Program on Education Policy and Governance Working Papers Series

Producing superstars for the economic Mundial: The Mexican Predicament with quality of education

Lant Pritchett and Martina Viarengo
Harvard University, John F. Kennedy School of Government

PEPG 09-01

January 2009

Program on Education Policy and Governance
Harvard Kennedy School
79 JFK Street, Taubman 304
Cambridge, MA 02138
Tel: 617-495-7976 Fax: 617-496-4428
www.ksg.harvard.edu/pepg/

Lant Pritchett is Professor of the Practice of Economic Development at the Kennedy School of Government at Harvard University.

Martina Viarengo is the 2008-2009 postdoctoral fellow at the Program on Education Policy Governance at the Kennedy School of Government at Harvard University.

Producing superstars for the economic Mundial: The Mexican Predicament with quality of education

Lant Pritchett and Martina Viarengo

Harvard University, John F. Kennedy School of Government

Abstract The question of how to build the capabilities to both initiate a resurgence of growth and facilitate Mexico's transition into a broader set of growth enhancing industries and activities is pressing. In this regard it seems important to understand the quality of the skills of the labor force. Moreover, in increasingly knowledge based economies it is not just the skills of the typical worker than matter, but also the skills of the most highly skilled. While everyone is aware of the lagging performance of Mexico on internationally comparable examinations like the PISA, what has been less explored is the consequence of that for the *absolute* number of *very highly skilled.* We examine how many students Mexico produces per year above the "high international benchmark" of the PISA in mathematics. While the calculations are somewhat crude and only indicative, our estimates are that Mexico produces only between 3,500 and 6,000 students per year above the high international benchmark (of a cohort of roughly 2 million). In spite of educational performance that is widely lamented within the USA, it produces a quarter of a million, Korea 125,000 and even India, who in general has much worse performance on average, produces over 100,000 high performance in math students per year. The issue is not about math per se, this is just an illustration and we feel similar findings would hold in other domains. The consequences of the dearth of globally competitive human capital are explored, with an emphasis on the rise of *super star* phenomena in labor markets (best documented in the USA). Finally, we explore the educational policies that one might consider to focus on the upper tail of performance, which are at odds with much of the "quality" focus of typical educational policies which are often remedial and focused on the lower, not upper tail of performance.

Forthcoming in the World Economic Forum Report 2008

Introduction

While we are leery of comparing economic "competition" with athletic "competition" we want to use the metaphor of the Mundial of *futbol* to motivate three points. First, in a Mundial global competitiveness matters as it stacks players of different countries up against each other on a level playing field. While victory in any given league is relative, one can be the best in a local league without being very good. Second, in the Mundial it is not the *average* quality of the futbol players that matters, it is the very upper tail—the best of the best. The quality of the players in the *upper tail* depends not just on the average of the distribution, but also how that distribution is shaped—variance and is it skewed towards the upper, high performance, tail. Third, the simple math of order statistics suggests that the absolute quality of the players depends in part on the size of the pool from which they are drawn—in a random draw of standard normal variates the best of 100 will be around 2.5, but the best of a million will be around 4.9. Every boy in Mexico believes that they are in the running to be on the Mundial selection, but can the same be said for the economic Mundial, does every child really believe they have a shot to rise to be the best of the best economically?

1) Quality of Mexican Education in an International Context

The low rates of schooling enrollments and educational attainment of the Mexican population are widely acknowledged: out of every 100 students entering primary school, 68 completed basic compulsory education whereas only 35 graduate from upper secondary[1] (Santibañez et al., 2005, p.17). Only 8.5 percent of the population aged 18 and older held a bachelor's degree in 2003 (Villa and Pacheco, 2004 and Santibañez et al., 2005). However, the more recent economic research has shown that what really plays a role in determining a country's competitiveness and economic

[1] This is calculated by taking completion indicators of the Secretaria of Educación, 2003. There is no data on how many of the upper secondary graduates enter college.

growth is the level of cognitive skills of the labor force as opposed to its schooling attainment[2]. That is the "quality" of education is more important than the mere expansion of schooling opportunities (i.e., the "quantity" of education, measured in terms of years of schooling). But this research also suggests that it is not just the average quality, but quality at the top matters as well. In this section we are going to examine how Mexico performs vis-à-vis of the world in terms of standards of education at secondary, both on average and at the top, and higher education levels.

1a) Quality at Secondary Level, Average and Upper Tail

In order to examine the quality of Mexican secondary education we compare it internationally. In this regard, the PISA test allows a consistent comparison across countries as it provides comparable measures of the knowledge acquired by 15-year old students who are close to the end of compulsory schooling in the majority of the participating countries. Moreover, the test is not curriculum-based (as is, for instance, the TIMSS) its focus being on "what people can do", rather than "what people know"—and while there are arguments for both types of tests, for our purposes the PISA raises less questions about the results being due to differences in curricular content.

The PISA aims to be effective in evaluating how education systems prepare students for life in a larger context. In figure 1, we compare Mexico's distribution of test scores in mathematics[3] to the United States and Korea. Test scores have been standardized, so that the OECD wide mean is equal to 500 and OECD wide student standard deviation equal to 100. The builders of the PISA test also

[2] The latest paper is Hanushek and Woessmann (2008). The literature about the role of schooling attainment (aggregated into measures of "schooling capital") on economic growth has shown mixed results, with the contribution to aggregate output at best equal to the contribution one would have expected from the microeconomic returns (Pritchett, 2006).

[3] We show here the figure for mathematics only as it appears to be the most readily comparable subject across countries. Analogous figures for science, reading and problem solving are available from the authors upon request.

distinguish six levels of proficiency[4]. An example of the sample questions in mathematics is provided in table A2.

The average Mexican student is performing significantly below the OECD average, far from his Korean and American counterparts. The average Mexican student scores below Turkish and Thai students[5]. That is, the average Mexican student achieves only proficiency Level One in mathematics which means they cannot do more than "carry out routine procedures according to direct instructions in explicit situations"[6]. We do not dwell on the average score as that Mexico is lagging other nations on average in learning competencies in fields like mathematics and science is not news. We focus on two features that the comparison of averages across countries does not highlight: the upper tail and the *absolute number* (not percent) of high performers.

A test score higher than 625 is considered to be "advanced" by PISA international standards. This is, by construction, 1.25 standard deviations above the OECD mean. This is the score near the middle of students in proficiency level 5 (from 607 to 668). Students above this benchmark of proficiency are capable, among other things, "of advanced mathematical thinking and reasoning and can interpret complex information about real-world situations"[7]. Figure 1 shows that only 0.29 percent of students who took the test have performed above the advanced international benchmark. This is compared to 18.2 percent of those tested in Korea and 6.5 percent in the United States. This implies that only 3 in 1000 Mexican 15 year-olds tested were "advanced" or above in mathematics. This is compared to roughly 100 in 1000 above that threshold OECD wide[8].

---

[4] These are: level 1 (358-420), level 2 (420-482), level 3 (482-545), level 4(545-607), level 5 (607-669), level 6 (above 669). In order to be assigned to a level of proficiency a student must provide the right answer to the majority of the questions of the related level. See table A1 in the appendix for a description of the students' skills and knowledge at each level of proficiency.

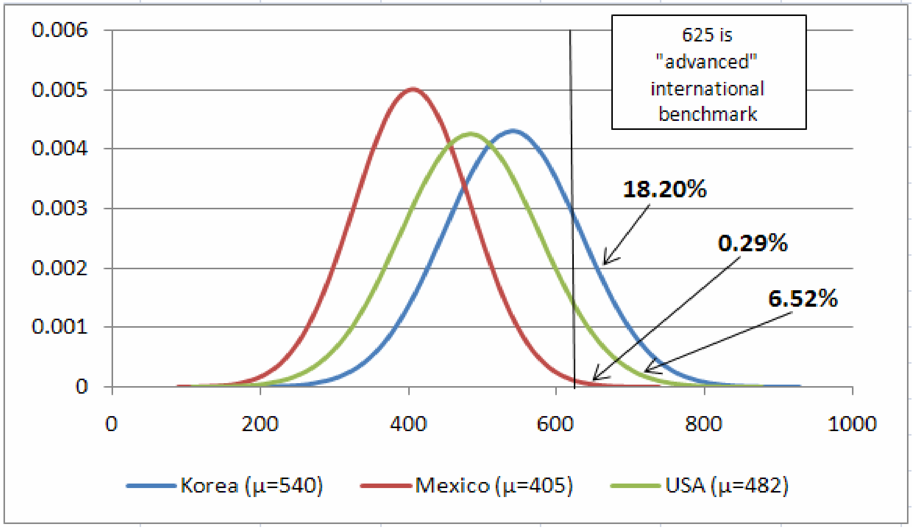
[5] PISA 2003, p.11.

[6] PISA 2003, Technical Report (p.261). Refer to appendix A1 for a description of the other levels of proficiency.

[7] PISA 2003, Technical Report (p.261). This corresponds to levels of proficiency 5 and 6.

