

Learning from the World: Achieving More by Doing Less

Countries that score higher in international comparisons than does the U.S. also require less time in school, assign less homework, and use less high-tech gadgetry. Mr. Baines argues that maybe it is time we learned from them.

BY LAWRENCE BAINES

AT THIS moment, in school districts throughout the United States, initiatives are being launched to extend the school day, increase homework, integrate technology, and require more high-stakes testing. The assumption underlying these initiatives is that more and more — more time in school, more homework, more technology, and more high-stakes testing — will produce smarter, better-prepared students who, in turn, will help guide the nation through the tumultuous and uncertain 21st century.

To realize the ideal of an educated, productive citizenry, however, many countries around the world are employing radically different approaches. Instead of executing a strategy of more and more, some countries have decided to educate their young people by doing less. Because the test scores of students from these countries routinely eclipse the scores posted by American students in two international comparisons of student

■ LAWRENCE BAINES is a professor of education at the University of Toledo, Toledo, Ohio.

achievement — Trends in International Math and Science Study (TIMSS) and Programme of International Student Achievement (PISA) — an investigation of educational practices in higher-achieving countries might prove instructive. Four areas where the policy and practice in high-achieving countries run counter to current practice and policy in the U.S. are as follows: 1) time spent at school, 2) homework, 3) technology, and 4) schools as agents of social change.

TIME SPENT IN SCHOOL

Students in public schools in most countries in Western Europe, Canada, Mexico, Korea, Japan, and Singapore — all members of the Organisation for Economic Co-operation and Development (OECD) — spend an average of 701 hours per year in school. In Finland, where students have scored near the top in international comparisons of achievement for a number of years, students spend only 600 hours in school. In the United States, by contrast, children go to school for six or more hours per day, five days per week, for approximately 185 days spread over a period of nine or 10 months.

The average time spent at school in the U.S. totals over 1,100 hours, almost double that of children in Finland. By the time children reach the age of 14 in Finland, they will have gone to school for 2,500 fewer hours than students in America (the equivalent of two to four years of schooling). Despite much longer school days, American students routinely score 10% to 20% lower than Finnish students on international tests of achievement.

Experimental studies have repeatedly found no correlation between time spent at school and levels of achievement.¹ Of course, as any teacher in American public schools can attest, time at school is often wasted on performing nonteaching tasks, organizing paperwork, maintaining discipline, and keeping students “busy.” Some of the more prestigious private secondary schools in America schedule classes in the fashion of universities — 90-minute periods that meet twice each week, with one day a week set aside for advising and one-on-one tutoring. If such a schedule were adopted in public high schools, for example, total instructional hours in America would drop sharply. But such a transformation would mean a departure from the traditional schedule and a retreat from the daily array of “professional development opportunities” such as hall duty, lunch supervision, bus detail, parking lot patrol, and detention hall supervision.

HOMEWORK

As with instructional hours spent in school, America also leads the world in assigning homework — a whopping 140 minutes per week in mathematics for secondary students. Despite this extra workload, American students are renowned for posting mediocre scores on math tests. For example, the average score for an eighth-grade American student on the mathematics portion of the TIMSS in 2003 was 502. In contrast, the average Korean eighth-grader scored 584. While many Americans may suppose that Korean teachers require more from their students, in actuality, Korean teachers assign *20 minutes less homework* per week than their American counterparts. Apparently, Korean students are learning more mathematics by doing less homework.

This should not be all that surprising. As a rule, time spent doing homework will be unconnected to academic achievement if the time is not spent productively. Because most American teachers tend to assign worksheets and exercises from textbooks for homework, a student’s level of engagement during the long evening hours of working at home may be less than optimal. Although much has been written about academic learn-

ing time (the time students are genuinely engaged in learning), many teachers are still more concerned with “keeping up” than with making learning interesting or relevant for their students. Obviously, as teacher salaries are increasingly tied to students’ performance on tests, the urge to “cover the curriculum” to be tested is understandable. However, lack of engagement inevitably leads to apathy, frustration, and boredom.

In examining homework policies around the world, researchers have concluded, “The relationship between national patterns of homework and national achievement suggests that . . . more homework may actually undermine national achievement.”² Many bleary-eyed American students would wholeheartedly agree.

