

Unexpected learning competencies of Grades 5 and 6 pupils in public elementary schools¹: A Philippine report

Abraham I. Felipe
Fund for Assistance to Private Education, Makati City afelipe@aya.yale.edu

The present study tested the assumption of a positive and linear relation between years of schooling and school learning in the Philippine setting. It replicated a 1976 study that had cast doubt on this assumption in the Philippine public educational system. It tested three competing hypotheses for that finding: common sense, the 1976 arrested development hypothesis, and the alternative accelerated development hypothesis.

To test these competing hypotheses, two factors were systematically varied: the grade levels of Ss and the levels of the tests used. The competing hypotheses have different predicted outcomes. A total n of 7097 from 96 schools participated in the study.

The results showed that on all tests Grade 5 showed more competencies than Grades 4 and 6, although Grade 6 continued to perform better than Grade 4. When sub-test level was held constant in multiple comparisons, Grade 5 was learning more Grade 6 competencies, whereas Grade 6 was losing not only Grade 6 but also Grade 5 competencies. It is noted that whereas Grade 6 enjoyed a slight superiority in achievement scores circa 1976, the present study shows that Grade 5 enjoys an impressive superiority over Grade 6 circa 2003. That in Grade 6 one knows more competencies than in Grade 5 seems to be a myth. The common sense hypothesis has been ruled out. The results are consistent with the accelerated development hypothesis.

Assessment, Philippines, basic education, comparative education, learning competencies

BACKGROUND

It is generally assumed that the more the years of elementary schooling, the more school competencies pupils learn. However, a search in the literature yielded no report systematically testing this assumption. This could mean that the assumption is being left untested, or that the outcomes of tests had not been reported because confirmation proves nothing interesting whereas deviations from expectations could be embarrassing. Yet, deviations are not implausible, and cases of deviations could even offer interesting data for comparative studies of educational systems of different countries.

To begin with, an appropriate test of the assumption is not easy to design. The minimum requirement is to have only one measure of schooling outcome (say, a test) appropriate to the

¹ Esperanza C. Buen, Jesus E. Sevilla, Venia E. Pepito, Lourdes M. Franco and Angeles A. Sampang of the Center for Educational Measurement contributed valuable inputs to the study. The Center funded the study.

different grade levels to be studied. By using only one measure, performance of different grade levels could then be meaningfully compared.

The prevailing practice, however, is to construct one level of test (as a measure of schooling outcome) for each level of schooling. Under this condition, the meaning of comparisons of different schooling levels would be lost. Differences in scores among grade levels would be ambiguous: they could be due to differences in the tests used or in what pupils from different grade levels actually learned.

Another relevant tradition is the approach in grade or age placement. In this tradition, only one test may be used and a subject's score on this one test is then expressed in terms of the performance of the age group or the grade level whose average score best resembles his score.

However, although this tradition satisfies the condition of using only one test for classifying scores according to age groups or grade levels, the method needed in developing the test precludes being able to disconfirm the assumption of a positive linear relation between years of schooling and competencies learned.

In the age or grade placement tradition, the test that is considered appropriate is that which satisfies empirically a positive and linear relation between age or grade level on one hand and test performance on the other. In developing the test, test scores are regressed on age or grade levels, or age or grade levels regressed on test scores. The test developer then imposes the criterion that the relation must be positive and linear in order for the test to be accepted. With such a criterion for developing a test, it will not be possible to disconfirm the assumption in question. An empirical study claiming no positive linear relation between years and outcomes of schooling would be simply dismissed as a case of a bad test.

In the Philippines, there have been claims that actual competencies in Mathematics and Reading of senior high school students are at the elementary school level only. Such claims could mean a non-linear relation (for example, quadratic) between competencies and years of schooling. Or, that the relationship could be linear but the slope is so low that it takes ten years of schooling to learn what could be done in six. Such claims call for empirical demonstration of whatever relation exists between these two variables. On this point, no published works have been found. Although the provenance of the claims is unclear, they nonetheless persist among critics of the educational system.