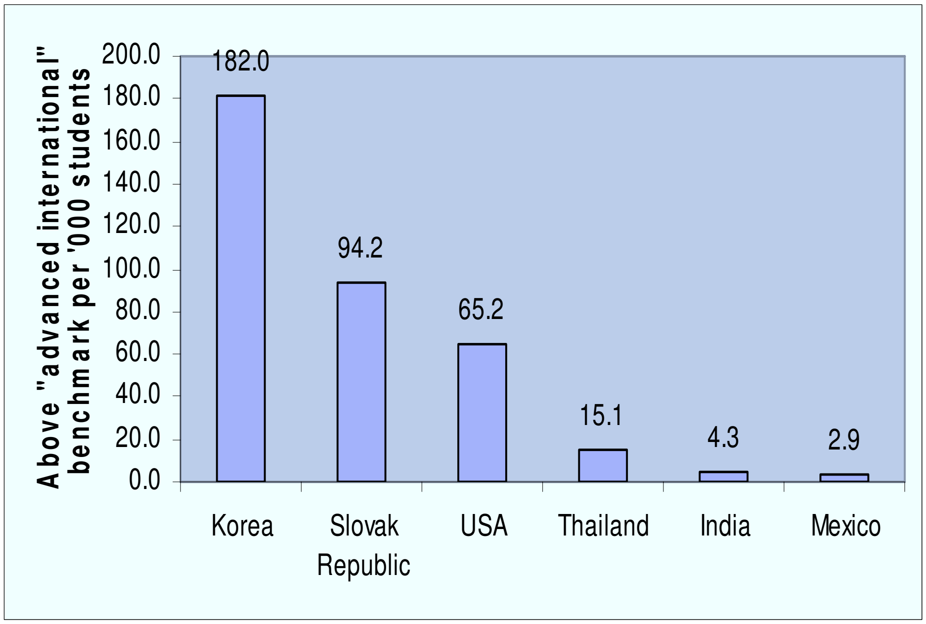
[8] Since the test is constructed to have mean 500 and standard deviation 100, the level 625 is roughly at the tenth percenteile.

Figure 1: Illustrating the distribution of test scores in the PISA 2003 Mathematics assessment for three countries, indicating the fraction above a high benchmark score.



If we compare Mexico's production of global performers per thousand people in the cohort we find this number to be extremely low for Mexico (figure 2). Again, the OECD standard is roughly 105 per thousand, Korea is well above that level, the Slovak Republic just below, Thailand is much below that level, with only 15 students per '000. But this is still five times higher than Mexico's level of 2.9 per '000. India has not participated in the TIMSS, but recently researchers have attempted to compare India to other countries using matched questions for two states and extrapolating, in this case from TIMSS comparable questions, but normalized in the same way to be crudely comparable. India, whose *average* is much lower than Mexico's still had a higher proportion above the threshold than Mexico.

Figure 2: Number of 1,000 students estimated to be above 625 on PISA 2003 in Mathematics (except for India)



Source: Author's calculations.

Making scores comparable across countries implies that usually the results are reported as percents or summary statistics of scores which do not depend on the absolute numbers. But it might be of interest how many students there are above a threshold. The very small share of students at the top of the distribution also implies a small absolute number of students above the advanced international benchmark. The difficulty is that we only have actual information on the tested population, which were intended to be a random sample of those in school. We can make two alternative assumptions. One is that the same proportion of the non-tested students would have scored above the threshold as tested students. This gives an "upper bound" on the total. Alternatively we can assume that none of the students not in school at age 15 would have scored above 625 if tested. In this case we calculate the total number by multiplying the cohort size by the gross enrollment in secondary schools to estimate the total enrolled population. This gives us a lower bound on total number. Work done in other contexts by one of the authors in connection with calculating the number of students above a potential Millennium Learning Goal that constructs cohort estimates from tested students and learning profiles (Filmer, Hasan and Pritchett, 2006) suggest that for Mexico the true number is more likely near the lower bound than upper bound as few drop-outs would be above the bound.

Table 1 shows that based on the lower bound estimates around 3,500 students are above the advanced international standard and even on the very optimistic assumptions of the upper bound the number is only 5,822. Of two million 15 year olds in Mexico every student above an advanced standard could fit in a small auditorium. Of course there are many other countries who also have small absolute numbers, but we calculated a the similar figure for a few other illustrative countries.

The Slovak Republic is a small country but with roughly OECD average quality but with only 85,000 in a cohort they produce absolutely more global performers than Mexico. Thailand is an emerging middle income country (still below Mexico's average GDP per capita) without particularly stellar schooling, but which produces over 10,000 per year. Korea is renowned for at least a type of academic excellence and with only 700,000 students produces 124,000 above this level. The USA, which happens to be close to Mexico, does not have good test scores by OECD standards but still produces almost a quarter of a million students a year with this level of capability. This means for every Mexican 15 year old above 625 there are 69 American students above that standard.

For India we only have very crude calculations[9], but the comparison is very instructive, particularly given India's sustained rapid growth and strong emergence in many IT and science related fields. The educational system in India as the basic (primary and

[9] These are based on a recent paper using TIMSS questions given to students in only two states of India, extrapolated to the national level. Although these calculations are the best that can be done, they are far from "official" and should be taken as rough approximations.

junior secondary) is extremely weak on average. A variety of recent assessments have shown that the typical Indian primary school child has extremely weak learning performance—much worse than Mexico. However, at the same time, at the higher levels there is strong competitive pressure on high stakes for the students examinations at grade 10 (for eligibility for 11-12 courses) and for university admission. This means that the upper tail has maintained very tough standards, a very high level of private sector participation, and very high student effort. This means that even though India has very weak *typical* student performance the upper tail is much thicker than one might expect. That, combined with a large cohort, produces a result that India produces roughly 100,000 students per year above this global benchmark—27 times as many as Mexico.

Table 1: Simple estimates of the total number of 15 year olds above an "advanced international benchmark" in mathematics for selected countries

	Cohort Size of 15 year olds	Gross Enrollment Rate in Secondary School	Estimated number of test takers (15 year olds, enrolled)	Test takers per 100 above the "advanced international benchmark" of 625 in Mathematics	Estimated absolute number of students above threshold	
					Lower bound	Upper bound
	A	B	C	D	C*D	A*D
Mexico	2,007,721	60	1,204,632	0.29	3,493	5,822
Slovak Republic	85,095	75	63,821	9.42	6,012	8,016
Thailand	1,021,145	71.2	727,055	1.51	10,979	15,419
India	21,994,737	52.3	11,503,247	0.83	95,659	182,904
Korea	701,056	97.2	681,426	18.2	124,020	127,592
USA	4,178,014	88	3,676,652	6.52	239,718	272,407

a) India has neither PISA nor TIMSS results, but a recent paper was able to estimate this number based on matching TIMSS methods. The percent is derived backwards from the raw lower bound estimate.

One way of illustrating the consequences for the average ability of the best performers is to calculate as we have done the fraction above some threshold. The other way is to compare the differences in performance of the very top. The International Mathematical Olympiad is a competition held every year for high school students. Each country can send up to six contestants. They are each given six very difficult questions and their answers of given a mark from o to 7, so that the maximum score for any individual student is 42 points. In table 2 we show the average score per eligible student for the contests between 2000 and 2007[10] for each of the countries in table 1, plus China. If one assumes the process of choosing the national contestants is reasonably effective then this is a comparison of how good the very best of each countries aspiring mathematicians are. As we see, these results are ordered exactly the same as the estimates of the number of students above the threshold. The typical Mexican contestant scores about as well as the typical Slovakian and only half as well as the typical Korean and almost a third the level of the typical Chinese contestant. Again, India, although it has on average worse education statistics outperforms at the top, in this case likely due to large size.

Table 2: Average scores of the six contestants from various countries in the International Mathematical Olympiad, 2001 to 2007

	Average score (out of 42 possible) of each eligible student on the contest examination	Cohort size of 15 year olds
Mexico	13.3	2,007,721
Slovakia	15.9	85,095
Thailand	19.2	1,021,145
India	21.3	21,994,737
Korea	28.0	701,056
USA	29.9	4,178,014
China	35.1	

Source: Results of International Mathematical Olympiad, various years.

[10] Except for 2005, as the web site that listed the results was not functional. Coincidentally, the 2005 contest was held in Mexico.

These results need to be read keeping in mind the properties of order statistics. If two countries had identical means and variances but were of different sizes then one expect the large country to have better performance at the top simply because of the larger sample from which it is drawn. The fact that Slovakia produces six students that outperform the Mexican six is striking when one considers the fact that Slovakia has less than a twentieth the number of high school aged population from which to generate the six. Therefore the modestly better scores in the Olympiad are consistent with substantially better typical performance[11].

We are *not* suggesting that mathematics is particularly central to either academic or economic performance. We also did these same calculations with both PISA reading and science scores with similar results. Nor is there anything particularly important about the "benchmark" of 625 that we use. The use of mathematics and the level of 625 are simply *illustrative* of the two issues that we think have not been sufficiently stressed in the discussion about education quality and which would hold true as an issue for any subject or any threshold.

The first issue is that the issue with low averages is not just low averages. A low average score without an elongated upper tail implies that proportionately very few students are high performers by an absolute or international standard of performance. This means the top Mexican students will only be at a level that is quite common in better performing countries. It also means the performance of the best will be substantially lower.

[11] One can simulate the differences in averages that would be consistent with the observed differences among the top six students assuming they are drawing from populations proportional to the cohort sizes. A simulation based just on these Olympiad scores produces inferred differences in central tendency (assuming equal variances) consistent with the rankings of large scale tests for these countries—e.g. shows that Slovakia has much better average scores than Thailand or India but does worse only because of size.

The second issue is the absolute numbers of high performers. Only a very small number of students are ready to go to on in higher education and to compete internationally.

