

## **The Effects of Professional Development: A View From Plato's Cave**

**Christopher H. Tienken**  
**Assistant Superintendent for the Monroe Township School District in**  
**Middlesex County, NJ and a part-time professor at the**  
**Rutgers Graduate School of Education**

**Charles M. Achilles**  
**Professor of Education Administration**  
**Eastern Michigan University and**  
**Seton Hall University**

---

### **Abstract**

**During the last 50 years federal legislation (e.g. NDEA, NCLB) and many staff-development consultants have emphasized Professional Development as the way to improve education, nearly to the exclusion of other approaches. Unfortunately, except for anecdotal, self-report information, there is little substitutive research to support those claims.**

**By re-analyzing the 2000 and 2003 NAEP database, the authors offer a contrarian view: The NAEP data offer little or no support for staff development as "improvement" when student test outcomes are the criterion measure.**

**The authors suggest that clear definitions of terms could improved research and evaluation on staff development. They provide minimum criteria for proclaiming staff development as "successful" and offer definitions to differentiate various approaches to staff development.**

---

## **The Effects of Professional Development: A View from Plato's Cave**

### **Introduction**

The No Child Left Behind (NCLB) legislation, most notably the “Highly Qualified Teacher” initiative, is a recent entry into federally mandated programs beginning with the National Defense Education Act (NDEA) in 1958 to propose teacher quality as the cure-all for perceived shortcomings in the education system almost to the exclusion of other improvement initiatives. Professional development (PD) has been a 50-year focus of many legislated programs, and yet it seems that each year brings additional cries about education's demise. Where are the assessments of prior PD efforts to show success?

The first section of this article refers to definitions for “professional development” and scientific-based research (SBR) that appear in Appendix A. The next section proposes criteria to judge the effectiveness of professional development, and some suggested criteria to judge evaluations of PD. The third section offers data from 2000 and 2003 National Assessment of Education Progress (NAEP) databases on the relationship of various teacher education and professional development structures to NAEP 4<sup>th</sup> and 8<sup>th</sup> grade mathematics and reading scores. The reader is encouraged to evaluate the NAEP data based on suggested criteria for effective PD and criteria for SBR from the NCLB as adopted by the National Research Council (NRC).

### **Importance of Clear Definitions of Terms**

Because consensus on research results, comprehension of prior research, and comparisons between present and other research endeavors require clarity of usage, definitions of terms as used in this article are provided in the Glossary (Appendix A). The terms are used consistently by the same authors in works that have been reported elsewhere (Covert, 2003; Tienken, 2003; Tienken & Achilles, 2003; Tienken & Achilles, 2005).

### **Effective Professional Development**

Educators and legislators should evaluate the emphasis on PD-type events in education against three criteria when they consider establishing rules, policy and/or support for professional development (Guskey, 1986; Tienken & Achilles, 2003):

1. Those who receive the professional development (e.g. teachers) demonstrate positive change in skills, knowledge, attitudes, and behaviors. The new professional practice will be sustained as part of the professional's armamentarium.

2. The desired changes and improvements should be measurable and observable in the short term and in the long term: They become the norm until better knowledge; skills (etc.) are available.
3. Because the ultimate beneficiary of PD should include the clients (i.e., students) the results of PD should lead directly to observable, measurable positive change in student outcomes on clearly defined criteria.

### **The Quality Paradigm**

Well-meaning educators have accepted the notion that teacher quality is the nucleus of education reform. School districts around the country spend large sums and considerable time each year trying to improve teacher quality. The NCLB Act will provide approximately 2.9 billion dollars during the 2005-2006 school year for increased PD to improve teacher quality (Viadero, 2005). An administrator's comment overheard at a professional development conference in New Jersey in 2005 captures the current paradigm: "If we could only fix the teachers." If the education leadership and local/state/national leaders accept teacher quality as the cause for shortcomings in the system and likewise accept PD as the solution, then they have avoided their share of any blame. Perceptions are reality in education. The perceptions are that education is broken, PD can "fix" teachers, and in time "fix" education. Concurrently, expectations for teachers continue to grow. A recent article provided an extensive list. Below is a partial account of a large set of expectations. Teachers should (Darling-Hammond, 2000):

- Have deep and flexible understanding of subject matter.
- Know how to represent ideas so they are accessible to others.
- Develop pedagogical content knowledge (as cited in Shulman, 1987).
- Understand child development
- Know how children and adolescents think and behave.
- Know how to help children grow in specific areas during specific life stages in specific contexts.
- Understand differences that come to pass due to culture, language, family circumstances, environment, prior experiences, and other student factors.
- Inquire sensitively and productively about children's understanding of content.
- Interpret curriculum through the eyes of children and deliver lessons that connect with the children
- Teach in ways that connect with students.
- Create learning situations in which students have opportunities to speak, write, and listen to each other's experiences.
- Reflect and analyze their practice in order to assess their effectiveness.