The Philippines was one country that officially tested if the positive linear relation of schooling outcomes and years of schooling held out in its public elementary school system. In 1976 the Department of Education of the Philippine government reported a comparative study of the performance on achievement tests of 4th, 5th and 6th Grade pupils, wherein it was reported that Grade 6 pupils did not perform much more than Grade 5 pupils on same tests while the latter performed distinctly better than Grade 4 (EDPITAF, 1976). In short, the study suggested that the assumed positive and linear relation might not obtain in the Philippines. The official reaction of the government was expressed in the Foreword by the then Secretary of Education who said that the finding was "disturbing and cause us no reason to rejoice".

Surprisingly, however, in spite of its social significance, the study was neither ever replicated nor its findings confirmed. After the 1976 study, the Philippine practice in tracking growth in school learning returned to the "one test for one grade level" approach (NETRC, 2004). In the meantime, new concepts and technologies have been put in place to better manage the schooling system. One such concept is that of "school learning competencies" which has led to specifying competencies and skills to be learned at each grade level. As the Department of Education curriculum guide puts it, "after undergoing (blank) years of the Program, the pupil is expected to have developed the following (listed) competencies" (Department of Education, 2004). The notion of a cumulative

school learning is clear. This system of specifying target competencies and skills for every grade level makes it easier to evaluate whether the system has been performing as intended. At the same time, the technology of diagnostic testing has been spreading. This technology permits more precise measurements of school learning than when “school achievements” were the object of measurement.

The present study is a follow up of the 1976 SOUTELE study with several differences. First, it used learning competencies or skills as the data to be studied instead of the older and rougher concept of “school achievement” in vogue in the 70’s. Second, it used diagnostic tests for measuring the learning competencies or skills instead of achievement tests. For actual test items, it used competency items prescribed by the Department of Education for each grade level. For design, it used the subject as his own control in establishing whether a difference in test scores between grade levels actually existed. In the present study, this was done with the use of multi-level tests.

PROBLEM AND HYPOTHESES

The main SOUTELE finding consisted of a report that on nine of ten achievement tests on elementary school subjects, Grade 6 pupils did not score much more than Grade 5 pupils. SOUTELE reported that the differences between the two grades were so small for them to be educationally significant. It was as if Grade 6 did not make any difference in student learning at all. The conclusion was that learning beyond Grade 5 had been arrested (the arrested development hypothesis), at least during that developmental stage. Several mechanisms (none critical for the present study) were proposed to explain the arrest.

The alternative view is that the arrested development conclusion was premature. To explain, refer to Figure 1. The heavy line going upwards from left to right in Figure 1 represents the common assumption of a positive and linear relation between school learning and years of schooling. That is the theory. The slope was exaggerated for expository purposes.

