relating years of inservice to supervisor ratings to find out exactly how many years of inservice a teacher would need to reach an acceptable level of competence in using the practices stressed.

The Table 1 data are not surprising in a second way as well. The supervisor was deeply invested in the project; she wanted it to be successful. But should in-house or self-ratings ever be trusted? I raise this issue because much of the current teacher certification and licensure procedures lie completely in the hands of those who do the training. Having an independent body such as the NBPTS certify and license teachers seems to make good sense if for no other reason, the use of an outside agent significantly reduces the conflicts of interest when universities are pressured to show success to maintain enrollments.

VI. Impact on Students

Having an impact on teachers is only important to the extent that it translates to impact on those teachers' students. Specifically, school district and project staff were interested in learning how the trait variable of years of professional development related to student perceptions of their science program, their attitudes toward science, and their achievement in science. The "student perceptions" of the teachers' behaviors and teaching strategies variable is interesting because of the long tradition in schools of using student feedback as a measure of teacher effectiveness. Even if the students don't perform any differently on achievement or attitude measures, can they perceive differences in the teacher and the science program as the teacher spends more and more time learning how to use special strategies such as those stressed in the case study project. Table 2 contains summary statistics for the analysis of the impact of years of inservice training on student perceptions of their science.

The inservice training focused heavily on getting teachers to pay attention to their student's ideas and challenging those ideas with inquiry activities (i.e.,

Table 2. Summary Statistics for Student Perceptions and Years of Teacher Inservice

Years	N	Use of my ideas	Parent interest	Use of stories
0	224	2.9	2.4	2.5
1	259	3.1	2.6	2.7*
2	150	2.6	2.4	2.8
3	183	3.0	2.6	2.8
4	23	3.5*	3.4*	2.8

Table 3. Summary Statistics for Student Attitudes and Achievement as a Function of Their Teacher's Years of Inservice Involvement

Years	N	Attitude	TIMSS (MC)1	TIMSS (CR)2
0	224	3.0	15.9	0.2*
1	259	2.9	15.5	-0.1
2	150	2.9	15.2	-0.2
3	183	2.8	15.2	0.0
4	23	3.0	15.1	0.1

[1] Multiple choice maximum score = 25
[2] Different scales on different test forms were converted to Z-scores for analysis

adopting a constructivist teaching approach). The children's literature and parent partners were part of the larger constructivist strategy. We were therefore interested in knowing whether or not the students felt that their teacher was asking for their ideas and using them during the lessons; whether or not the students saw the stories that their teacher was using as being an important part of the science instruction; and whether or not the students saw their parents as active partners in their science instruction. As can be seen in Table 2, students rated the 4-year teachers significantly higher in "use of my ideas" and in getting parents involved or "interested." Students rated the "use of stories" significantly higher after their teacher had only one year of inservice. These data suggest that students were actually seeing and feeling a difference on those dimensions that were stressed in the project.

However, is it enough that students perceive differences in how they are being taught science? Do these differences translate in improved student attitudes or achievement, i.e., do students like science better, have more confidence in their ability to do science, and do they score higher on achievement measures as a result of their teacher having more years of professional development? Table 3 contains summary statistics for the analyses relating years of inservice training on student attitudes and achievement.

There are two things that stand out in Table 3: (1) the flat means on the attitude measure and on the multiple choice section of the TIMSS test, and (2) the drop in performance on the constructed response portion of the TIMSS test among students in classes of teachers with one and two years of inservice training. The data in Table 3 clearly show that "cardiac" data (i.e., everyone in the school district "feeling" that the project was having a great impact) are probably not valid as an indicator of program success. The data also cast doubt on models of teacher certification and licensure that are based on the teacher trait of years of participation in professional development-a person is not necessarily a better teacher (as measured in student performance) just because she/he spent X years in preparation or professional development!