TECHNOLOGY

A study of the integration of technology into American classrooms over the past century reveals that claims for new paths to achievement come as a matter of course with the development of new machines. In the past, some researchers have claimed academic gains associated with the use of film, radio, the tape recorder, videotape, television, and even the overhead projector. Apparently, after the novelty of a machine fades, so do claims that interactions with it will yield dramatic gains in achievement. For example, few researchers anymore would contend that an overhead projector enhances student achievement through the sheer power of its technology. Yet many schools in America have spent billions of dollars over the past 20 years under the illusion that providing students with access to computers and the Internet would somehow enhance achievement. While the universe of knowledge available via the Internet is indisputably vast, schools have been forced to restrict student access because too many websites feature pornography, ultra-violent images, or other material unsuitable for children. As a result, if they are used in schools at all, computers have taken on the role formerly occupied by a multivolume set of encyclopedias — a storehouse of concise, neatly categorized information used once or twice per year for research projects.

Undeniably, having access to the latest technologies is preferable to being relegated to a barren one-room schoolhouse with only a small, cracked chalkboard. However, technologies come with a bundle of benefits and tradeoffs. Ten years ago, the reason some high schools and universities began requiring students to come to class with laptops is that administrators believed laptops would enhance student achievement. Ten years later, the reason these same high schools and universities have stopped requiring laptops is that no evidence

has surfaced to substantiate that they made any difference.

In the 2003 administration of PISA, the factor most strongly associated with high scores on reading, problem solving, and mathematics was not the presence or absence of technology, but the number of books to which a student had access. Across categories of race, gender, and nationality, the more books present in the home, the higher a student's level of achievement.

Unfortunately, in most American schools today, books are handled as if they were artifacts from a museum. Consider the following policies now enforced in many schools:

- Students are often forbidden to take books (even textbooks) home.
- If students are allowed to take books home, no more than one may be checked out of the library, and it may be checked out for only a short duration.
- Books should be used with care (students may not write in them).

School libraries, once repositories for books, have morphed into multifunctional media centers. As a result, budgets for print materials have been reduced in order to keep the computers running. Although school libraries might serve as the sole access point for books in a particular community, libraries in high-poverty urban and rural areas may have precious few books to lend. In addition, school libraries in America usually close soon after the dismissal bell, so that students, parents, and members of the community have no time to browse the shelves or simply sit down and read. Bookless homes remain bookless homes.

In most OECD countries books are not treated as artifacts but are given to students to use as they wish. They can take them home, share them, and — believe it or not — scribble notes in the margins without penalty.

SCHOOLS AS AGENTS OF SOCIAL CHANGE

Perhaps only in America could a strict regimen of standardized testing be considered an antidote to the social problems of the poor and disenfranchised. But No Child Left Behind gained widespread, bipartisan political support by using precisely this logic. While the federal and state governments have focused upon the establishment of school-based initiatives — setting curricular standards, specifying performance outcomes, and integrating technology — other countries have taken a broader approach to social problems. Perhaps leaders of those countries are more familiar with the research that substantiates that differences in academic achieve-

ment are more attributable to differences in social background than to variations in standardized testing.³


Three dubious distinctions characterize America's poorest students: most hail from one- or no-parent households, they are the least healthy children in the country, and they score at the very bottom on achievement tests. On international achievement tests, more than one in four American students score at the lowest possible level. In Korea, only 9.6% of students score at the lowest tier; in Finland, only 6.8%.

The poverty rate in Finland is 5%, in Korea it is 15%, and in America, it is 12%. From this information, we can infer that America not only is doing an inadequate job of educating students in poverty but also is failing with significant numbers of the nonpoor. In recent decades, underachievement in America has been wholly perceived as a "school problem," and solutions have focused solely on interactions with students during school hours. The latest thinking in the United States has not been directed toward creating more family-friendly policies (such as the Canadian and European tax incentives for stay-at-home parents) or broader social initiatives, but toward putting in place more rigorous and frequent testing. A kid can try to hug a test, but the test will never hug back.

An examination of scores on standardized tests in the United States over the past 50 years reveals no discernible change in student achievement despite myriad efforts at reform. The initiatives of an extended school day, more homework, increased technology, and vigorous standardized testing, in vogue for decades, have done little to enhance achievement, promote positive attitudes, or foster good citizenship. Perhaps it is time to learn from the world, to stop thinking in terms of more and more, and consider what might be achieved by doing less.