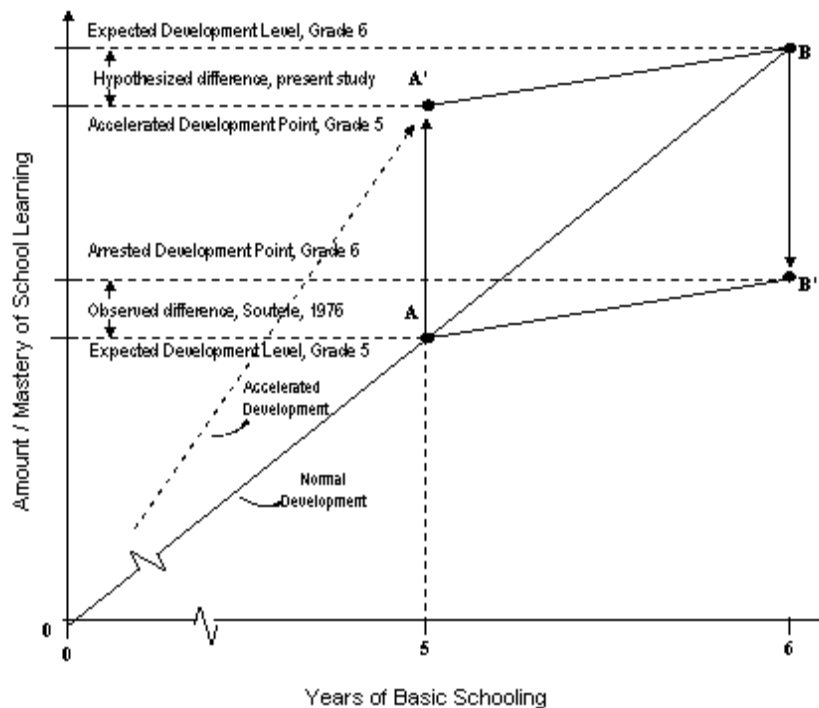


Figure 1. Relationship of schooling and learning

By Figure 1, learning among Grade 6 pupils as reported by SOUTELE did not proceed to Point B as predicted by the model but stopped at Point B' only, which was just a bit more than that expected for Grade 5. Hence the conclusion that their leaning had been arrested at the Grade 5 level.

The competing explanation however says that this reasoning is specious, since there was no independent evidence showing that Grade 6 pupils did not reach Point B. The alternative hypothesis is equally plausible: Grade 6 reached Point B as expected, but Grade 5 pupils learned more than they were expected, performed almost as well as the Grade 6 pupils and even reached Point A. Their learning, in other words, was accelerated (the accelerated development hypothesis) for which reason only a small difference was observed between Grades 5 and 6. From this viewpoint, the SOUTELE report should not be a cause to be disturbed but, instead, a cause to rejoice, using the language of the Secretary of Education.

In support of its position, the accelerated development hypothesis points out that the tests used were really Grade 6 tests, because they were pre-tested for the Grade 6 level and not for Grade 5. From this viewpoint, they were therefore the Grade 5 pupils who performed at the level of Grade 6, rather than the other way around.

Critical Tests of the Hypotheses

This reasoning suggests the value of using multi-level tests to determine whether it was the Grade 6 pupils who performed at the level of Grade 5, or the other way around. In a multi-level test, the items would be applicable to different grade levels instead of only one.

What was needed was a design that would enable one to say of Grade 6, as SOUTELE did, "They are **only** as good as Grade 5" (in order to support the arrested development hypothesis); or, alternatively, to say of Grade 5, as the accelerated development hypothesis would, "They are **even** as good as Grade 6".

To be able to make these conclusions, two types of items of learning competencies² were needed. One type consisted of Grade 5 competencies normed at the Grade 5 level (or simply, Level 5 test), and another type consisted of Grade 6 competencies normed at the Grade 6 level (or, Level 6 test, for short). The two types of items were mixed in one test form and all Grade 5 and Grade 6 pupils had to answer all items. Hence, the repeated measures design. To gain a better perspective on the issue, explore more aspects of elementary education, and provide better experimental control, items of Grade 4 were included in the mix.

Predictions

In this situation, the following are the respective predictions of the three hypotheses:

1. Common sense predicts that performance of Grade 6 will be as good as Grade 5 on Level 5 tests but will be better on Grade 6 level tests.

² Examples of competencies for the different subject areas prescribed by the Department of Education:

<i>English:</i>	Level 4:	"Uses the time expression/adverb of time to complete a sentence"
	Level 5:	"Selects the appropriate demonstrative pronoun to complete the given sentence"
	Level 6:	"Uses the positive form of adjective to make sentence complete"
<i>Mathematics:</i>	Level 4:	"Subtracts 5-digit numbers from 6-digit numbers"
	Level 5:	"Lists the prime factor of given numbers"
	Level 6:	"Adds decimals through ten thousandths"
<i>Science:</i>	Level 4:	"Explains how muscles cause body movement"
	Level 5:	"Cites the conditions that bring about changes in materials"
	Level 6:	"Identifies the nervous system and its major parts"

2. The arrested development hypothesis predicts Grade 6 will perform only as good as Grade 5 on both Level 5 and Level 6 tests.
3. The alternative accelerated development hypothesis predicts that even on Level 6 tests, Grade 5 performance will approach that of Grade 6.