1b) Low Quality at Higher Levels of Education

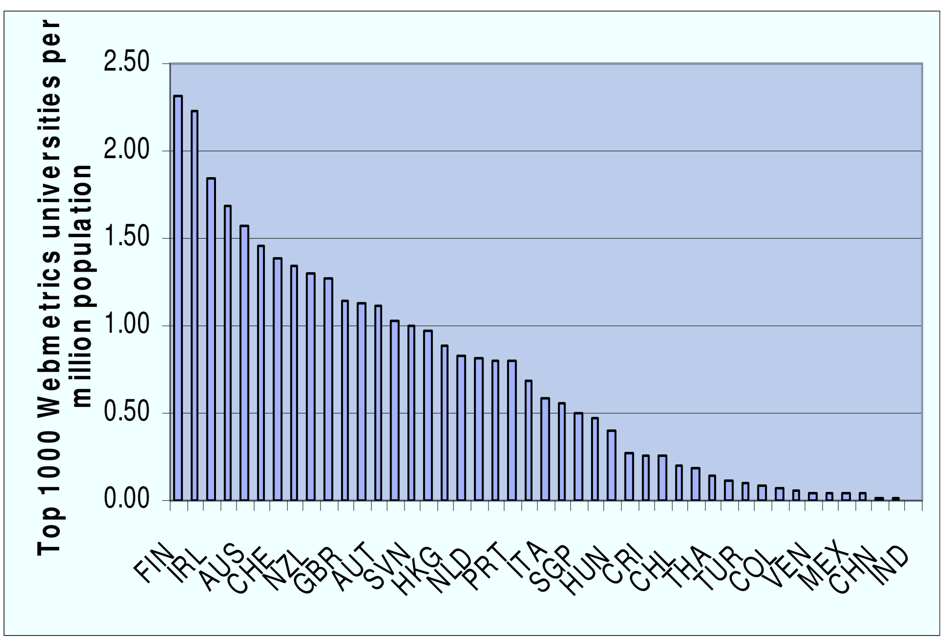
One might conjecture that these deficiencies in quality of secondary students are made up for in higher education. This is almost certainly not true, especially if the comparison is with students in the USA. It is of course very difficult to rank institutions of higher education, and one must take existing rankings with some considerable caution and grains of salt. However, the broad pattern one finds is so striking it is unlikely that other methods would over turn the results.

According to the Shanghai Academic Ranking of Top World Universities, there is only *one* Mexican university in the top 500. This ranking assesses broad based universities based on faculty, facilities, research etc. They find that only Universidad Nacional Autonoma de Mexico (i.e., UNAM, which ranks 185, between University of Miami and University of Nebraska - Lincoln). This of course leaves off several Mexican institutions of higher education that almost certainly are of high quality but may be too specialized to make a list aimed to assess general universities (such as ITAM or Monterrey Tech). But even if one were to add these, the point remains the *typical* higher education experience in Mexico is unlikely to overcome initial gaps, and if any thing exacerbate them.

There are other rankings of top universities. According to the London Times Higher Education Supplement 2007 (THES-QS) among the top 400 higher education institutions, Korea has 7, Brazil has 3 and Mexico only one. Again this is UNAM ranked at 192. This is not saying that Mexican universities don't rank with the world's best (in the top 20 are universities from US, UK, Australia, Japan, Hong Kong, Canada) but that they are not superior to middle tier US universities that round out the 400—University of Kentucky, Georgia State, University of Missouri.

The third global ranking focuses on the universities Web presence, which is probably also distorted in various ways, but provides a cross check on the others and enables us to go even deeper than 400 and 500. If we look at the number of Mexican universities in the top 1,000 web presence universities we only find four Mexican institutions. This implies there is only 1 per 25,000,000 people (figure 4). This implies that the very small number of global performers at secondary level which has been described in the previous section does not have the possibility to receive instruction from world-leading universities. At best, those high achievers can benefit from an average tertiary education.

Figure 3: Ranking of Universities based on Web visibility, numbers of universities in top 1000 globally per million population



Source: Authors' calculations.

Moreover, education research is still very limited. On a per-capita basis, Mexico produces a very small number of highly-qualified workers: in 2002 Mexico graduated 1.4 Ph.Ds per million inhabitants, against 22 Ph.Ds in the U.S. (Santibañez et al., 2005, p.ix). And out of more than 200 graduate programs, only four are recognized by the Consejo Nacional de Ciencia y Tecnología to be of "high-quality" by international standards.

Of course this is just examines domestic universities and a large number of Mexicans seek degrees in the US or UK or other destinations. In 2006/07 it was estimated there were about 14,000 Mexican students studying in the USA (of which only some fraction were undergraduates). This still remains a small fraction of the overall education outcomes in the USA.

2) Rise of returns to "*super-stars*" in the United States

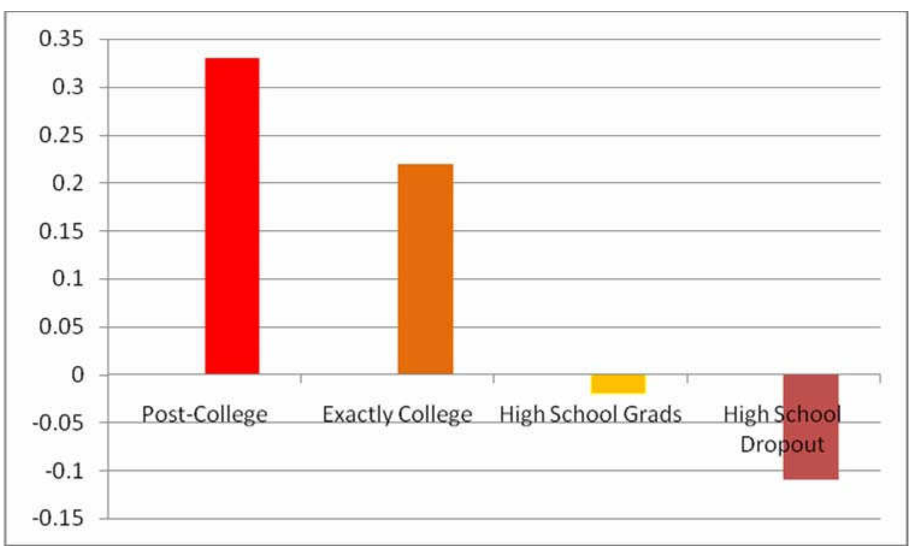
We now shift attention from the quality of Mexican education to the changing returns to various levels of skills, primarily in the US labor market. We will come back and link these two in good time.

The demand for workers at various skill levels has changed over time and we examine the US. experience in some detail, as it shows that this expansion in inequality and rise in the returns to "quality" has, as Krugman once argued, a fractal like aspect—not matter where you look inequality was increasing—not just been "unskilled" and "skilled" but within occupations, and within the top of the distribution among the educated as much (or more) than in the bottom. Labor market inequality has increased in recent decades in the United States. This increase was virtually non-existent at the bottom, moderate in the middle and strong at the top of the distribution (Autor, 2007). That is, this wage growth appears to be "polarized" at the high-end of the wage distribution.

The mostly widely remarked upon and research phenomena is the rise of the premia to a college education as the differential growth of wages of those with and without a college degree has caused the rations to expand. But examining more closely the relationship between wages and education, and we look at the change in wages by education group (figure 4) we can observe that both high school dropouts and high school graduates have experienced falling real wages whereas college graduates have experienced a significant increase in labor income. Autor, Katz and Kearney (2007) suggest that this may be partly explained by the introduction of information technology which by complementing abstract and complex tasks increases the demand for highly educated workers and by substituting routine tasks reduces the demand for less qualified workers.

However, what is also noticeable is that the wages of those with a *post-college* degree has risen by more *in percentage terms* than those with just a college degree. This suggests that the degree of skills being rewarded is not just have some analytical capability but even more than what is gained from four years of college in the USA.

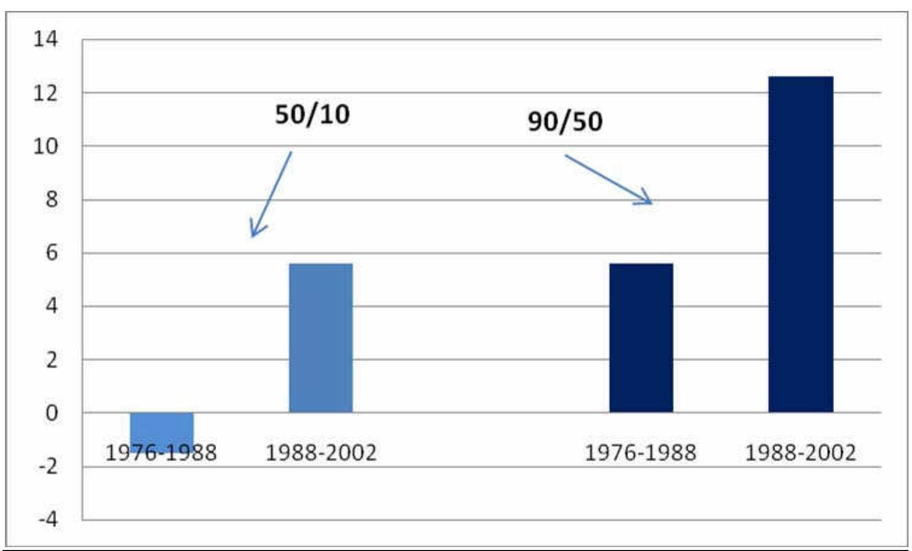
Figure 4: Changes in Composition-Adjusted Real Log Weekly Full-Time Wages by Education, 1981-2005, U.S.



Sources: wages from Autor (2007), *Structural Demand Shifts and Potential Labor Supply Responses in the New Century;* average years of schooling from Barro and Lee (2000), *International Data on Educational Attainment: Updates and Implications*

One can follow this up by examining the distribution of wages among only those who have a four year college degree. Inequality can increased in a variety of ways, either "radially" and symmetrically or asymmetrically if either the middle pulls away from the bottom or if the top pulls away from the middle. Figure 5 shows how upper- and lower-tail inequality (summarized by 90th-50th percentile and 50th-10th log wage differential) have evolved over the periods from 1976-1988 and then from 1988-2002. In the earlier period the top pulled away from the middle but the bottom gained on the middle. In the latter period (from 1988 to 2002) the inequality increased in both ways, but the increase of the top (90th percentile) pulled away even more rapidly from the middle. This suggests the skills that were increasingly in demand were not just "having a college degree" but even among those people with a college degree the more skilled (or at least at the top of the earnings distribution) were even more in demand.