1. Charles W. Fisher and David C. Berliner, eds., *Perspectives on Instructional Time* (New York: Longman, 1985); and Rita Mulholland and Michelle Cepello, "What Teacher Candidates Need to Know About Academic Learning Time," *International Journal of Special Education*, vol. 21, no. 2, 2006, pp. 63-73.

2. David Baker and Gerald Letendre, *National Differences, Global Similarities: World Culture and the Future of Schooling* (Stanford, Calif.: Stanford University Press, 2005), p. 130.

3. See Rosie McNiece, Penelope Bidgood, and Peter Soan, "An Investigation into Using National Longitudinal Studies to Examine Trends in Educational Attainment and Development," *Educational Research*, vol. 46, 2004, pp. 119-36; Jerry Westermeyer, "Predictors and Characteristics of Erikson's Life Cycle Model Among Men: A 32-year Longitudinal Study," *International Journal of Aging and Human Development*, vol. 58, 2004, pp. 29-48; and Betty Hart and Todd Risley, *Meaningful Differences in the Everyday Experience of Young American Children* (Baltimore: Brookes Publishing, 1995). 

Are International Tests Worth Anything?

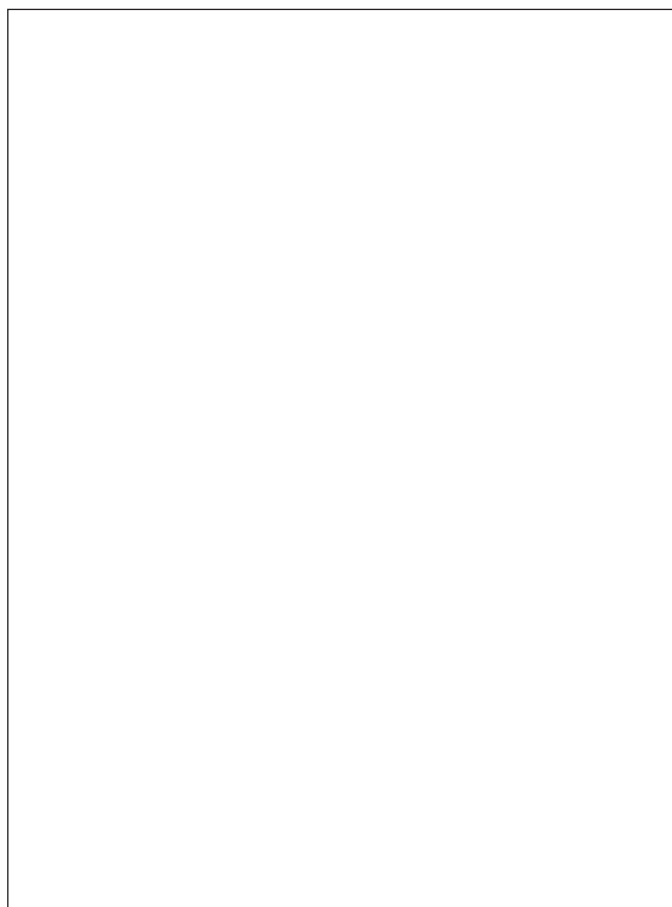
Do the U.S. rankings on international achievement tests signal doom for the country's future standing in the world? Mr. Baker sets out to answer that question by looking beyond the test scores to other dimensions on which nations can be compared.

BY KEITH BAKER

THE IDEA THAT America was being harmed because our schools were not keeping up with those in other advanced nations emerged after Sputnik in 1957, took a firm hold on education policy when *A Nation at Risk* appeared in 1983, and continues today. Policy makers justify this concern by pointing to evidence showing that, for individuals within the U.S., higher test scores predict a number of important life advantages, such as going on to college and making more money as an adult. From this they extrapolate that higher *national* test scores correlate with global success. The origins of the notion that education is crucial to the nation date back to the Founding Fathers, especially Jefferson, who held that a well-educated citizenry was the foundation of a nation's, especially a democracy's, success in the world.

Since Sputnik, the evidence driving worries about the performance of U.S. schools has come primarily from a series of international achievement testing programs that started in 1964 with the First International Mathematics Study (FIMS). This was followed by the Second International Mathematics Study (SIMS), the Third International Mathematics and Science Study (TIMSS), and, most recently, the Programme for International Student Assessment (PISA).