METHOD

Design

The basic design was a 3 x 3 factorial, with repeated measures on one factor. The first factor was grade level (involving three levels – Grades 4, 5 and 6). The second factor was sub-test level (also three levels – Subtest Level 4, Level 5 and Level 6). Repeated measures were taken on the second factor. This basic design was replicated three times, once each for English, Mathematics and Science.

The Test Instruments

A battery of three multi-level tests, one each for English, Mathematics and Science, was used. Each test had 90 items of learning competencies prescribed by the Department of Education of the Philippine government for elementary curriculum. Each test had three subtests: one each for Levels 4, 5 and 6. Each subtest had 30 items. Thus, there were nine subtests altogether in the battery (see Table 1). For any given subject area, the items were jumbled by grade level and item difficulty. The tests were developed by a professional testing organisation.

Table 1. The Multi-level tests: Number of items

Subtest Level	Math Test	Science Test	English Test	Total
Level Subtest 4	30	30	30	90
Level Subtest 5	30	30	30	90
Level Subtest 6	30	30	30	90
Total	90	90	90	270

Subjects (Ss)

Ss were 7097 pupils from 96 schools of two administrative regions of the Department of Education (2320 Grade 4, 2350 Grade 5, and 2364 Grade 6). All 28 school divisions in the two regions participated.

Because the study called for competencies learned at the end of a certain grade level, testing was done at the beginning of the following school year. As a consequence, the Grade 6 pupils did not exactly come from the same schools as the Grades 4 and 5 pupils: “Grade 6” consisted of those enrolled as freshmen in the nearest public high schools (that is, nearest to the participating elementary schools) during the time of testing. Most of the high schools were regular public high schools, but five were science high schools contributing a total of 250 students to the sample.

RESULTS: COURSE OF GROWTH OF INDIVIDUAL COMPETENCIES FROM GRADE 4 TO GRADE 6

The most compelling findings on the comparative performance of Grade 5 and Grade 6 are graphically presented in Figure 2 (a, b and c) for the three English subtests, Figure 3 (a, b and c) for the three Mathematics subtests, and Figure 4 (a, b and c) for the three Science subtests. Each point in these figures represents the percentage of pupils in a grade level who answered correctly an item. Each line corresponds to one “competency” prescribed by the Department of Education for the elementary curriculum. Each line corresponds to one test item. A percentage, therefore, can be interpreted as the percentage of the pupils who “had” the competency.