Consider this partial list of expectations in a class of 26 or more fourth-grade students. The teacher is preparing at least five distinct lessons a day along with monitoring homework. This could mean at least 130 papers to check, exclusive of homework assignments. Perhaps fourth-grade is not a good example. Consider a ninth-grade Algebra teacher who teaches five sections with 25 students in each section. For argument's sake, accept that the teacher needs minimum preparation to teach the same subject all day. During the day the teacher meets approximately 125 students. Class-work, plus homework, parent contact (125 sets of parents) and other high school duties leave little time to address expectations as expressed above. We are not making excuses for teachers, nor do we expect less than the list proposed by Darling-Hammond (2000). The conditions (such as organization for instruction) for teachers and students to "do" and to "know" must be in place as a precondition to improvement.

### **NAEP Data on Professional Development and Outcomes**

The National Assessment of Education Progress (NAEP), often called The Nation's Report Card on Education, relies on nationally representative yearly samples of the academic achievement of students in grades 4, 8, and 12. Assessments are conducted in reading, mathematics, writing, science, US history, civics, and the arts. The US Department of Education, under the direction of the Secretary of Education, supports the assessment project. The NAEP is guided by a 26-member board consisting of businesspeople, governors, legislators, state school officials, and select (political friends) members of the public. NAEP results are reported based on the national sample of students. The results are used by state, national, and others as a gauge of student performance (USDOE, 2003). NAEP databases provide indications about subject-matter achievement, instructional experiences, teacher characteristics, and school environment.

The NAEP data set has limitations. Results are aggregated to the national level creating large sample sizes. The large sample sizes can produce statistically significant results even when changes in scale scores are only two or three points in any direction. Although some results are statistically significant, they may not be educationally important. The lack of reported effect sizes is another NAEP limitation. Because of the crude representation of the NAEP data it is difficult to judge or validate the quality of the PD experiences reported or whether the "PD" is PD as defined here, in-service sessions, or staff development<sup>1</sup> (NCES, 2003).

### **Results and Analysis**

The results reported here originate from a search of the NAEP Data Tool, an online database. Results include data from the national sample of public school teachers and the achievement of public school students. The PD structures reviewed for this article were found in

the “Teacher Factors” section of the NAEP Data Tool. Results for the most common types of PD, as reported by Garet, et al, (2001) were searched for this article.

NAEP reports student achievement according to four categories. Student scores can range from 0-500. Table 1 contains the categories and just-proficient score for each category. For example, NAEP considers a 4<sup>th</sup> grade student scoring 206 on the reading portion “below basic” in reading achievement whereas a student scoring 240 is proficient. Problems with NAEP proficiency levels have been identified elsewhere (Bracey, 2001) and are discussed in the author’s notes.

Table 1

NAEP Achievement Categories and Cut-Scores for 4<sup>th</sup> and 8<sup>th</sup> Grades Reading and Mathematics

<u>Cut Scores for Achievement Categories</u>				
Achievement Category	4 <sup>th</sup> Reading	4 <sup>th</sup> Math	8 <sup>th</sup> Reading	8 <sup>th</sup> Math
Below Basic	<208	<214	<243	<262
Basic	≥208	≥214	≥243	≥262
Proficient	≥238	≥249	≥281	≥299
Advanced	≥268	≥282	≥323	≥333

**Graduate Work Beyond Initial Degree**

Data in Table 2 compare test-score results of students who had a teacher with a Masters degree with the scores of students who had a teacher with a Bachelors degree (4<sup>th</sup> and 8<sup>th</sup> grade reading scores). Although graduate work is not considered professional development as per the definition of terms, it is discussed because of the weight given it by policy makers. Average achievement for students of teachers with Masters degrees did not rank in the proficient range. The NAEP finding is similar to the results of Michel (2004) who reported that having a Masters degree showed little relationship to student achievement when compared to the socio-economic status (SES) of the community in which a student attends school and the rate of student mobility.