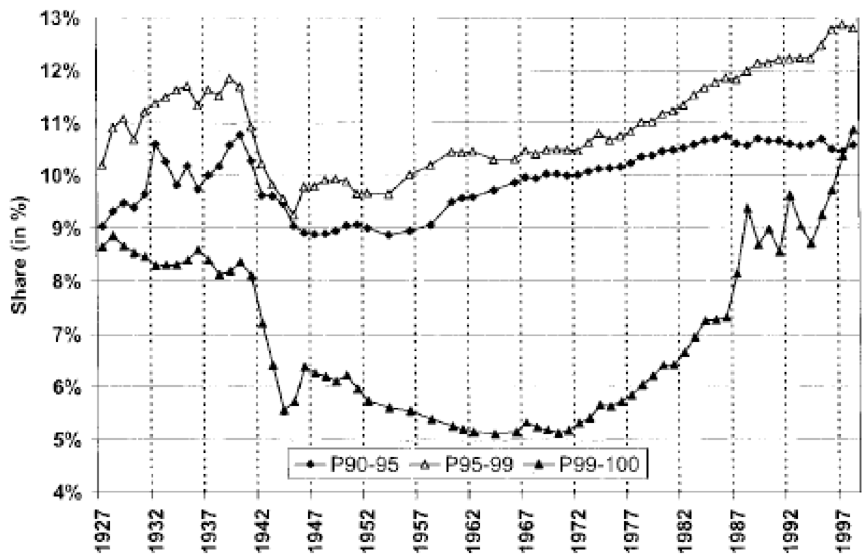
Figure 5: Changes (x100) in Male Hourly Earnings Inequality among College Graduate Males with 24-26 Years of Experience: 1976 to 1988 and 1988 to 2002, U.S.



Source: Autor, Katz and Kearney (2005), *Rising Wage Inequality: The Role of Composition and Prices* Note on data: March CPS 1976-2004. Statistics pool three years of data centered on indicated year. College graduates are those with 16 or 17 years of completed schooling (surveys prior to 1992) or a baccalaureate degree only (1992 forward).

Finally, if we look at the long run evolution of wages by percentile, we can observe that the very top has really pulled away from the rest of the distribution. The share of total wage income going to the top 1 percent earners increased from lows of round 5 percent of total earnings to almost 10 percent of total wages by 1998, when these particular data end (figure 6)[12].

Figure 6: Top Wage Shares in the Long Run: Wage Income Shares for P90-95, P95-99, and P99-100, 1927-1998, U.S.



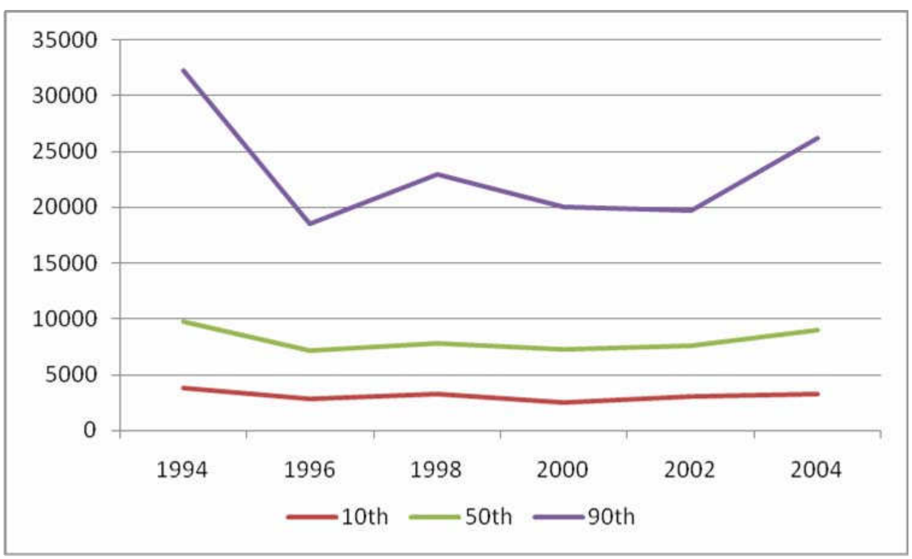
Source: Piketty and Saez (2003), *Income Inequality in the United States, 1913-1998*

As Rosen (1981) first suggested, the growing inequality can be explained by the fact that thanks to the modern technologies the "*super-stars*" are greatly rewarded whereas the runners-up get far less. This affects the demand for skills along the distribution and causes the top to pull away from the rest.

[12] This appears even more striking when we look at how fast the relative change has occurred: in 1970, 0.01 percent of taxpayers only earned 70 times as much as the average whereas one in 1998 the richest 13,000 US families had income 300 times greater than the average family (Krugman, 2002).

There is evidence that this phenomenon is not just limited to the United States but is taking place in the majority of the OECD countries, including Mexico. In figure 7 we can observe that among the college graduates, those at the bottom and middle of the wage distribution have experienced a rather flat wage profile over 1994-2004. On the other hand, those at the high-end of the distribution experienced a decline in wages after *el Error de Diciembre* due to the severe macroeconomic conditions brought on by the peso devaluation that adversely impacted the better educated workers. But over time, trade liberalization as well as market-oriented reforms have increased demand for more educated workers, and the increased demand for college graduates increased in *all* industries and was a result of the "within" industry shifts (Cragg, Epelbaum and Meza-Gonzales, 1997). After 2002 it is possible to observe that the top is pulling away from the middle and bottom of the wage distribution even among college graduates.

Figure 7: Average Male Wages among College Graduates for P10, P50 and P90, 1994-2004, Mexico



Source: Encuesta Nacional de Ingresos y Gastos de los Hogares, 1994-2004 and CPI from the IMF International Financial Statistics

3) Implications for Mexico

In this section we are going to compare two general points of view with regard to improving the quality of education in Mexico and its potential consequences. This has two dimensions. One is to think through policies and their impact on the *distribution* of skills and by how much. A second is to envision Mexico embedded in a global economy and analyze whether Mexico's education policies are expanding the supply of a range of skills for which there is robust evidence there is expanding demand. Let us be the first to warn the reader we are going to say that two very unpopular things. We are going to say these things without definitive proof, but as suggestive and provocative. They should at least be considered and examined as alternatives to the overwhelming messages about education that do not make these two points explicit.

First, the array of options to improve quality suggest that expansion of "business as usual" policies for improving the quality of education, while they may be justified on narrow cost-benefit grounds, are unlikely to have transformational effects on quality. Second, radial expansions of quality from Mexico's current levels of quality will augment a segment of the range of skilled labor for which there are at least serious questions as to whether global demand is expanding fast enough to accommodate.

3a) Improve quality of education: Trapped in a flat bowl

Let us just illustrate what we mean about a "flat bowl" on a general level before delving into the specifics of the evidence about individual interventions. This illustration just links together three facts that are widely acknowledged but seldom considered together.

First fact is that Mexico is far from the international frontier in terms of student quality. There are enormous differences across students in measured competencies and hence the "student standard deviation" is very high—in the three areas of PISA it is typically between 80 and 95. Across the board Mexico is roughly a full student standard deviation behind the OECD average (which is often roughly the US level) and more like 1.5 student standard deviations behind the "cutting edge" countries like Korea. This means that students near the very top in performance in Mexico (the 95th percentile) would be roughly the average performer in Korea. This means the average performer in Mexico would have to have massive improvement to be average in the USA or Korea

Table 3: How much will it take to bring average performance in Mexico up?			
	Scores on the 2006 PISA	How many Mexican student standard deviations[a] Mexico is behind (or "effect size" needed for catch-up)	Ratio of country/region *average* student scored to Mexican 95th percentile student score
Mexico	**406**		
USA	474	0.8	87.0%
OECD	498	1.1	91.2%
Korea	547	1.7	100.4%
	Science		
Mexico	**410.00**		
USA	489	1.0	89.9%
OECD	500	1.1	91.9%
Korea	522	1.4	96.0%
	Reading		
Mexico	**410.50**		
USA	NA		
OECD	492	0.86	88.0%
Korea	556	1.53	99.4%

a) The Mexican student standard deviations were calculated as the 5th-95th range divided by 1.642*2 (under the assumption of a normal distribution). The results were Math 84.9, Science 79.9, Reading 94.9 (the OECD student standard deviation is 100 by construction).

Second fact is that the absolute *magnitude* of the learning gains that are demonstrated in the typical proposed educational improvement scheme are very small (see Filmer and Pritchett, 1996; Pritchett, 2004). The literature on education often uses "effect sizes" to have a common metric for evaluating the magnitude of learning gains (as otherwise test instruments with different absolute scales would have different apparent absolute impacts)[13]. The typical effect size in the literature of the standard "business as usual" expansion of inputs is roughly zero. No definitive conclusion has been reached by scholars on what education policies and reforms may be most effective in improving the overall quality of schooling. There is certainly no clear causal relationship between expenditure and students' achievement (Hanushek, 2003, 2006). This does not mean that "money does not matter" or "money cannot matter" (Hanushek, 2007, p.9). On the other hand, it reveals the importance of making an effective use of resources to produce positive results. In this regard, there is general agreement on basic aspects such as the importance of teacher quality, the need for standards and accountability as well as the possible benefits of incentives and market-oriented reforms. However, empirical findings are not conclusive and often show that the impact of a specific policy is highly dependent on the institutional context where the reform is implemented and the time of assessment. In what follows, we provide an overview of significant education policies, most of which have been recently implemented in Latin and Central America to improve the quality of schooling. They can be grouped according to their focus under three categories: teacher quality, resources and school-based management reforms.

The point of this is that, even when it comes to available options about which there is even semi-conclusive evidence about their efficacy in practice in raising scores, these do not provide a definitive guide for substantial improvements in performance and the *magnitude* of the impacts even of those that are demonstrated to be statistically different from zero is often very small. An effect size of an intervention of any kind of a tenth of a standard deviation is considered very large.