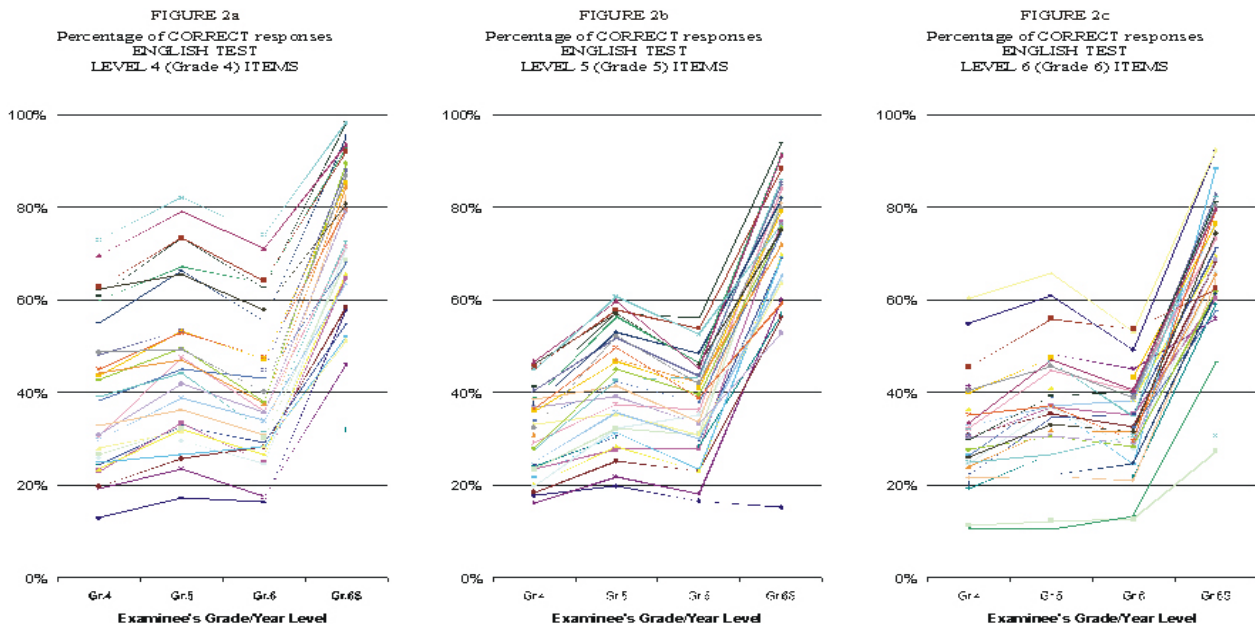


Figure 2. Percentage of correct responses for the three English subtests

Each subtest represents thirty competencies; therefore, each figure contains 30 lines. The grade levels of the pupils are shown in the horizontal axis. Grade 6 was divided into two, the “regular Grade 6” and the “science Grade 6”. A line that moves upward means that the percentage who answered the item correctly increased from one grade to the next. A line that moves downward means that the percentage decreased. For more impact, the reader is advised to look at each figure in its totality rather than scrutinise the details.

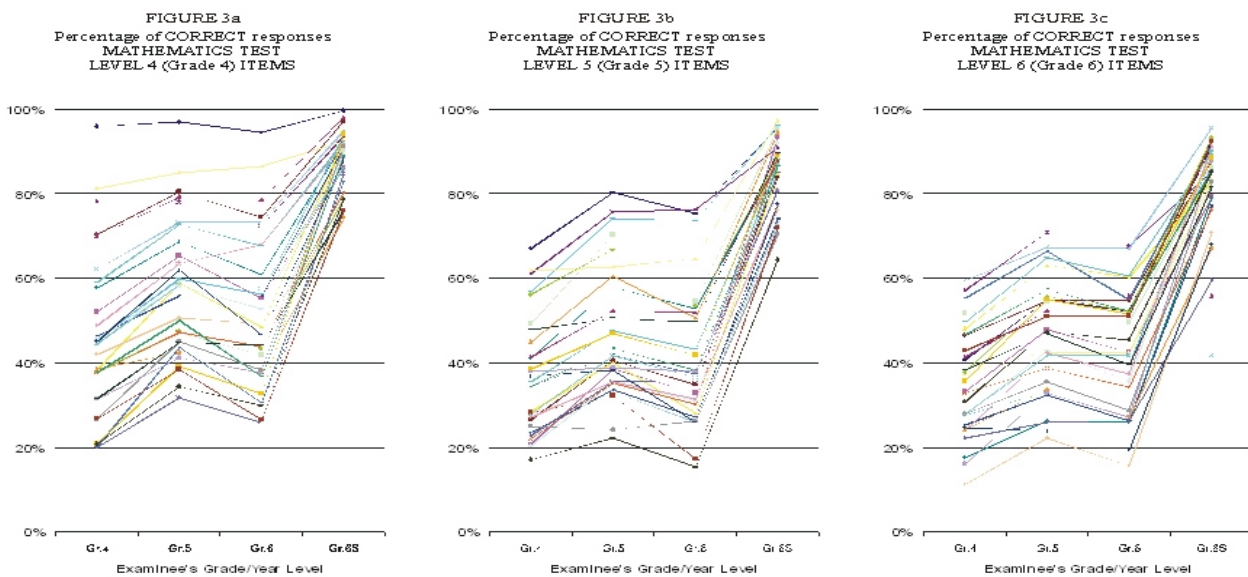


Figure 3. Percentage of correct responses for the three Mathematics subtests

In 261 out of the 270 competencies represented in the tests, Grade 5 scored more than Grade 4. This was expected. In the remaining nine competencies, the differences between the two grades were very minimal.