VII. Linking Teaching to Learning

If years of training (and I would also argue, traits such as years of classroom experience and content knowledge) are not reliable predictors of student performance, is it possible (and should we try) to link teacher behaviors, skills, and practice to student performance? Again, I think definitely, yes! The supervisor ratings of all teachers in the school district presented an opportunity in the project to link teachers individually and collectively to their students' performance. Again, the idea was to look specifically at how the level of teachers' PCK (as defined by the supervisor rating scale items) were related to students' perceptions of their science program, their attitudes toward science, and their achievement.

Students were asked to identify their current year science teacher and their science teachers from the previous two years. Knowing a student's teachers for three consecutive years provided the opportunity to examine two different links between teacher performance and student performance. The district supervisor ratings of the teachers allowed us to link the current year teacher to his/her students directly. Table 4 contains summary statistics for the analyses of student perceptions, attitudes, and achievement as a function of the rating or quality of instruction received from their current year teacher.

Table 4. Summary Statistics for Student Perceptions, Attitudes and Achievement as a Function of Supervisor Ratings of the Students' Current Year Teacher

Quality	N	Use of my ideas	Parent interest	Use of stories	Attitudes	TIMSS (MC)
Low	181	2.7	2.4	2.5	2.9	15.2
Medium	402	2.9	2.6	2.7	2.9	15.0
High	263	3.2*	2.7*	2.8*	2.9	15.3

Table 5. Summary Statistics for Student Perceptions, Attitudes and Achievement as a Function of Quality of Instruction Received by Students Across Three Years

Quality	N	Use of my ideas	Parent interest	Use of stories	Attitudes	TIMSS (MC)
Low	121	3.0	2.7	2.7	3.0	15.7
Medium	403	3.0	2.6	2.7	3.0	15.8
High	47	3.2	2.8*	2.8	2.8	13.2*

The Table 4 results are strikingly similar to the Table 2 results. Students were able to detect differences in how their teachers used student ideas, stories in science, and parents as partners in classrooms taught be the highest rated teachers, but these students showed no more-positive attitudes in science or no higher performance on the multiple choice section of the TIMSS test.

VIII. Aggregate Teacher Effects

The supervisor ratings of all three teachers were also combined to yield a "teacher team rating" score for each student. This score allowed the project staff to examine the cumulative effect of exposure to different qualities of science teaching across a three-year period on student performance. Each of the three teacher ratings for a given student were combined into a three year score for both the "general reform" and the "project-specific" subscale ratings. Three years scores that totaled 54-60 on the science reform subscale and 39-45 on the project-specific subscale (meaning all three teachers had to receive 4s and 5s on the 5-point rating scale) defined science instruction of "high quality." Lower point totals were used to define three year ratings of "medium" and "low" quality instruction. This cumulative effect study is interesting because it begins to approximate what might be called an aggregate "school" effect on student performance-the variable that is implied in reports comparing student performance across different school districts. Table 5 contains summary statistics for the analyses of student perceptions, attitudes, and achievement as a function of the quality of instruction received by a student across three years.

The Table 5 data are important for the significantly higher "parent involvement" results and the significantly lower "achievement" results for students having experienced three years of the best teachers in the district. These students saw their parents as more involved, but these same students performed more poorly on the TIMSS test. Think about these results: combinations of the "best" teachers (i.e., rated "best" by the science supervisor) produced students who scored no better on measures of perception and attitude and significantly lower on achievement than teacher combinations that were not rated as highly. This was disappointing news for the project staff who were certainly hoping that the teachers considered best by the district's science supervisor would turnout students with the most positive attitudes and the highest achievement scores. While it may be that the TIMSS simply did not assess elements of science literacy nurtured by the "best" teachers, these results should, at the very least, force the district to search for student assessment strategies that will support the science supervisor's ratings or force the science supervisor to redefine the criteria used to rate elementary science teachers in the district.