The third fact is that even once one finds interventions that have a substantial effect size the *scope* for the application of the intervention is often limited. Many interventions that are proposed

[13] Of course this far from solves the problem as even the student standard deviation depends on the underlying evaluation instrument—a test that was far too hard for the tested population might return as very low student standard deviation as the scores cluster on zero. Often in empirical studies the standard deviation of the assessment is itself normalized and then impacts are reported as effect sizes but the effect sizes may or may not be comparable.

are remediation of shortages (e.g. large class sizes, under-trained teachers, lacking facilities). But in these cases the impact on the *total* or *average* score is the treatment effect times the potential scope of the treatment. For instance, suppose one found that under qualified teachers could be brought up to par with training. Then the total gain is the gain per trained teacher times the number of potentially trainable teachers. So even if the training were to show an effect (which they often do not) and even if this has a huge effect size of .1 (which is even less plausible) then if the percent of teachers for whom this training is effective is 20 percent of the teachers, it adds up to a gain of .02 student standard deviations.

The upshot is the *trip is long, the vehicle is slow and you are almost out of gas.* We illustrate this somewhat more concretely, again this is an illustration, not concrete proof, by examining the empirical magnitudes of the potential gains from the type of educational reforms being discussed in Latin America generally. We then summarize the potential gains in table 4 and figure 8.

<u>Performance-based pay bonuses for Teachers:</u> teachers play a key role in students' learning (Hanushek at al., 2005; Vegas and Umansky, 2005) and teachers' salary represents the largest share of educational expenditure. This explains why teacher incentive reforms are one of the main education challenges faced by Latin American countries. Specifically, two programs have been recently implemented in Mexico and Chile to improve teaching quality by providing teachers with bonuses linked to their performance. The *Carrera Magisterial* was introduced in Mexico in 1992 to modernize primary schooling. Among other things, the program replaced the five-year seniority teacher pay scale with a new pay structure where improvements in students' performance represented 20 percent of the total weight (Mizala and Romaguera, 2004, table 2). It consists of a promotion system where teachers and principals are evaluated on an individual basis. Empirically, no positive effect has been found on students' performance and this may be partly due to the weak

incentives faced by teachers and the significant role played by unions in determining the final teachers' pay improvements (Vegas, 2005).

On the other hand, the Chilean Program introduced in 1996, *Sistema Nacional de Evaluacion de Desempeno de los Establecimientos Educacionales* (SNED), seems to have been more effective. This group-based incentive is assigned to the highest performing schools that enroll at least 25 percent of students in each region; this award represents 5-7 percent of teachers' annual wages. It assigns a greater weight to improvements in students' performance in determining teachers' award: 28 percent. Moreover, the effectiveness of students' performance is included in the evaluation and counted for 37 percent of the total weight (Mizala and Romaguera, 2004, table 2). There is evidence of a positive effect of this reform on students' performance especially for those schools more likely of winning the award (Vegas, 2005). What the reforms undertaken in Chile and Mexico show is that the political context and the unions play a significant role in the design and implementation of teacher reforms (Vegas, 2005).

Finance equalization: The *Fundo para Manutenção e Desenvolvimento do Ensino Fundamental e Valorização do Magistério* (FUNDEF) was introduced in Brazil in 1998, with the aim of reducing the inequality of the education system. Specifically, FUNDEF main aim is to redistribute resources from the richer to the poorer regions and to increase public teachers´ wages. The program has led to an increase in teachers' wages and to a relative improvement of the public schooling system (Menezes-Filho et al., 2004), also for specific demographic groups (Gordon and Vegas, 2005).

Class size reduction: 75 percent of the existing studies have found no effect of a decrease in the pupil-teacher ration on students' performance (Hanushek, 2007). Among the remaining 25 percent, the evidence is mixed[14]. The Brazilian finance equalization program, FUNDEF, while

[14] Surveys of the evidence on class size include Hanushek (1996, 1986), Card and Krueger (1996).

redistributing revenue led to changes in educational inputs and in particular to a reduction in class size. The available empirical evidence does not show improvements in students' performance resulting from a reduction in the pupil-teacher ratio (Gordon and Vegas, 2005). If we want to estimate what would be the largest possible gain from this policy we can focus on Krueger's evaluation of the Project STAR, an experiment carried out by the State of Tennessee in the mid-1980s, which involved a comparison of achievement by students randomly assigned to classes of different size. Krueger (1999) finds significant and large gains from a reduction in school size.[15]

School Awards: These incentives have been recently introduced in Latin America so that most of the existing programs have not been evaluated yet. The collective incentives appear to be particularly effective with respect to the individual ones as they promote cooperation to achieve common objectives (Mizala and Romaguera, 2004). The programs implemented in Latin America differ in the structure and requirements. The school award introduced in Bolivia in 2001 is to provide a monetary compensation to teachers' principals and staff based on the overall improvement of school performance. On the other hand, in the framework of the school award PLAN implemented in El Salvador standards are set by the Ministry of Education and the school personnel are remunerated accordingly (Mizala and Romaguera, 2004). In addition to the teachers' award previously described (which represents 90 percent of the SNED bonus), the Chilean Program provides the remaining of the SNED bonus to schools as "an excellence subsidy". Schools have autonomy with respect to the use of this award (Mizala and Romaguera, 2005).

[15] We acknowledge the limitations of Krueger's study (Hoxby, 2000; Hanushek, 2007). The purpose here is only to show the maximum possible gain in case the policy was effctive.

Table 1: Policy Interventions, Effect Size and Maximum Gain

Quality Intervention	Examples	Effect Size on Students' Performance	Maximum Gain
Teacher quality			
Performance-based pay bonuses for Teachers	The *Sistema Nacional de Evaluacion de Desempeno de los Establecimientos Educacionales* (SNED) was introduced in Chile in 1996. Among the objectives, key is the improvement in teacher quality. It offers bonuses to schools that show excellent performance in terms of students' achievement. 90% of the SNED bonus is divided by teachers in the school (which represents 5-7% of the annual wage) *Carrera Magisterial* was introduced in Mexico in 1993. It provides teachers with large financial rewards that are based, among other factors, on students' test scores. Participation is on a voluntary and individual basis.	Positive impact on students' achievement especially in those schools more likely to win an award. (also positive effects on teachers' attitudes and quality of entrants into teacher education programs increased). (Mizala and Romaguera, 2005) None (not robust) (McEwan and Santibanez, 2005)	0.05 increase in school performance; teachers' average salaries rose 156% over 1990-2002
Resources			
Finance equalization	The *Fundo para Manutenção e Desenvolvimento do Ensino Fundamental e Valorização do Magistério* (FUNDEF) was introduced in Brazil in 1998. It aimed to promote greater equity in educational opportunities between states and across municipalities by providing a minimum per pupil expenditure in primary schools throughout the country	Mixed: Reduction in spending inequality positively affects nonwhites students and students at the bottom of the distribution (Gordon and Vegas, 2005). Improvements in students' test scores forstudents in public schools with respect to their counterparts in private schools appear to be partly related to teachers' increased wages. The effects appear to be concentrated in the Northeast of the country (Menezes-Filho and Pazello, 2004)	
Class size reduction	The *Student-Teacher Achievement Ratio* (STAR) was a randomized experiment undertaken in Tennessee in the 1980s FUNDEF: increased resources were partly used to reduce class-size (in Brazil starting from 1998)	Positive and significant effect on students' performance (Krueger, 1999) No effect (Gordon and Vegas, 2005)	0.2 standard deviations of test performance in reading and math

Table 1 (cont'd): Policy Interventions, Effect Size and Maximum Gain

School-based management reforms			
Performance-based Pay bonuses for Principals	*Carrera Magisterial* (introduced in Mexico in 1993), allows Principals to receive an award based on the overall school performance. Participation is on a voluntary and individual basis	None (not robust) (McEwan and Santibanez, 2005)	
School Awards	The *Incentivo Colectivo a Escuela* (ICE) was introduced in Bolivia in 2001 to encourage collaboration between principals, teachers and staff in primary schools The *Plan de Estimulos a la Labor Educativa Institutional* (PLAN) was introduced in El Salvador in 2000 to encourage public school teachers to work together to solve the problems affecting their schools and improve the quality of educational services that they offer the community SNED bonus (10%) is given to schools as an "excellence subsidy"	No assessment available of the effects on students' performance No assessment available of the effects on students' performance (Discussed above)	
School autonomy	The *Educacion con Participation de la Comunidad* (EDUCO) Program was established in El Salvador in 1992 with the aim of increasing decentralization and delegating the school decision-making authority of pre-schools and primary schools to community organizations and parents *Autonoma Scolar* started in Nicaragua in 1993. It introduced decentralization of the schooling system with a financial-administrative focus	Positive effects on teacher behavior. Teachers may have more motivation (e.g., by dedicating more time teaching, being absent less and by spending more time meeting with parents). No conclusive evidence on the effects of these policies on students' performance (Sawada and Ragatz, 2005) Differences between autonomous and centralized schools do not seem to affect students' outcome, results are not robust to different specifications (Parker, 2005)	

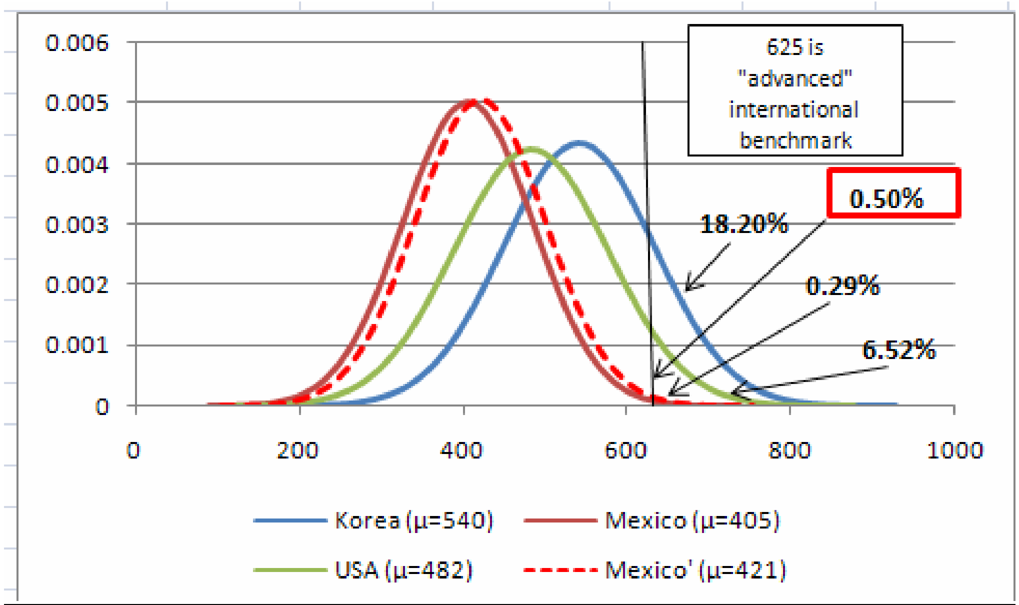
Source: Authors' elaborations based on Vegas (2005), Mizala and Romaguera (2005), McEwan and Santibanez (2005), Gordon and Vegas (2005), Krueger (1999), McEwan and Santibanez (2005), Menezes-Filho et al. (2004), Sawada and Ragatz (2005), Parker (2005).