On the other hand, in 232 out of the 270 items, Grade 5 scored more than “regular” Grade 6. The number of competencies that registered decreases from Grade 5 to Grade 6 were 83 out of 90 cases in English, 75 of 90 cases in Mathematics and 74 of 90 cases in Science. Even with just the

Level 6 tests (see *Figures 2c, 3c and 4c* representing 90 items of Grade 6 competencies), Grade 5 scored more than Grade 6. Also, competencies gained in Grade 5 were subsequently lost in Grade 6. The growth of the competencies from Grade 4 through Grade 5 to Grade 6 followed an inverted “V” pattern. This was not expected.

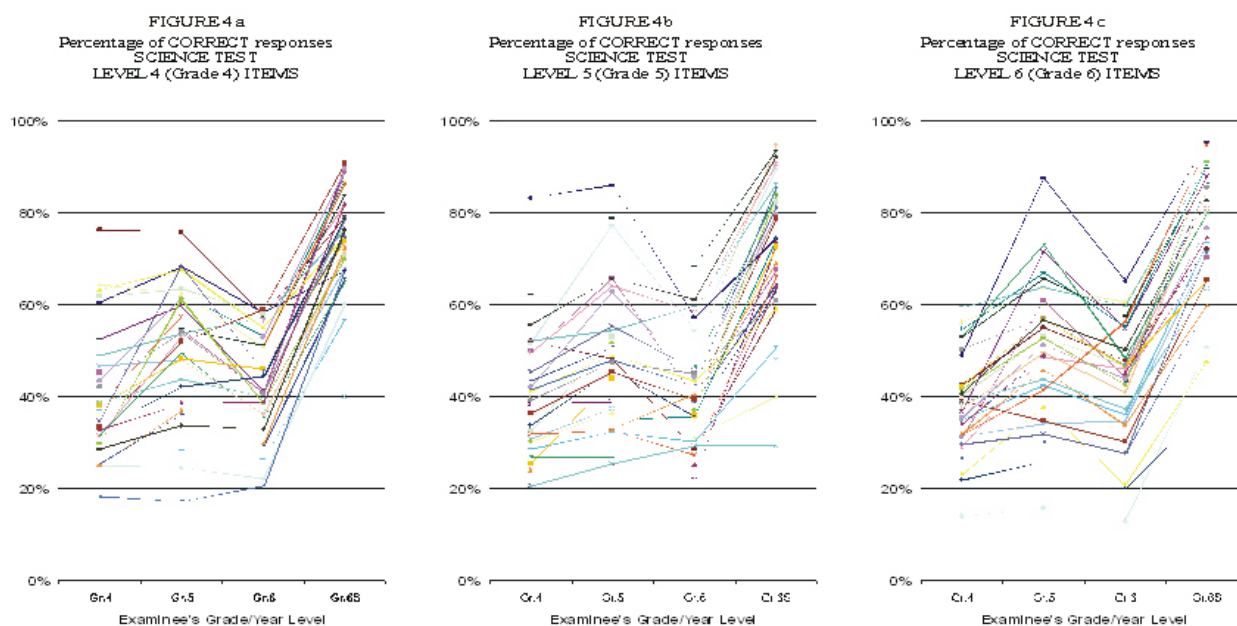


Figure 4. Percentage of correct responses for the three Science subtests

The science Grade 6 group was very distinct. They showed significantly more competencies compared to any of the groups studied, including the “regular” Grade 6.

F-tests showed a significant main effect attributed to Grade Level (for English, $F = 254.39$, $df = 2$, 6826, $p < 0.001$; for Mathematics, $F = 242.89$, $df = 2$, 6826, $p < 0.001$; and for Science, $F = 348.39$, $df = 2$, 6826, $p < 0.001$).

Similarly, there was also a significant main effect attributed to Sub-test level in all analyses (for English, $F = 5594.93$, $df = 2$, 13652, $p < 0.001$; for Mathematics, $F = 5052.37$, $df = 2$, 13652, $p < 0.001$; and for Science, $F = 501.84$, $df = 2$, 13652, $p < 0.001$). When subtest levels were held constant, multiple comparisons by grade levels using Duncan C revealed that Grade 5 had significantly higher scores than both Grade 4 and Grade 6 in all tests, including the crucial Level 6 tests. Unexpectedly, Grade 6 had significantly lower scores than Grade 5 even on Levels 4 and 5 tests. Although the scores of Grade 6 remained higher than those of Grade 4 in all comparisons, the results document an apparently ‘worsening’ Grade 6 *vis-a-vis* Grade 5 from 1976 when SOUTELE was published to the time of the present study.