Table 2

The Educational Impact of a Teacher with a Masters Degree in Mathematics on NAEP 4<sup>th</sup> and 8<sup>th</sup> Grade Mathematics Scores (2000)

Degree	n <sup>a</sup>	4 <sup>th</sup> Grade			8 <sup>th</sup> Grade			
		Scale Score	(p value)	Diff.	n <sup>a</sup>	Scale Score	(p value)	Diff.
Bachelor	3696	223			4510	273		
Master	2294	225	.37	+2	3060	275	.51	+2

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

There was not a significant ( $p \geq .05$ ) difference in the mathematics scores in large samples of 4<sup>th</sup> or 8<sup>th</sup> grade students taught by teachers with a Masters degree in mathematics compared to students taught by teachers with Bachelors degrees. The NAEP data show what Darling-Hammond (2000) concluded: The percentage of teachers with full certification and a major in their teaching field is a stronger predictor of student achievement than is the percentage of teachers with Master's degrees. Goldhaber and Brewer (1997) reported that teachers' advanced-degree work in grades 8-10 did not translate into higher levels of student achievement across the board, but did influence achievement in specific subjects like mathematics and science. They concluded that there was not a consistent and positive correlation between teachers with advanced degrees in subjects other than those they teach and student achievement.

As data in Table 3 show, different graduate concentrations of teachers did show some mixed (statistically significant) positive results for students' grade-8 NAEP scores for the 2000 sample. Students of teachers who reported "yes" to studying in mathematics and mathematics education produced some higher scale-score results but not all students of teachers with advanced education produced significantly higher scale-score results ( $p \leq .05$ ).

Table 3

The Educational Impact of Types of Graduate Work on NAEP 8<sup>th</sup> Grade Mathematics Scores (2000)

Type of Grad. Work	n	Scale Scores <sup>a</sup> (Yes)	n	(No)	(p value)	Scale Score “Yes” Diff.
Major in Math	1213	278	6871	273	.04*	+5
Major in Educ.	1904	276	6144	273	.12	+3
Major in Math Educ.	1536	277	6548	273	.03*	+4
Major Other	1293	274	6791	273	.88	+1
Major Secondary Ed.	1213	276	6871	273	.21	+3
Major Elem. Educ.	728	266	7356	274	.02*	-8

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

<sup>a</sup> = n calculated using the NAEP total n multiplied by the row percentages.

\* = statistically significant result

A major in elementary education was associated with negative student test-score ( $p \leq .05$ ) results. Results are consistent with those reported by Darling-Hammond (2000), but note that large sample size may influence statistical significance. Effect sizes were not reported.

### **Time Spent Developing Professionally**

Teacher-reported time spent in PD seminars or workshops had no significant relationship to students' 4<sup>th</sup> grade mathematics scores (See Table 4). Based on student NAEP scale-score results, more teacher-reported time spent in PD was negatively related to achievement. Similar results can be found for 8<sup>th</sup> grade NAEP outcomes. Kennedy (1998) reported similar findings from a meta-analysis of 93 studies of the effect of professional development on student achievement in mathematics and science. She found longer time periods spent participating in a specific PD program decreased the likelihood that the programs retain their constancy. Table 4

raises interesting policy questions given that many states have professional development requirements. For example, beginning in 2000, New Jersey required teachers to accrue 100 hours of professional development every five years. Is requiring teachers to accrue hours a wise investment of public funds and teacher time? Does accruing PD hours meet the criteria of effective PD based on both observed performance change and student improved outcomes?

Table 4

Reported Hours Spent in Mathematics Education Seminars and /or Workshops (2000) Related to Student NAEP 4<sup>th</sup> Grade Mathematics Scale Scores (n=6014)

			<u>Reported Hours Spent in PD Seminars/Workshop (p value)</u>			
			None	< 6	6-15	16-35
<u>Hours</u>	<u>Scale Score</u>	n				
None	229	1263				
< 6	226	2285	p=.34	-	-	-
6-15	224	1323	p=.13	p=.42	-	-
16-35	225	722	p=.27	p=.68	p=.58	-
35+	225	421	p=.36	p=.82	p=.71	p=.88