It is possible to observe that only the very costly and highly controversial class size policy appears to have a considerable effect on students' achievement, as found by Krueger (1999) for the randomized experiment carried out in the 1980s in Tennessee (and even then the effect was concentrated). If we consider the other policies we can observe that when effective they appear either to have a more moderate effect or to improve the performance of specific groups of students. If we do not consider any issue related to the reliability and external validity of Krueger's (1999) study[16] and we relate the effect of this policy to the distribution of test scores presented in a previous section (figure 1) we can observe that at best the class-size policy would lead to a modest gain in the mean test score.

Even in this "best-case scenario" that interventions were able to increase quality by .2 student standard deviations (and again it is worth stressing these are at the outer range of any class size effects estimated nearly anywhere and that achieving these gains would be costly and take a long time) the average *quality* of education in Mexico would remain far below those of Korea and the United States. The percent of Mexican students above the international benchmark would increase from .29 percent to .50 percent (figure 8) and hence in the low estimate the total number of global high performers would increase from 3,500 to only 6,000—after years of effort and huge increases in expenditures one would need a slightly larger auditorium to hold the global high performers.

[16] A study by Woessmann (2000) summarized in Pritchett (2004) shows that, even using plausible techniques for identifying the causal impacts of class size reductions, examining the evidence across more than a dozen OECD countries using the TIMSS data, *none* of them find an effect as large as Krueger suggests and most of them are very near zero. While identification is an issue, even class size impacts identified with randomized experiments in contexts like India and Kenya find essentially no impact.

Figure 8: Simulated effect on the PISA 2003, Mathematics scores of the *maximum* possible effect of class size reduction



3b) Expansion of "business as usual"

The other popularly recommended education policy is expanding schooling, and in much of Latin American this constitutes a call for making higher and higher levels of secondary schooling universal. Increasing the average education level of the Mexican population would imply expanding primary and secondary over tertiary. There are many reasons one might want to make secondary schooling universal and we make no arguments again that as a social policy that Mexico may wish to pursue.

The question we raise is whether that would likely have much impact on Mexico's economy or economic position. Consider the global demand and supply for various types of skills and skilled labor. If one is producing "stuff"—that is, engaged in routine manufacturing—then all countries have been coping with the massive effective expansion of relatively low skilled labor that has happened as first China and then India have each added their billion person populations to the effective supply of unskilled and semi-skilled labor capable of producing manufacturers. Any sustained wage advantage over these economies must be grounded in higher productivity labor that leads to competitive unit labor costs. The question, for which we have no answer or evidence, is whether moving the typical Mexican worker from 8 to 9 or 9 to 10 years of schooling is going to make a substantial difference in Mexico's dynamism as an economy. This is expanding the supply of a factor that world markets (for tradables, which impact labor markets) have been suggesting is hardly in excess demand.

If one is not competing for producing "stuff"—the application of routine manufacturing production techniques to add value—then perhaps one can compete in the market for "ideas" which broadly taken is the addition of value through design, invention, innovation, creativity, first mover advantages, etc. In the market for ideas in tradables (either directly as in service industries such as finance or indirectly in creativity embedded in goods) one is competing with the USA and other advanced market economies (and in fields like engineering East Asian economies). Again, you should ask yourself—will the expansion of the education of the typical young person from 8 to 9 years of school at existing (or feasible) levels of quality really be transformative in equipping Mexico to raise productivity based in competing in the global market for ideas? As we demonstrated above, one of the key social issues in the United States is that among *college graduates* the demand for skills was shifting towards the upper tail of skills. The markets for ideas often display "super-star" features.

4) Discovering the Discoverers

We do not wish to overstate our case. We are merely suggesting a new range of policies that should be considered in discussions of what education policies should be pursued. We are suggesting that, however desirable for social policy reasons educational policies aimed at addressing broad based quality and promoting equality of access are (and we take no issue with these arguments) they are unlikely to be of great significance in the short to medium run economically. This is part just pointing out the unpleasant but obvious—during all of the "lost decades" of stagnating (or falling) levels of output per worker in Latin America, the average levels of educational attainment have been moving steadily upward. If the "business as usual" expansions of existing education systems were capable of producing a growth acceleration then the impacts should have been widely noticeable by now (Pritchett 2006).

What we are proposing is at least the consideration of policies that have three features:

1) Encourage better performance among the top performers,

2) Emphasize broadening the base of talent across socio-economic groups by pro-actively identifying and encouraging academic excellence outside of the currently well-off,

3) Creating a conducive environment for entrepreneurship to that new ideas in the production of tradables can flourish

4.a) Discovering the discoverers

The adoption of imitative technology requires a country to develop the *social capability*[17] to effectively adapt and use the technologies in the production system. That is, the transfer of existing technology needs the *appropriate institutions* to be successful and entrepreneurs who decide how to use it in the most effective way given the other inputs of production[18]. Optimal production strategies greatly differ across sectors. To produce manufactures Mexico must be able to compete with China and Vietnam, to produce low-end portable services Mexico must be able to compete with India and Ireland but to be able to produce high end and "ideas" Mexico must be able to compete with the United States and Israel.

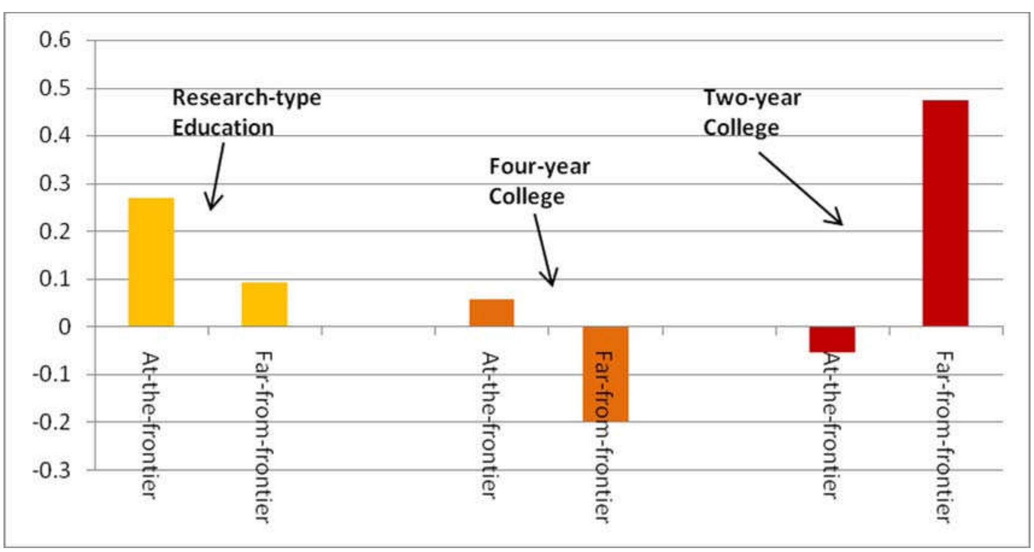
[17] Abramovitz (1986, p.387).

[18] Hausmann and Rodrik (2003) describe how even in case of complete information on technology, entrepreneurs would play the key role in deciding what to produce with them.

This implies that to foster economic growth a country has to “learn what is good at producing” (Hausmann and Rodrik, 2003) and the level of education affects what there is to be discovered in a country’s capability set as discoverers lead “self-discovery”. In the process of Schumpeterian entrepreneurship one needs a critical mass of people of high ability to put together factors in a new way. And Mexico is failing to achieve this.

Aghion et al. (2006) have shown how the composition of human capital and the distance from the technological frontier affect a country’s economic growth. They show that countries that are closer to the technological frontier may benefit the most from investments in research education as this may foster the creation of knowledge and the process of innovation. On the other hand, research education can also have a significant impact on growth and development in far from the frontier countries. In the case of the U.S. it is possible to observe that even in far from the frontier States increased spending in research type education has a positive effect on economic growth (figure 9). As described in the previous section, even in Mexico the “*super-star*” phenomenon shows that returns to the top are high and the demand for highly educated and able workers is increasing.