In this article I will show that for the U.S. and for the



top dozen or so most-advanced nations in the world, standings in the league tables of international tests are worthless. There is no association between test scores and national success, and, contrary to one of the major beliefs driving U.S. education policy for nearly half a century, international test scores are nothing to be concerned about. America's schools are doing just fine on the world scene.

BACKGROUND

When policy makers and politicians infer that the same relationship holds *between* nations as is found *with-*

■ *KEITH BAKER is retired as a researcher for the U.S. Department of Education and now lives in Utah.*

in nations, they commit the logical error known as the ecological correlation fallacy. Evidence of the effects of education within nations does not transfer to differences among nations.

To see the ecological fallacy at work, picture fans doing “the wave” at a football stadium. Watching only

he said to look at “life, liberty, and the pursuit of happiness.” To find out how the FIMS nations are doing on life, liberty, and the pursuit of happiness, I looked at seven indicators of national success. I related them to FIMS scores.

Wealth. First, and perhaps most important to a na-

Jefferson told us where to look to see if a nation is a success. He did not say to look at test scores. Instead, he said to look at “life, liberty, and the pursuit of happiness.”

the up and down movements of individual “citizens” of “Stadium Nation,” however, tells us nothing about the direction in which the wave circles the stadium, its “national” movement. If we had two such stadiums side by side, and our view from the Goodyear Blimp showed one wave circling to the left and the other circling to the right, neither wave nor both would tell us how the citizens are moving. Going down into the crowd and watching citizens move up and down tells us nothing about how the wave appears from the blimp — or what is going on in the neighboring stadium. Likewise, the effects of high test scores on the individuals within a nation tell us nothing about the relationship of those test scores to national success.

The mathematics of the ecological correlation fallacy is a proof that generalizing from the relationship between variables at the individual level to larger aggregate levels, such as nations, is indeterminate.¹ That is, maybe the generalization holds, maybe it doesn't. Therefore, when such a generalization is made, we must treat it as a hypothesis, never as established fact, until it has been confirmed at the level of nations. Only then is it wise to act on the hypothesis.

FIMS

To see if the leap from within-nation results to between-nation results is justifiable, I looked at how well test scores on FIMS, the first international comparison study, predicted national success in the first half-decade of the 21st century. FIMS was administered in 1964 to samples of 12-year-olds in 11 nations. Today's world is largely a world created and operated by the now 55-year-old FIMS generation. If there is a connection between high test scores and national success, it will show up in looking at how well the 1964 FIMS scores predicted where nations are today. Among the 11 FIMS nations, the U.S. finished second to last (ahead of Sweden).

Jefferson told us where to look to see if a nation is a success. He did not say to look at test scores. Instead,

tion, is the creation of wealth. The best measure of generating wealth is per-capita GDP adjusted for cost of living differences, or purchasing power parity (PPP-GDP). The wealth of nations scoring higher than the U.S. on FIMS averaged 73% of the per-capita income in the U.S. in 2002.² FIMS scores in 1964 correlate at $r = -0.48$ with 2002 PPP-GDP. In short, the higher a nation's test score 40 years ago, the worse its economic performance on this measure of national wealth — the opposite of what the Chicken Littles raising the alarm over the poor test scores of U.S. children claimed would happen.

Rate of growth. One can argue that since the U.S. had a big post-WW II economic lead over the rest of the world, the rate of economic growth is at least as important as GDP as an indicator of national achievement. The nations that scored better than the U.S. in 1964 had an average economic growth rate for the decade 1992-2002 of 2.5%; the growth rate for the U.S. during that decade was 3.3%. The average economic growth rate for the decade 1992-2002 correlates with FIMS at $r = -0.24$. Like the generation of wealth, the rate of economic growth for nations improved as test scores dropped.

Productivity. GDP is a measure of a nation's total economic output. Productivity — GDP per hour worked — might be a better measure of a nation's economic success than GDP, since nations differ in the number of hours a year that the average worker spends at work.³ There is no relationship between FIMS scores and hourly output, $r = -.03$. In 2004, the average hourly output of those nations that outscored the U.S. in 1964 was 3.4% lower than U.S. productivity, though the three nations with higher hourly output all had higher test scores than the U.S. However, on the PISA test, which I discuss below, none of these three nations scored higher than the U.S.