**Effects of Different Structures**

As well as compiling time spent in PD, NAEP provides data on various forms of professional development (See Table 5). In the NAEP data, workshops and training sessions had the strongest relationship to student achievement ( $p \leq .00$ ). Interestingly, the NAEP data contradict findings on the ineffectiveness of traditional workshops and seminars as reported in research and evaluation studies, 1997-2003 (See Table 6) that support the 1983 work of Joyce and Showers that demonstrated one-day events were largely ineffective relative to classroom implementation of one-shot training content. Teachers did not include the content of the in-service activities in future lesson planning or implementation. Showers and Joyce (1996) confirmed that a low percentage of teachers implement what they encounter at in-service training

sessions. Covert (2003) showed that the content of even continuous workshop-type training, with some follow-up was both poorly implemented and almost never observed being used six months after the training, and not at all one year later. Indeed, Guskey and Sparks (1996; Sparks, 1995) called for a paradigm shift in staff development to get away from disconnected and isolated PD. Darling-Hammond (1997) echoed their request by calling for an end to “drive by” professional development.

Table 5

Relationship of Various Structures of Professional Development to Student 4<sup>th</sup> Grade NAEP 2003 Mathematics Scores.

Type of PD	n <sup>a</sup> (Yes)	Scale Score (Yes)	n <sup>a</sup> (No)	Scale Score (No)	(p value)	Scale Score Diff. (Yes)
Study Group	43,833	234	131,500	234	.88	0
Individual/ Collaborative Research	29,806	235	145,527	234	.52	+1
Peer Coaching	47,340	235	127,994	234	.03*	+1
Workshop	106,954	235	68,380	233	.00*	+2
Post Certification Coursework	29,806	235	145,527	234	.19	+1
Grad Major in Math Education	3140	230	144,445	235	.01*	-5
Conference or Prof. Assoc. Meeting	49,094	236	126,240	234	.00*	+2
Consult with Math Specialist	49,094	235	126,240	234	.18	+1
Math Teacher Network/ Collaborative	19,287	234	156,047	234	.27	0

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

\* = statistically significant result

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

Table 6  
Studies that Support Professional Development (PD) Without Empirical Evidence of Teacher Change or Student-Improvement Gains.

<u>Source/Focus/Design</u>	<u>Design/Outcome/Quote</u>
Guskey, T. R. (1997)	“Research needs to link PD and student learning.”
Haller, Brent & McNamara (1997) Advanced training in EdAd.	No measurable difference on “Effective Schools” indicators of advanced training.
Newmann et al. <u>AERJ</u> (2000) and <u>American J. of Education</u> . Broadly theoretic and conceptual paper. School capacity focus.	“The case for substantial investment in [PD] is vulnerable because of an absence of research that links specific forms of [PD] to changes in teacher learning and practice and to student achievement gains. . .” (p. 53).
Bodilly et al. (2003). Rand Evaluation of New American Schools (NAS) implementations	“...reforms such as NAS –including teacher-reported collaboration, professional development, and revised instructional practices—were <u>not</u> rated to student achievement...”
Garet et al. (2001). Focus on teacher learning.	National probability sample (n=1027). Teachers self-report on a survey. No evidence of student gains.
Covert, S. (2003) Review of research for a dissertation on PD. (E. Michigan U.)	“Given the lack of studies which demonstrate PD effects on teachers or student outcomes...” (p. 25). Argues for attention to theory to guide PD.
Guskey, T. R. (2003), Review of 13 lists of “effective PD.” Dearth of SBR work and evidence of outcomes. Mostly “surveys of opinions” (p. 749)	“...lists could be described as “research based.” But that research rarely includes rigorous investigations of the relationship between the noted improvements in instructional practice or learning outcomes.” (p. 749).

Note: Table adapted from Tienken & Achilles, 2005 (p.307), and used with permission of *Current Issues in School Leadership*.

The NAEP 8<sup>th</sup> grade results are similar to those for 4<sup>th</sup> grade students. (See Table 7). Regardless of the data quality, the consistency of these results should raise policy questions about PD to “fix” schools and about its cost in time and money.

Table 7

Relationship of Various Structures of Professional Development on 8<sup>th</sup> Grade NAEP Mathematics Scores 2003.