Figure 9: Effect on per-employee growth rate of a $1,000 per person in additional spending at different levels of education in states at-the-frontier and far from the technological frontier

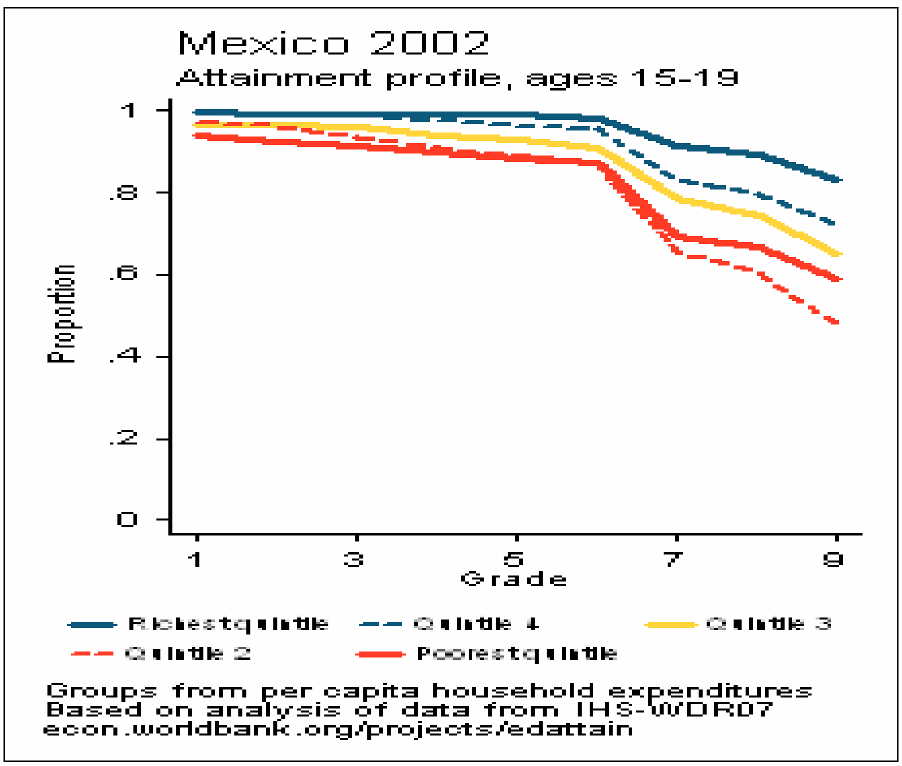


Source: Aghion, Boustan, Hoxby and Vandenbussche (2005)

4.b) Expansion of opportunity through pro-active identification of talent

In this regard, some 45.1 percent of the population aged 15-19 years is not in education in Mexico (OECD, 2008, p.22) and among these only 62 percent is employed whereas the remaining 38 percent is not engaged in any productive activity (i.e., employment, education or training). As we can see from figure 10 educational attainment stratifies sharply on household income, only half the poorest 20 percent even reach ninth grade.

Figure 10: Attainment profile, ages 15-19, by per capita household groups



This means that not even all children in the age cohort are taking the PISA test. This implies that by sorting on socio-economic status, not ability, Mexico is recruiting from a narrow base. If potentially high ability children from low income background drop-out there is a loss to the pool of potential discovers. Mexico would never attempt the Mundial of *futból* recruiting only from a small stratum of the population, why is Mexico attempting the *economic* Mundial?

4.c) Allocation of talent to global tradables

All these factors combined significantly affect the "allocation of talents", that is the relationship between the reward structure of a society and the way individuals allocate their talent between productive and unproductive activities. The allocation of talent in a society is an important determinant of output and growth. Shleifer, Murphy and Vishny, (1991) show how *general talent* is not occupation-specific but its allocation critically depends on the returns to ability between different sectors and the set of incentives faced by individuals. That is, if we exclude those individuals who have exceptional *natural talent* in a specific task like singing opera or playing basketball, other individuals may have higher intelligence and ability that gives them a competitive advantage in any occupation they choose. In the case of Mexico highly educated individuals would choose professional activities insulated from the international competition given the low average quality education they have received at secondary and tertiary levels. These are the "non-tradable" occupations described in table 5.

Table 5: The Allocation of Talent

	Tradable	**Non-tradable**
Innovation	Growth-enhancing professions (e.g., entrepreneurs, engineers, designers)	Nationally regulated professions (e.g., doctors)
Re-distribution		Rent-seeking professions (e.g., lawyers, lobbyist, "connection with wealth")

Moreover, the institutional setting, legal framework and social esteem attached to different occupations will affect how individuals choose their profession. As previously shown, education in Mexico is stratified on income and ability to pay and not on talent and ability. This reinforces the low expansion of high quality education, but also the low social mobility and perpetuation of high inequality. This sorting also affects the way individuals perceive the fairness of institutions and how society rewards their effort and commitment. A self-replicating elite is more likely to be in favor of the *status quo*. Therefore, productive ability may be less socially valuable than rent-seeking behavior[19]. Thus, in this context talent would end up by being concentrated in the "non-tradable" and rent-seeking activities (table 5). These are the low-growth professions as opposed to the growth enhancing ones. If institutions do not encourage private initiative, social mobility and productive activities, talented individuals would choose occupations that do not face competition and low global quality education means the high ability allocate into non-tradable occupations to limit competition. This would create a significant distortion in the allocation of people as highly productive individuals would choose socially unproductive occupations.

We are aware that these fly in the face of the vast majority of recommendations about education. Again, we are not disputing that the usual recommendations, which tend to focus on system expansion in access, broad based improvements in quality, and, if anything, reducing inequality in outcomes by focusing on the low performing schools and students, are correct as *education* policies for a variety of social and internal educational reasons. However, the typical (average scoring) Mexican 15 year old student is roughly at the 18th percentile of the skills distribution of OECD 15 year olds. Moreover, that is roughly average school completion in Mexico while other countries have much higher average levels. So the typical *school leaver* and labor market entrant in Mexico is likely near the bottom 10-15 percent of the skills distribution of the typical labor market entrant in the USA or other OECD countries. It is just difficult to believe that the available, marginal and gradual, improvements in the skills of the typical school leaver will

[19] It refers to "the socially costly pursuit of wealth transfers" (Tollison, 1997, p. 506).

have immediate, growth accelerating effects by in any way facilitating a structural transformation in the Mexican economy or an expansion in productivity.

Conclusion

Education, research, scientific discovery, innovation and economic growth are closely related. The paper has shown that traditional education policies that focus on expanding the education system at the *average*, while they may have many benefits, are unlikely to make Mexico competitive in the knowledge-based global economy. On the other hand, the paper has suggested the exploration of measures to enhance the country's productive capabilities and foster economic growth. An effective development strategy would not simply raise the *average* schooling levels of the population, but would rather enhance the *top*. In this regard, a combination focus on the upper tail and on the expansion of opportunity is needed to position Mexico to compete. The emphasis should be on both "discovering the discoverers": developing the educational system on top quality, fostering global standards of performance, and also on pro-actively identifying low-income and high ability individuals to facilitate educational and economic mobility.

References

Abramovitz M. (1986), "Catching Up, Forging Ahead and Falling Behind", *Journal of Economic History,* Vol. 46, No. 2, pp. 385-406.

Aghion P., J. Vandenbussche and C. Meghir (2006), "Growth, Distance to Frontier and Composition of Human Capital", *Journal of Economic Growth,* Vol. 11, No. 2, pp. 97-127.

Aghion P., L. Boustan, C. Hoxby and J. Vandenbussche (2005), "Exploiting States' Mistakes to Identify the Causal Impact of Higher Education on Growth", *Unpublished Paper.*

Autor D. (2007), "Structural Demand Shifts and Potential Labor Supply Responses in the New Century", Federal Reserve Bank of Boston Conference Series No. 52: Labor Supply in the New Century, forthcoming.

Autor D. H., L.F. Katz. and M.S. Kearney (2007), "Trends in U.S. Wage Inequality: Revising the Revisionists", *Unpublished Manuscript*.

Barro R. and J. Lee (2000), "International Data on Educational Attainment: Updates and Implications", Oxford Economic Papers, 53(3): 541-563.

Card D., and A. Krueger, (1996) ''School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina,'' *Journal of Economic Perspectives,* X: 31–40.

Das J. and T. Zajonc (2008), "India Shining and Bharat Drowning: Comparing Two Indian States to the Worldwide Distribution in Mathematics Achievement", *World Bank Policy Research Working Paper 4644.*

Filmer D., A. Hasan and L. Pritchett (2006), "A Millenium Learning Goal: Measuring Real Progress in Education", Center for Global Development *Working Paper 97.*

Gordon N. and E. Vegas (2005), "Educational Finance Equalization, Spending, Teacher Quality, and Student Outcomes" in Vegas E. (ed.) (2005), "Incentives to Improve Teaching", Washington D.C.: The World Bank, 151-186.

Hanushek, E. A. (1986), ‘‘The Economics of Schooling: Production and Efficiency in Public Schools,’’ *Journal of Economic Literature,* XXIV: 1141–1177.

Hanushek, E. A. (1996), ‘‘Measuring Investment in Education,’’ *Journal of Economic Perspectives,* X: 9–30.

Hanushek (2003). "The failure of input-based schooling policies." *Economic Journal,* 113(485) (February): F64-F98.

Hanushek E.A. (2006), “School Resources” In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 865-908.

Hanushek E.A. (2007), “Incentives for Efficiency and Equity in the School System”, *Perspektiven der Wirtschaftspolitik,* 9(1): 5-27.

Hanushek E. A., J.F. Kain, D.M. O’Brien and S.G. Rivkin (2005), “The Market for Teacher Quality”, NBER Working Paper 1154.

Hanushek, E. A. and L. Woessmann (2008), “The Role of Cognitive Skills in Economic Development”, *Journal of Economic Literature*, 46(3): 607–68.

Hausmann R. and D. Rodrik (2003), ‘Economic Development As Self-Discovery’, *Journal of Development Economics* 72(2): 603-633.

Hoxby C. M. (2000), ‘The effects of class size on student achievement: new evidence from population variation’, *Quarterly Journal of Economics,* 115(4): 1239-1285.

Krueger A. B. (1999), ‘Experimental estimates of education production functions’, *Quarterly Journal of Economics,* 114(2): 497-532.

Krugman P. (2002), “For Richer”, *The New York Times*, October 20.

McEwan P.J. and L. Santibanez (2005), “Teacher and Principal Incentives in Mexico” in Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank, 213-254.