VX. Conservative Implications

It is fairly certain that certification and licensure processes will continue to depend heavily on teacher traits that are assumed to be effective in helping students learn, be it science, mathematics, history, or whatever. Given this reality of certification practices, the task is very straightforward: collect data that show how various teacher traits and skills are linked to and impact student learning. One approach is to identify teachers who consistently produce students who outperform other students on outcomes considered valuable and to profile those teachers. This is the approach that Yager and Penick effectively followed in their "Exemplary Programs" series (1983). Another approach is to synthesize the research on the links between teacher traits, behaviors and practices and student outcomes as was done in the meta-analyses lead by Anderson (1983). As promising as either of these techniques might be, the research base on teacher effectiveness has grown very little in the past 15 years (Andrews & Barnes, 1990, p.595). Moreover, the research which is being done tends to focus on generic teaching traits and skills, despite research that suggests that teaching is very subject and situation specific (Shulman, 1987, 1988).

The task of documenting the effects of teacher traits and skills on student outcomes in specific situations is daunting. A disturbing reality that complicates the task is that teachers are very reluctant to participate in studies of their own classroom teaching and to be judged individually on their students' performance.

While it may be human nature to avoid evaluation, education cannot hope to survive as a profession unless teachers are willing to do research themselves and provide access to others to study their impact on student learning. What is needed is a comprehensive "map" of how teacher traits, skills and practices impact student performance-- a "Teacher Impact" (Tea-Pact) project akin to the "Human Genome" project in the field of genetics. Like the Genome project which is attempting to identify all the estimated 80,000 genes in the human DNA and determine the sequences of the 3 billion chemical bases that make up the DNA and their links to human diseases, a Tea-Pact project would try to identify patterns of critical teacher traits, behaviors and practices and link them to student outcomes. As I have said, the task is daunting. But as Andrews & Barnes have stated, "until the link between teacher behavior and impact on students is forged, the assessment of teaching will remain a developmental science" (1990, p.595).

X. Radical Implications

Perhaps it is not necessary to map all the complex teacher effects-on-student performance connections at the individual teacher trait, behavior and practice level. After all, even the most reductionist of those among us would probably still admit that a good part of effective teaching is an "art." But if we even accept even a modest notion that teachers should be judged in some way on student performance, can student data such as those collected in the inservice project previously described help us certify science teachers? Yes and no. We are not at a place yet where we can hold teachers directly accountable for how their students perform on tests of importance in any given year (e.g., a TIMSS-type test). There are simply too many extenuating circumstances which are beyond the control of one teacher. Two important ones are the resources available to the teachers and the differences in student populations. Though it might be possible to equalize available resources through school funding, accounting for differences in students is not quite so easy. Even with these limitations however, it is critical that educators begin to incorporate student performance in teacher certification decisions.

At least one US state, Kentucky, has experimented with an accountability system based on student performance data. About 5 years ago state officials in Kentucky implemented statewide testing of students in several areas of the curriculum including science. The idea of state testing was not particularly new, but the idea of using test results to certify schools was. Under the Kentucky plan, baselines of student performance were established over a period of two years. These baselines became the benchmark against which student performance was to be judged in the first and subsequent years of formal statewide testing. Test score targets were set for schools whose students were judged to be performing below a minimum standard across the benchmark setting years. The state allocated special teacher training and instructional material funds to help low-performing schools meet their performance targets. Schools not meeting targets or not making significant progress toward meeting their performance targets were given even more special professional development. If targets were still not met, the school could eventually be taken over by the state. Schools meeting or exceeding their performance targets received special incentives including salary bonuses.

In the Kentucky plan it was never intended that individual teachers would lose their certification if the school was taken over by the state. But if student performance targets are not met, the state focuses remediation efforts on teachers, not students. Teachers must attend special inservice training programs to enhance their subject knowledge and teaching skills. The schools are also given special support and monitored continually. Once schools have met their performance targets, they can reclaim local control.

The Kentucky program has not been without its problems. Critics are quick to point out that the student performance tests being used may not be measuring all that is learned or things that are really important. This however, is an equally valid argument for the continued improvement of tests and their interpretation as it is against the idea of using student performance in school and teacher certification.