These results were consistent with the accelerated development hypothesis and not consistent with both the arrested development hypothesis and common sense.

DISCUSSION

The Three Hypotheses and the Grade 5-6 Paradox

The common sense view that learning competencies increase with years of schooling was not supported by the results. This view might be correct in general, but it did not apply to Philippine public elementary education during 2003.

Circa 2003 in the Philippines, while Grade 5 had overwhelmingly more competencies than Grade 4 (expected), Grade 6 did not do as well compared to Grade 5. On the other hand, it was the

opposite. Not only did Grade 6 have less Level 5 competencies, but Grade 5 had more Level 6 competencies (not expected). That in Grade 6 one knew more than in Grade 5 seemed to be a myth. The reality of Philippine public elementary education reality was the opposite. That is why it is not inaccurate to call the finding a *paradox*.

The poor performance of Grade 6 did not seem to be a matter of Level 6 competencies being unable to develop beyond the Grade 5 level, as the arrested development hypothesis proposed. Compared to Grade 6, Grade 5 had more competencies of the Level 4, Level 5 and even Level 6 types. This means that competencies continued to be acquired (and not arrested) during Grade 5. Grade 5 pupils could answer Level 6 tests better than Grade 6 pupils themselves.

Present findings suggest (a) an above par Grade 5 education, and (b) a below par Grade 6 education. The performance of Grade 5 is unexpected; it deserves the official attention of the Department of Education.

A poorly performing Grade 6, the main point of the arrested development hypothesis, was first documented in 1976. After 29 years, there is no evidence that its performance *vis-a-vis* Grade 5 has improved. On the other hand, its present performance now seems to be even worse. Grade 6 is not what we supposed it to be. It had morphed.

Hypothesis for the Grades 5-6 Paradox

The more plausible explanation for the findings involved the operation of two mechanisms, one on the side of Grade 5 and the other on the side of Grade 6.

Hypothesis for the above-par Grade 5 performance

On the side of Grade 5, the mechanism involved an acceleration of the learning of Grade 6 competencies. A plausible factor for this is a change in the instructional process that takes place starting Grade 5. Beginning Grade 5 there is a shift from the "one classroom, one teacher" teaching assignment to the "departmentalisation" of teaching, wherein specialised teachers teach their respective subjects of specialisation. This could have led to a change in the way competencies are taught. As noted by SOUTELE itself, when the same teacher teaches the same subject area for both Grades 5 and 6, it may turn out that the teacher, even without intending it, would eagerly teach Level 6 competencies to Grade 5 pupils. This could account for the acceleration of learning in Grade 5.

Hypothesis for the below-par Grade 6 performance

On the side of Grade 6, the mechanism involved two components: depressing the learning of competencies "across the board" (that is, in several subject areas) and loss of competencies already learned.

The depression of learning "across the board" component is equivalent to the case of Grade 6 teachers not teaching enough of the Grade 6 curriculum to their pupils. This might be related to the practice of "teaching to the test", a practice wherein teachers spend a lot of time reviewing their students for various government-sponsored assessment tests throughout the school year, including the year-end National Achievement Test. Time for instruction is used instead for drills and practice for the examinations. This could lead to the acquisition of fewer Grade 6 competencies.

"Teaching to the test" could also lead to a loss of interest in and subsequent loss of competencies earlier learned because of the focus on "passing" the examinations. This then is what was registered as a loss of Grade 5 competencies among Grade 6 pupils. A longitudinal design could better study a hypothesis of loss.

However, even if the damage that “teaching to the test” might have inflicted is extensive and alarming, it is still repairable. Being rooted in a prescribed program, its adverse effects should disappear when the Department of Education would cancel the program in the future.