Type of PD	n <sup>a</sup> (Yes)	Scale Score (Yes)	n <sup>a</sup> (No)	Scale Score (No)	(p value)	Y-N Diff.
Study Group	54,779	277	66,955	278	.058	-1
Individual/ Collaborative Research	33,911	278	81,031	278	.64	0
Peer Coaching	59,676	278	62,112	277	.09	+1
Workshop Session	111,653	278	12,406	274	.00*	+4
Post Certification Coursework	46,148	279	75,293	277	.06	+2
Team-teaching	53,427	278	67,997	277	.057	+1
Math Consult	57,351	276	64,672	279	.00*	-3
Grade Major in Math Education	18,517	281	78,698	276	.00*	+5
Conference or Prof. Assoc. Meeting	82,107	279	40,441	275	.00*	+4
Consult with Math Specialist	57,350	276	64,672	279	.00*	-3
Act as Lead Teacher, Mentor, Trainer	45,959	280	78,238	276	.00 <sup>a</sup>	+4

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

\* = statistically significant result

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

Although 6 of the 11 PD structures reported statistically significant differences ( $p \leq .05$ ), positive results for PD participants did not exceed five scale-score points for the type of PD listed (See Table 8). The negative results did not exceed three scale-score points. The data suggest that PD has a limited relationship to student reading achievement. Students of teachers with Masters degrees did not have significantly higher 4<sup>th</sup> and 8<sup>th</sup> grade reading scores than did students of teachers with Bachelors degrees.

Table 8

The Educational Impact of Masters Degree in Reading on NAEP 4<sup>th</sup> and 8<sup>th</sup> Grade Reading Scores (2000)

Degree	4 <sup>th</sup> Grade			8 <sup>th</sup> Grade		
	Scale Score	n <sup>a</sup>	(p value)	Scale Score	n <sup>a</sup>	(p value)
Bachelor	211	5182	.33	257	7248	.24
Master	214			259		

Scores range from 0-500

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

Half, 4/8, of the PD formats examined for 4<sup>th</sup> grade reading had a statistically significant relationship to 4<sup>th</sup> grade students' reading scores (See Table 9). Positive scale-score gains did not exceed three points. The other half did not significant relationships.

Table 9

Relationship of Various Structures of Professional Development on 4<sup>th</sup> Grade 2003 NAEP Reading Scores.

Type of PD	n <sup>a</sup> (Yes)	Scale Score (Yes)	n <sup>a</sup> (No)	Scale Score (No)	(p value)	Y-N Diff.
Study Group	71,540	216	98,794	217	.15	-1
Individual/ Collaborative Research	51,100	217	119,234	217	.99	0
Peer Coaching	68,1334	216	102,200	217	.46	-1
Workshop Session	148,190	217	221,143	215	.00*	+2
Grad Major Reading/Language Arts	18,759	219	115,680	217	.01*	+2
Advice from Curriculum Specialist	78,354	216	91,980	217	.34	-1
Curric./Instr. Committee	86,870	218	83,464	215	.00*	+3
LA Teacher Network/ Collaborative	30,660	215	139,674	217	.00*	-2

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

\*= Statistically significant result

LA= Language Arts

Only one of the four structures related to improved reading scores on the 4<sup>th</sup> NAEP, demonstrated a positive relationship to 8<sup>th</sup> grade scores (See Table 10).

Table 10  
Impact of Various Forms of Professional Development on 8<sup>th</sup> Grade NAEP Reading Scores (2002 & 2003).

Type of PD	n <sup>a</sup> (Yes)	Scale Score (Yes)	n <sup>a</sup> (No)	Scale Score (No)	(p value)	Y-N Diff.
Study Group	55,872	263	65,589	263	.26	0
Individual/ Group Research	55,580	263	65,246	263	.23	0
Peer Coaching	63,397	263	58,518	262	.12	+1
Workshop Session	118,001	263	7,532	262	.38	+1
Grad Major Reading/Language Arts	22,675	257	69,158	259	.00*	-2
Post Cert. Coursework	54,584	263	66,732	263	.38	0

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

\*= Statistically significant result

The students of teachers who reported that they had influence over their professional development scored higher than did students of teachers who reported that they did not feel they had influence of their professional development (Table 11). The data suggest that teacher influence over their professional development is related to higher student achievement.

Table 11  
Does Teacher Influence Over the Content of Professional Development in Reading on 4<sup>th</sup> and 8<sup>th</sup> Grade Influence Scale Scores (2002)

Grade Level	n <sup>a</sup> (No)	Scale Score		n <sup>a</sup> (Yes)	Scale Score		(p value)	Y-N .Diff
		No	Yes		Yes	No		
4th	8,527	213		14,619	221		.00*	+8
8th	8,593	261		10,502	265		.00*	+4

Note: “Yes” or “No” refers to respondents who participated (or not) in the type of graduate work listed.