The Kentucky plan is effectively a "school certification" system. Can such a plan work at the individual teacher level? My response: a tentative, "yes," at the inservice teacher level. In the United States, most states have done away with the concept of permanent teacher certification and implemented programs of *continuing certification*. Under this program teachers must update their knowledge and skills every 5-10 years. Even so, no re-certification process requires teachers to produce any evidence that students have learned under their instruction in the previous certification period. If we are to take seriously the idea that student performance <u>is</u> the product of teaching, we need to figure out ways to include such evidence in the re-certification process at the very least. One approach that seems both reasonable and feasible is to have practicing teachers provide evidence of student learning across the previous period of certification at the

time they seek re-certification. New teachers to the field would not be required to provide such evidence at the time of initial certification, but they would be required to provide student data at the time of their first re-certification. Target student performance goals for the 5-10 year certification period could be negotiated. Teachers whose students fail to meet performance targets could be given special support and inservice. At the other end of the spectrum, teachers who are seeking or wish to retain "master teacher" status for a certification period should be expected to provide evidence of significant student performance gains beyond the target performance levels set or maintenance of already high student performance levels.

XI. Concluding Remarks

In this paper I have argued for the need to include evidence of student performance in the certification of teachers. Requiring such evidence will demand our continued development of valid and cost-effective student performance assessments. In arguing for the inclusion of student performance data in certifying teachers, I am not arguing against the use of teacher trait and skills in the training or initial certification of teachers. In fact, a greater emphasis on using student performance in certification decisions will force us to research the linkages between teacher traits and skills and student performance and improve our teacher education and initial certification programs.

References

Anderson, R.D., Kahl, S.R., Glass, G.V. & Smith, M.L. (1983). Science education: A meta-analysis of major questions. *Journal of Research in Science Teaching*, **20**(5), 379-385.

Andrews, T.E., Barnes, S. Assessment of teaching. In W.R. Houston, M. Haberman & J. Sikula (Ed.), *Handbook of Research on Teachers Education* (569-598). NY: Macmillan.

Gitomer, D. (1999) "National Board for Professional Teaching Standards: Adolescence through Young Adulthood Science." Paper presented at the Annual International Conference of the Association for the Education of Teachers in Science, Austin TX.

Guo, C. J. (1995). Exploring science teacher education and certification from science education reform perspectives. *Proceedings of the 28th Seminar on Science and Technology: Science Education* (pp.91-102). Tokyo, Japan.

Henry, T. (1999). "Education chief: License teachers." *Iowa City Press Citizen.*

Medley, D. (1985) "Issues and Problems in the Validation of Teaching and Teacher Professional Behaviors." Paper presented at the Annual Meeting of the American Educational Research Association, Chicago IL.

Shulman, L. (1987). Assessment for teaching: An initiative for the profession. *Phi Delta Kappan*, **69**(1), 38-44.

Shulman, L. (1988). Speech presented at the meeting of the New Teacher Assessment and Support Consortium, Stanford, CA.

Walsh, E. (1999). "Dole: Teacher training lagging." *Iowa City Press Citizen.*

Yager, R. E. (1983). *Exemplary Programs in Physics, Chemistry, Biology and Earth Science*. Washington DC: NSTA.

Role of Student Outcome Data

學生表現的資料在師資檢定中所扮演的角色

James A. Shymansky

E. Desmond Lee Professor of Science Education
University of Missouri-St. Louis
St. Louis, MO, U.S.A.

摘　要

本文旨在論述在科學師資檢定中納入學生表現資料的必要性。首先簡要綜論支撐現今檢證制度的基本模式，進而指出，儘管學生、家長、社會大眾與學校行政人員對於學生的表現極感興趣，但是現行的檢定制度仍忽略了這方面的資料。其次討論針對教師行爲與教學措施對於學生表現所產生之影響進行周延研究的必要性。最後，本文呈現一個以學生表現來評鑑教師在職進修效能方案之實施結果，並描述美國某州致力於使用學生方面資料來檢定學校的做法，以支持在現行之整體或個別師資檢定的過程中，可納入學生學習成果資料的主張。