In 1976, Grade 6 still enjoyed a slight superiority in achievement over Grade 5; in 2003, it was the other way around except that the difference was no longer slight. This change signals that something undesirable was happening to the education of Filipino school children.

Need for Replication

Because of the serious social implications of the present study, it is suggested that it be repeated using a wider and more representative sample. The purpose of the replication should be to document the many other dimensions of the phenomenon, including its extent and magnitude and not to merely confirm it. It is expected that a more representative sample will magnify the Grade 6 depreciation that had been observed because the sample in the present study came from the more economically developed regions and better schools and was thus biased against finding a below par Grade 6 from the very start.

For an initial follow up of the implications of the present study, it might be better to look into all levels of the whole basic education ladder and for all the subject areas, and not only in Grades 4 to 6 English, Mathematics and Science. Education authorities should know if other grade or year levels are under-performing like the present Grade 6. They should know if the depreciation of Grade 6 is being recovered at all or not somewhere along one’s basic education. Studies should be done to determine if the depreciation means only a simple delay in formal learning. The authorities should evaluate if the system of schooling is seriously diminished by it. Those interested in compensatory education, should assess the place of special programs, such as the pre-baccalaureate or a Grade 7 program, catch-up approaches (such as lengthening instructional time), and even quick fixes (such as reviews) in recovering or compensating for the loss. Local governments should know the status of Grade 5-6 education in their jurisdictions, should they want to focus some local resources to the problem.

Curriculum

The present findings raise questions about some assumptions in current curriculum planning and design in the Philippine educational system. It appears that Grade 5 pupils could learn competencies intended for higher grade levels. It appears that the Department of Education has been setting curricular standards that are far too low. Filipino school children seem developmentally ready to learn competencies being assigned by curriculum makers for higher grades. Elementary pupils might even be already set to learn high school materials. The study raises issues about the curricula in teacher training institutions in order to make the basic education level more responsive.

The Below-Par Performance of Grade 6 and Philippine Development

The more serious implications are those that pertain to the larger issue of development. Education, including basic education, is viewed as a major tool for development. If the tool was defective, development must have been diminished. A defective system of basic education would not be able to perform the function for which it was designed. It would impair higher education and the correlated efforts to improve the lot of Filipinos. It will reduce the quality of human resources for local production and services, as well as for global competition. Grade 6 does not have the worth it is supposed to have in the market.

Of course, one could not yet tell the precise extent to which a below standard Grade 6 education has been affecting the larger development picture. However, because there is urgency in dealing

with it, one should not wait until his knowledge becomes full before acting. It is best to immediately look at the source of the problem and solve it there as soon as possible. Besides being expensive, quick fixes as well as catch-up compensatory programs might only mislead one to thinking that something is being done about the problem. Finally, just to ensure that the Grade 6 problem is kept engaged, the Philippines should monitor its system of schooling in terms of what children learn.

Improving education is standard concern for politicians and other national leaders. However, no current leader of note has ever shown awareness of SOUTELE's 29-year old report, that learning at the elementary school level was being held at the Grade 5 level. No one has ever inquired whether the situation has since changed. Politically, the issue has been completely neglected. Neglect might be why the situation seems to have become worse. The added concerns now are the vertical and long-term effects of a worsened situation. A Grade 6 level that performs below par might be related to the poor standing of the Philippines in the *Trends in International Mathematics and Science Study* (Gonzalez, *et al.*, 2004). It is also possible that the performance of other academic year levels has been dragged down.

REFERENCES

- Department of Education (2004), *Learning Competencies in English, Mathematics, and Science: Filipino and Makabayan*. Pasig City.
- EDPITAF (1976), *Survey of the Outcomes of Elementary Education, (SOUTELE)*. Department of Education, Manila.
- Gonzalez, P., *et al.* (2004), *Highlights from the Trends in International Mathematics and Science Study (TIMSS) 2003*. National Center for Education Statistics (NCES), US Department of Education.
- NETRC (2004), *2004 Annual Report of the National Educational Testing and Research Center*. Department of Education, Pasig City.