<sup>a</sup>= n calculated using the NAEP total n multiplied by the row percentages.

\*= Statistically significant result

### Conclusions

Overall, the NAEP results have many inconsistencies. Taken as a body of evidence, the professional development practices reported in the NAEP database did not have an observed scale-score or statistically significant relationship to improved student achievement. Statistically significant scale-score gains were small, generally five or fewer points, very modest given the points available (0-500) and the number of points needed to move up one category on the NAEP scale. Jacobs and Lefgren (2002) found similar results in their analysis of the impact of professional development on student achievement in the Chicago Public Schools. They identified that moderate increases in PD had no effect on students’ math and reading scores. Kennedy (1998) who determined there was a lack of a clear relationship between PD programs and improved student achievement. She called for careful testing of the claims being made about the ability of PD to reform the education system.

## Policy Implications

The NAEP data set represents a national sample of teachers. Given NAEP's nationwide reach, it is appropriate to examine and criticize the results for implications on professional development policy. Students of teachers with Masters degrees did not outperform students of teachers who lacked advanced-degree work. Students of teachers who had participated in various PD activities did not outperform students of teachers who did not participate. Although some statistical differences did appear, the positive educational significance of PD activities was limited to five or fewer scale-score points. This is a relatively modest gain, given that it takes between 30-37 points for students to move from the Basic category to the Proficient category. Some (5) of the PD activities had a statistically significant negative effect on student achievement at various levels and subjects.

Limitations exist when analyzing aggregated data for tens of thousands of participants. Some teachers and their students probably benefited a great deal from the PD experiences. But taken as a group, the overall relationship between reported PD and student achievement is weak. Policymakers should consider the lack of evidence supporting PD and advanced degrees when developing national, state, and local PD requirements and policy related to "teacher quality" and PD. Neither PD nor advanced degrees seem to influence student achievement when certification and licensure status are accounted for in the data. It has been reported that a teacher certified to teach the subject or area of his/her current assignment has a significant impact on student achievement when compared to a teacher teaching outside of his/her certificated area (Darling-Hammond, 2000). Given the amount of money spent by districts for college tuition reimbursement, PD costs, and pay incentives based on advanced work or PD, policymakers may want to examine research-based ways to use those funds better.

## Future Research

Researchers should examine the structure and content of PD activities presented in the NAEP data base to determine their quality. Does the self-report nature of the data-collection format provide an accurate database? A detailed investigation may provide more clues into the types of structures and activities that produce the greatest results for teachers and students.

## Closing Thoughts from Plato's Cave

As an analogy, consider Plato's Republic and its poignant image of The Cave. The cave image seems particularly timely given the shadowy obsession with professional development as the solution to improved teacher quality. *And Socrates said to Glaucon,*

*Imagine an underground chamber like a cave, with a long entrance open to the daylight and as wide as the cave. In this chamber are men who have been prisoners there since they were children, their legs and necks being so fastened that they can only look straight ahead and cannot turn their heads. Some way off, behind and higher up, a fire is burning, and between the fire and the prisoners and above them runs a road, in front of which a curtain-wall has been built, like the screen at puppet shows between the operators and their audience, above which they show their puppets... For, tell me, do you think our prisoners could see anything of themselves or their fellows except the shadows thrown by the fire on the wall of the cave opposite them? Would they not assume the shadows they saw were real things? Glaucon replied, "Inevitably."*

Plato continued,

*And so in every way they would believe that the shadows of the objects we mentioned were the whole truth. Glaucon replied, "Yes, inevitably." (Plato, Republic. 514a-515c. Translated by Kamtekar, 2003).*

Substitute the words "education leaders" or "administrators" for prisoners in the cave and replace the shadows in the cave with "teacher quality/professional development". Have the unquestioned shadows of teacher quality initiatives been cast for too long? Have leaders become intellectually anesthetized to research-based innovations for education reform in the rush to blame teachers and teacher education?

Teacher quality certainly is one factor in effective education of children. The overwhelming reliance upon PD as the primary education improvement vehicle with little or no SBR to support its effects needs review and critique. If PD is to be claimed as a superior solution, its quality must be judged by empirical strength.