Quality of life. Some argue that GDP is too simple a measure of national goals, that there is more to the good life than money. The United Nation's Quality of Life Index addresses this concern. Those who worry about

international test score standings base their worries on an assumption that high-scoring nations are more successful at doing the things nations should be doing, and offering a good quality of life to citizens is one of those things. But again, they are wrong. The average rank on the Quality of Life Index for nations that scored above the U.S. on FIMS was 10.8. The U.S. ranked seventh (lower numbers are better). FIMS scores correlated with Quality of Life at $r = -0.57$.

Livability. An alternative to the Quality of Life Index, the Most Livable Countries Index, shows that six of the nine countries that scored higher on FIMS than the U.S. are worse places to live. Livability correlates with FIMS scores at $r = -.49$.

Democracy. Jefferson also held that a well-educated citizenry is necessary for good democratic government. On the Economy Intelligence Unit's Index of Democracy, those nations that scored below the median on FIMS have a higher average rank on achieving democracy (9.8) than do the nations that scored above the median (18). Once again, the U.S. scored higher on attaining democracy than did nations with higher 1964 test scores.

Creativity. A good school system should foster creativity. The number of patents issued in 2004 is one indicator of how creative the generation of students tested in 1964 turned out to be. The average number of patents per million people for the nations with FIMS scores higher than the U.S. is 127. America clobbered the world on creativity, with 326 patents per million people. However, FIMS scores do correlate with the number of patents issued: $r = .13$ with the U.S. and $r = .49$ without the U.S.

THE FIMS PREDICTIONS

The hypothesis that low scores on international tests lead to national disaster, or at least inferior performance as a nation, predicts that the nine nations scoring higher than the U.S. on FIMS should outperform the U.S. on measures of national success. If the hypothesis is correct, nations with higher FIMS scores than the U.S. should be doing better than the U.S. on the seven indicators of national success in a world that is now run by the FIMS generation.

What's the bottom line? Altogether, there are 61 possible comparisons between the U.S. and a higher-scoring nation across the seven indicators. According to the hypothesis, 100% of these comparisons — or, at the very least, an impressive majority — will show the U.S. doing a worse job than the higher-scoring nations. In fact, the U.S. comes out on top in 74% of the comparisons.

In the face of such evidence, we can do more than reject the widely held hypothesis that high test scores lead to national success in the future. We can also hypothesize that high test scores are damaging to nations. That the U.S. comes out on top in national success in 74% of the comparisons with higher-scoring nations is statistically significant ($p < .0001$, binomial test).

Sputnik went up, and America's test scores went down compared to other advanced nations. But there was no need to panic or to proclaim, as so many did, that America's schools were in a crisis of poor performance. In looking at the world four decades after FIMS, the U.S. turned out more than just okay compared to nations with higher test scores. No matter how you look at it, high test scores in 1964 were not positive predictors of how the world would turn out. At best, international test scores are useless and may well be harbingers of failure, rather than success.

The logic of the ecological correlation fallacy warned that jumping to policy conclusions from international tests was a dubious enterprise. Since this logical fallacy was known by 1950, there was no excuse for policy makers at the time of FIMS or at any time since to proclaim the existence of problems in U.S. schools because some other countries posted higher test scores.

PISA

PISA, a second and more recent international testing program, included more than twice as many nations ($n = 27$) as FIMS ($n = 11$).⁴ Like FIMS, PISA shows

no connection between high test scores and how well a nation does at achieving wealth, growth, democracy, or quality of life for its people.

On these indicators of success, the nations that scored at the PISA average generally outperformed those scoring either above or below average. For example, per-capita GDP was \$22,495 for the 11 nations scoring above average, \$34,414 for the five average nations, and \$16,375 for the 11 below-average nations. The same pattern holds for quality of life, democracy, and creativity as measured by patents.

International comparisons on many factors show that Norway is the best place in the world to live, and, like the U.S., Norway scored right at the PISA average. Mediocre test scores correlate with better, more successful countries than do top scores (or lower scores). Mediocrity in test scores is, for nations, a good thing! This finding is highly counterintuitive. Why should it be so?

CONCLUSIONS

Among high-scoring nations, a certain level of educational attainment, as reflected in test scores, provides a platform for launching national success, but once that platform is reached, other factors become more important than further gains in test scores. Indeed, once the platform is reached, it may be bad policy to pursue further gains in test scores because focusing on the scores diverts attention, effort, and resources away from other factors that are more important determinants of national success.