Menezes-Filho N. and E. Pazello (2004), “Evaluating the Effects of FUNDEF on Wages and Test Scores in Brazil”, *CEP Seminar Paper 02-03-04.*

Mizala A. and P. Romaguera (2004), “School and teacher performance incentives: The Latin American experience” *International Journal of Educational Development,* 24: 739-754.

Mizala A. and P. Romaguera (2005), “Teachers’ Salary Structure and Incentives in Chile” in Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank, 103-150.

Murphy K. M., Shleifer A. and Vishny R. W. (1991), “The Allocation of Talent: Implications for Growth”, *Quarterly Journal of Economics,* 106(2): 503-30.

OECD (2003), “PISA 2003: Technical Report”, Paris: OECD.

OECD (2003), “First Results from PISA 2003”, Paris: OECD.

Parker C.E. (2005), “Teachers Incentives and Student Achievement in Nicaraguan Autonomous Schools” in Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank, 359-382.

Piketty T. and E. Saez (2003), “Income Inequality in the United States, 1913-1998”, *Quarterly Journal of Economics,* 118(1): 1-39.

Rosen S. (1981), “The Economics of Superstars”, *American Economic Review,* 71(5): 845-858.

Santibanez L., G. Vernez and P. Razquin (2005), “Education in Mexico: Challenges and Opportunities”’, *RAND Corporation Documented Briefing*.

Sawada Y. and A. B. Ragatz (2005), “Decentralization of Education, Teacher Behavior, and Outcomes” in Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank, 255-299.

Tollison, R.D. (1997), “Rent seeking”, in *Perspectives on public choice*, D.C. Mueller (ed.), Cambridge: Cambridge University Press, 506-525.

UNESCO (2003), “Statistical Yearbook”.

Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank

Vegas E. and I. Umansky (2005), “Improving Teaching and Learning through Effective Incentives” in Vegas E. (ed.) (2005), “Incentives to Improve Teaching”, Washington D.C.: The World Bank, 1-20.

World Bank (2002), “Household Expenditure Indicators” ”, Washington D. C.: World Bank, Development Data Centre.

World Bank (2003), “World Development Indicators 2003”, Washington D. C.: World Bank, Development Data Centre.

Appendix

Table A1: PISA 2003, Levels of Mathematics Proficiency

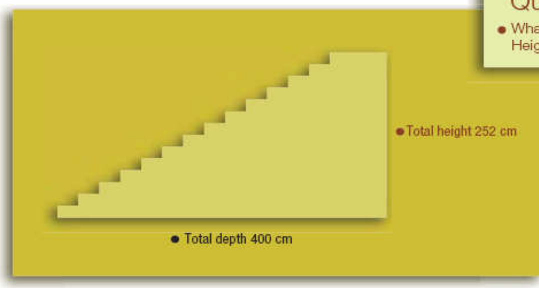
6	At Level 6 students can conceptualise, generalise, and utilise information based on their investigations and modelling of complex problem situations. They can link different information sources and representations and flexibly translate among them. Students at this level are capable of advanced mathematical thinking and reasoning. These students can apply their insight and understandings along with a mastery of symbolic and formal mathematical operations and relationships to develop new approaches and strategies for attacking novel situations. Students at this level can formulate and precisely communicate their actions and reflections regarding their findings, interpretations, arguments, and the appropriateness of these to the original situations.
5	At Level 5 students can develop and work with models for complex situations, identifying constraints and specifying assumptions. They can select, compare, and evaluate appropriate problem-solving strategies for dealing with complex problems related to these models. Students at this level can work strategically using broad, well-developed thinking and reasoning skills, appropriate linked representations, symbolic and formal characterisations, and insight pertaining to these situations. They can reflect on their actions and formulate and communicate their interpretations and reasoning.
4	At Level 4 students can work effectively with explicit models for complex concrete situations that may involve constraints or call for making assumptions. They can select and integrate different representations, including symbolic, linking them directly to aspects of real-world situations. Students at this level can utilise well-developed skills and reason flexibly, with some insight, in these contexts. They can construct and communicate explanations and arguments based on their interpretations, arguments and actions.
3	At Level 3 students can execute clearly described procedures, including those that require sequential decisions. They can select and apply simple problem-solving strategies. Students at this level can interpret and use representations based on different information sources and reason directly from them. They can develop short communications reporting their interpretations, results and reasoning.
2	At Level 2 students can interpret and recognise situations in contexts that require no more than direct inference. They can extract relevant information from a single source and make use of a single representational mode. Students at this level can employ basic algorithms, formulae, procedures, or conventions. They are capable of direct reasoning and making literal interpretations of the results.
1	At Level 1 students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined. They are able to identify information and to carry out routine procedures according to direct instructions in explicit situations. They can perform actions that are obvious and follow immediately from the given stimuli.

Source: PISA 2003, Technical Report (p.261)

Table A2: PISA 2003, Mathematics, sample questions

STAIRCASE

The diagram below illustrates a staircase with 14 steps and a total height of 252 cm :



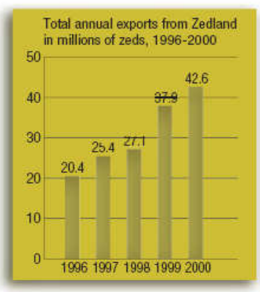
Question

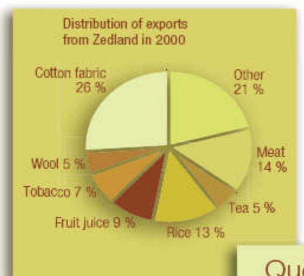
- What is the height of each of the 14 steps?
  Height: .... cm

This short response question is situated in a daily life context. The student has to interpret and solve the problem which uses two different representation modes: language, including numbers, and graphical. This question also has redundant information (i.e., the depth is 400 cm) which can be confusing for students, but this is not unusual in real-world problem solving. The actual procedure needed is a simple division. As this is a basic operation with numbers (252 divided by 14) the question belongs to the reproduction competency cluster. All the required information is presented in a recognisable situation and the students can extract the relevant information from this. The question has a difficulty of 421 score points (Level 2).

EXPORTS

The graphics below show information about exports from Zedland, a country that uses zeds as its currency.





This multiple-choice question is situated in a public context and is in the uncertainty content area. It consists of reading data from a bar chart and a pie chart, and combining that data to carry out a basic number operation. Specifically, it involves decoding the charts by looking at the total of annual exports of the year 2000 (42.6 million zeds) and at the percentage coming from Fruit Juice exports (9%). It is this activity and the process of connecting these numbers by an appropriate numerical operation (9% of 42.6) that places this question in the connections competency cluster. The question has a difficulty of 565 score points (Level 4).

Question

- What was the value of fruit juice exported from Zedland in 2000?
  A. 1.8 million zeds.
  B. 2.3 million zeds.
  C. 2.4 million zeds.
  D. 3.4 million zeds.
  E. 3.8 million zeds.

WALKING

The picture shows the footprints of a man walking.
The pacelength P is the distance between the rear of two consecutive footprints.

For men, the formula, $\frac{n}{P} = 140$, gives an approximate relationship between n and P where:

n = number of steps per minute, and
P = pacelength in metres.

Question

- If the formula applies to Heiko's walking and Heiko takes 70 steps per minute, what is Heiko's pacelength?
  Show your work.

This open-constructed response question is situated in a personal context. The question is about the relationship between the number of steps per minute and pacelength, which means that it is in the change and relationships content area. Students need to solve the problem by substitution into a simple formula and carrying out a routine calculation: if n/p = 140, and n = 70, what is the value of p? The competencies needed involve reproducing practised knowledge, performing routine procedures, the application of standard technical skills, manipulating expressions containing symbols and formulae, and carrying out computations. With this combination of competencies, and the real-world setting that students must handle, the question has a difficulty of 611 score points (Level 5).

Source: First Results from PISA 2003 (pp.6-7)

Table A3: PISA 2003, Number of students above the Advanced International Benchmark in Mathematics

Country	**% of Test Takers**	**Test Takers as % of cohort**	**Cohort Size**
India	52.3	11,503,247	21,994,737
Korea	97.2	681,426	701,056
Mexico	60.0	1,204,632	2,007,721
Slovak Republic	75.0	63,821	85,095
Thailand	71.2	727,055	1,021,145
USA	88.0	3,676,652	4,178,014
Country	**% of Test Takers >625**	**Test Takers >625 as % of cohort**	**Cohort Size**
India	1.00	95,659**	21,994,737
Korea	18.2	127,592	701,056
Mexico	0.29	5,822	2,007,721
Slovak Republic	9.42	8,016	85,095
Thailand	1.51	15,419	1,021,145
USA	6.52	272,406	4,178,014

**TIMSS 2003 estimate for India is 101,000, here it is adjusted by the US TIMSS/PISA ratio and is equal to 95,659. This is because India did not participate in PISA 2003, so that we are using the test scores directly comparable to the TIMSS 2003, developed by Das and Zajonc (2008) and we are adjusting these test scores in order to take into account the differences between TIMSS and PISA (i.e., TIMSS is taken by 4th and 8th grade students and has questions more closely related to the curriculum whereas PISA is taken by 15 year-old students and measures literacy in the subject)

Description of Table construction:

- % of Test Takers: gross enrolment in secondary education in the country (as net enrolment was not available for the Slovak Republic), (UNESCO Statistical Yearbook)
- Test Takers as % of cohort: percentage of test takers in the maths test as a share of the total number of 15 year-olds in the country, (PISA 2003)
- % of Test Takers>625: percentage of test takers who achieved a score>625 in the maths test, (UNESCO Statistical Yearbook, PISA 2003)
- Test Takers >625 as % of cohort: percentage of test takers who achieved a score>625 in the maths test as a share of the total number of 15 year-olds in the country (PISA 2003)
- Cohort Size: Total number of 15 year-olds in the country, (World Development Indicators)