## APPENDIX A

## Glossary of Terms

- 1) In-service training refers to one-time or short-term training, usually a specific workshop or large-group session to present information or a basic skill, easily learned, usually delivered primarily via one-way communication. The training may include job-embedded elements, but there typically is no follow-up.
- 2) Job-embedded staff or professional development is planned and continuous training that in education specifically emphasizes teaching or instructional skills and knowledge related to student outcomes. Teachers learn by doing. Examples include action research and evaluation, structured study groups, peer coaching, and mentoring.
- 3) Professional development (PD) is ongoing, planned, continuing education through which certified, qualified teachers and other education professionals improve skills, knowledge, and attitudes/dispositions related to assisting students achieve the goals of the organization (i.e., improved student performance and outcomes). A primary interest is to improve the professional's workplace performance and increase individual long-term value. Interaction and two-way communication are an integral part of the long-term effort. Note that PD, as defined here, excludes formal advanced work, such as for a degree or certification. Experience, a type of on-the-job training (OJT), is also excluded.
- 4) Staff development (SD) involves workshops, training and knowledge related to the workplace, and offered to both professional and support personnel. The focus may not be on classroom performance but rather on personal and job-related topics of interest and value to staff and to organization maintenance or health (e.g., retirement planning, first aid training, diversity training, conflict resolution, policies/procedures related to law, etc.). This often helps in the management and smooth operation of the organization.
- 5) Teacher Education is the highest level of education attained by the teacher. Advanced course work is included.
- 6) Scientifically-based Research (SBR) is research that involves the application of rigorous, systemic, and objective procedures to obtain reliable and valid knowledge relevant to education activities and programs. It includes research that:
  - a) Employs systemic, empirical methods that draw on observation or experiment;
  - b) Involves rigorous data analyses that are adequate to test the stated hypotheses and justify the general conclusions drawn;
  - c) Relies on measurements or observational methods that provide reliable and valid data across evaluators and observers, across multiple measurements and observations, and across studies by the same or different investigators;
  - d) Is evaluated using experimental or quasi-experimental designs;
  - e) Ensures that experimental studies are presented in sufficient detail and clarity to allow for replication or, at a minimum, offer the opportunity to build systemically on their findings, and;
  - f) Has been accepted by a peer-reviewed journal or approved by a panel or independent experts through a comparably rigorous, objective, and scientific review [ESEA, 2002, Section 9101 (37)]

### Author Notes

<sup>1</sup> Although we acknowledge the potential weaknesses of massaging large databases, using survey data, and reporting cohort data as trends, the NAEP is used as the “Nation’s Report Card.” The following quotes regarding the flaws in the NAEP achievement levels were taken from the 2002 Executive Summary NAEP Reading Report Card (USDOE, 2003):

“As provided by law, NCES, upon review of a congressionally mandated evaluation of NAEP, determined that achievement levels are to be used on a trial basis and should be interpreted with caution” (USDOE, p.XI).

“In 1993, the first of several congressionally mandated evaluations of the achievement level setting process concluded that the procedures used to set the achievement levels were flawed...In response to the evaluation and critiques, NAGB conducted an additional study of the 1992 reading achievement levels before deciding to use them for reporting the 1994 NAEP results. When reviewing the findings of this study, the National Academy of Education (NAE) panel expressed concern about what it saw as a confirmatory bias in the study and about the inability of the study to address the panel’s perception that the levels had been set too high.” (USDOE, p. 14).

The NAE panel summarized its concerns as follows:

“First, the potential instability of the levels may interfere with the accurate portrayal of trends... it is noteworthy that when American students performed very well on an international reading assessment, these results were discounted because these results were contradicted by poor performance against the possibly flawed NAEP reading achievement levels in the following year.” (USDOE, p. 14)

“Although new directions were presented and discussed (for setting achievement levels) a proven alternative to the current process has not yet been identified.” (USDOE, p. 15)

***“The most recent congressional mandated evaluation conducted by the National Academy of Sciences (NAS) relied on prior studies of achievement levels...The panel (NAS) concluded NAEP’s current achievement-level-setting-procedures remain fundamentally flawed. The judgment tasks are difficult and confusing; raters’ judgments of different item types are internally inconsistent; appropriate validity evidence for cut scores is lacking, and the process has produced unreasonable results.”***  
(USDOE, p.15)