The fixation on test scores has so dominated policy that little attention has been paid to finding out what makes America's schools the best in the world with regard to international economic competition. But a recent conversation I had with a Swede now living in Los Angeles seems to point in the right direction. He holds a high position in a bioscience company and has lived in 10 different nations. He told me, "There is no doubt that graduates of European high schools know a lot more than American grads, but I prefer my kids go to school in America because Americans acquire a spirit that the other countries lack." Other anecdotal sources suggest this "spirit" involves ambition, inquisitiveness, independence, and perhaps most important, the absence of a fixation on testing and test scores.

"Imagination is more important than knowledge," Einstein observed, and this principle applies to physics, to science, to what makes a modern economy succeed, and to what schools should teach. As to the relative importance of test scores and that "spirit" that U.S. schools

seem to cultivate better than those anywhere else, Einstein again is on the mark: "The true sign of intelligence is not knowledge but imagination." How America's schools beat the rest of the world in developing imagination may not yet be clear, but that — rather than raising test scores — should be the focus of both policy and research.

For more than a quarter of a century, the American public has been barraged by politicians and pundits claiming that America's schools are disaster zones because we are not at or near the top of the league standings in test scores. This claim is flat out wrong. It is wrong in fact, and it is wrong in theory. For almost 40 years, those who believe this fallacious theory have been leading the nation down the wrong path in education policy. It turns out that the elementary teachers who have said all along that there is more to education than what is reflected in test scores were right and the "experts" were wrong.

Trying to raise America's test scores in comparison to those of other nations is worse than pointless. It looks to be harmful, for the only way to do it is to divert time, energy, skill, and resources away from those other factors that propel the U.S. to the top of the heap on everything that matters: life, liberty, and the pursuit of happiness.

The fixation with test scores also harms the nation by diverting time, attention, and resources away from America's real educational problems, such as too few minorities graduating from college, the run-down schools in the nation's inner cities, misdirected parental interference in schools, and the lack of parental and administrative support for teachers. There are more, of course, but nowhere on the list of our educational problems should we ever again find worries over our performance on tests compared to that of other nations.

1. Nor, as I showed in my article, "Yes, Throw Money at Schools," *Phi Delta Kappan*, April 1991, pp. 628-32, do relationships found among larger aggregates, such as nations or schools, generalize to individuals.

2. Although I began the analysis with statistical tests of the hypothesis that high scores lead to high national success in the future, these results are not presented, since statistical testing turned out to be unnecessary because the hypothesis underlying the policy concern about America's poor test scores is a directional hypothesis. Being a directional hypothesis, it is sufficient to conclude that the null hypothesis cannot be rejected when we find negative relationships between the independent and dependent variables.

3. Hourly output is available only for the OECD nations, which include nine of the 11 FIMS nations.

4. Using PISA scores to examine the effects of high test scores on national success has both problems and advantages. The main problem is that the students tested in 2000 have not had time to have much effect on their nations. PISA's advantage is that it included many more countries than FIMS. Inspection of the international test score data suggests it is reasonable to assume that national test scores are stable over time, a conclusion confirmed by the fact that similar patterns show up for PISA and for FIMS.



File Name and Bibliographic Information

k0710bai.pdf

Lawrence Baines, Learning from the World: Achieving More Doing Less, Vol. 89, No. 02, October 2007, pp. 98-100.

Copyright Notice

Phi Delta Kappa International, Inc., holds copyright to this article, which may be reproduced or otherwise used only in accordance with U.S. law governing fair use. MULTIPLE copies, in print and electronic formats, may not be made or distributed without express permission from Phi Delta Kappa International, Inc. All rights reserved.

Note that photographs, artwork, advertising, and other elements to which Phi Delta Kappa does not hold copyright may have been removed from these pages.

Please fax permission requests to the attention of KAPPAN Permissions Editor at 812/339-0018 or e-mail permission requests to kappan@pdkintl.org.

For further information, contact:

Phi Delta Kappa International, Inc.
408 N. Union St.
P.O. Box 789
Bloomington, Indiana 47402-0789
812/339-1156 Phone
800/766-1156 Tollfree
812/339-0018 Fax

<http://www.pdkintl.org>

Find more articles using PDK's Publication Archives Search at
<http://www.pdkintl.org/search.htm>.