## References

- Bodilly, S. J., Gill, B. P., Berends, M., Kirby, S. N., Dembowski, J. W. & Caulkins, J. P. (2003). Hard lessons learned from educational interventions. *Rand Review*, 27(1), 1-10. Retrieved 5/1/03 from <http://www.rand.org/publications/randreview/issues/spring2003/crashcourses.html>.
- Bracey, G. (2001). *The 11<sup>th</sup> Bracey report on the condition of public education*. *Phi Delta Kappan*, 83(2), 157-169.
- Covert, S. L. (2003). Transferring Professional Development in the Classroom. Unpublished EdD Dissertation. Eastern Michigan University, Ypsilanti, MI 48197 (pp. 16-17).
- Darling-Hammond, L. (1997). *The right to learn: A blueprint for creating schools that work*. San Francisco: Jossey-Bass Publishers.
- Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives*, 8(1).
- Garet, M.S., Porter, A.C., Desimone, L., Birman, B.F., & Yoon, K.S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, 38(4), 915-945.
- Goldhaber, D. & Brewer, D. (1997). Evaluating the effect of teacher degree level on educational performance. In W. Fowler (Ed.), *Developments in school finance, 1996* (pp. 197-210). Washington, DC: U.S. Department of Education. National Center for Education Statistics. (ED 409 634).
- Guskey, T.R. (1986, May). Staff development and the process of teacher change. *Educational Researcher*, 5-12.
- Guskey, T.R. (2003, June). What makes professional development effective? *Phi Delta Kappan*, 84(10), 748-750.
- Guskey, T.R. (1997) Research needs to link professional development and student learning. *The Journal of Staff Development*, 18(2).
- Guskey, T.R. & Sparks, D. (1996). Exploring the relationship between staff development and improvements in student learning, *Journal of Staff Development*, 17(4).
- Haller, E.J., Brent, B.O. & McNamara, J.H. (1997, November). Does graduate training in educational administration improve America's schools? *Phi Delta Kappan*, 79(3), 222-227.
- Jacobs, B.A. & Lefgren, L. (2002). *The impact of teacher training on student achievement: Quasi-experimental evidence from school reform efforts in Chicago*. Working Paper 8916. National Bureau of Economic Research.
- Joyce, B. & Showers, J. (1983). *Power in staff development through research on training*. Arlington, VA: Association for Supervision and Curriculum Development.
- Kamtekar, R. (2003). *Plato: The republic*. London: Penguin Books.
- Kennedy, M. M. (1998, April). *Form and substance in inservice teacher education*. Research Report from the National Institute For Science Education, University of Wisconsin, 1998.
- Michel, A. (2004). What is the relative influence of teacher educational attainment on student NJASK4 scores? Dissertation Abstracts International-A, 65/06 p.2173. (UMI No. AAT 3136106).

- National Center for Education Statistics (n.d.) NAEP 2000, 2003 national/public reading and mathematics assessments – Data Tool – Grade 4 and Grade 8: Teacher Factors weighted percentages and composite scale-score means (National/Public school). (Available online at <http://www.NCES.ed.gov/nationsreportcard/y25alm/almanac.shtml>).
- Newmann, F.M., King, M.B. & Youngs, P. (2000). *Professional development that addresses school capacity: Lessons from urban elementary schools*. Paper presented at AERA, April 28, 2000.
- No Child Left Behind (NCLB) Act. PL 107-110 (Signed 1/8/02). Amended the Elementary and Secondary Education Act (ESEA), PL 89-10.
- Showers, B. & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership*, 53(6), 12-16.
- Sparks, D. (1995, Winter). A paradigm shift in staff development. *Professional Staff Development: The ERIC Review*, 3(3), 2-4.
- Tienken, C.H. (2003). Tienken, C.H. (2003). *The effect of staff development in the use of scoring rubrics and reflective questioning strategies on fourth-grade students' narrative writing performance*. UMI No. AAT 3081032
- Tienken, C.H., & Achilles, C.M. (2003). Changing teacher behavior and improving student writing. *Planning and Changing*, 34(3-4), 153-168.
- Tienken, C.H., & Achilles, C.M. (2005). *Staff development and school outcomes*. In L. Hughes, (ed.) *Current Issues in School Leadership*. Mahwah, NJ: Erlbaum.
- Viadero, D. (2005, July 27). Pressure builds for effective staff training. *Education Week*, pp. 1, 18-19, 21.
- United States Department of Education. Institute of Education Sciences. National Center for Education Statistics, (2003). *The nation's report card: Reading 2002*, NCES 2003-521, by W.S.Grigg, M.C. Daane, and J.R. Campbell. Washington, DC, 2003.