The Anti-Flynn Effect: Implications for teachers

During the CSMS research programme (1974-1980) we conducted a survey of around 12,000 students between the ages of 10 and 16, using three Piagetian tests, in order to find how their cognitive development varied with age.

As some of you may remember these tests were used as pre- and post-tests to check the effects of our various CA programmes (CASE, CAME, CASE@KS1, and Thinking Maths), usually successful compared to control schools or classes. However, round about 1998 we found that schools registering for professional development at King’s College University seemed to be becoming lower and lower in their apparent CSMS National percentile standings at their pre-test in Y7 and that the national standards of the tests had changed from what we had found in 1976.

This is where the Flynn-effect comes in: Flynn had reported that since the early 20s most major tests had shown higher and higher standardization results of between 0.5 and 0.8 standard deviations each decade, and so they had to re-standardize each 10 years or so because children were getting cleverer. His interpretation of this is that it reflects the gradual pressure that the move toward total industrialization in Western countries exerted on children for increasing competence on skills that IQ tests actually test. But our results ran contrary to that: that since 1995 they had been getting less able.

Michael and Flynn investigated this, and reported that, in the four Scandinavian countries and Britain, the Flynn effect had disappeared and that on all tests, as well as our Piagetian ones, national standards on all tests had got progressively worse each year, beginning in 1995 until at least the year 2004. Hence the adaptation to complete industrialization had been finished, and had now been replaced by another environmental pressure.

What significance do these results have for teachers today?

The initial CSMS research had found that in order for students to obtain C-grade or higher in science or maths they had to be at least at the early formal level by the end of Y9. This led to the CASE research intended to find if it were possible to accelerate the children's development to enable more students to achieve success in these demanding subjects.

Here first are the data for the Pendulum test at the end of Y8:

**Pendulum 2007**

<table>
<thead>
<tr>
<th>Year</th>
<th>%&lt;2B</th>
<th>%2B</th>
<th>%2B*</th>
<th>%3A</th>
<th>%3B</th>
<th>Test means</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>29.4</td>
<td>25.2</td>
<td>23.6</td>
<td>14.8</td>
<td>7.1</td>
<td>5.82</td>
</tr>
<tr>
<td>2007</td>
<td>18.65</td>
<td>29.7</td>
<td>40</td>
<td>8.60</td>
<td>3.05</td>
<td>5.93</td>
</tr>
</tbody>
</table>

Effect size  
-0.33sd  +0.12sd  +0.42sd  -0.34sd  -0.49sd  0.10 sd

The two striking results for us are first that the proportion of students showing formal operational thinking by entry to Y9 had halved since 1976 - remember the evidence from
Durham University that the difficulty of science and maths GCSE had been becoming steadily lower - and secondly that the proportion showing concrete operational thinking had markedly increased.

And the data for Equilibrium in the Balance at the end of Y9:

**EQUILIBRIUM IN THE BALANCE 2006**

<table>
<thead>
<tr>
<th>Year</th>
<th>%&lt;2B</th>
<th>%2B</th>
<th>%2B*</th>
<th>%3A</th>
<th>%3B</th>
<th>Test means</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>36.05</td>
<td>22</td>
<td>21.65</td>
<td>14.65</td>
<td>5.7</td>
<td>(5.82)pend</td>
</tr>
<tr>
<td>2007</td>
<td>15.73</td>
<td>41.05</td>
<td>38.05</td>
<td>4.3</td>
<td>0.825</td>
<td>5.84</td>
</tr>
<tr>
<td>Effect size</td>
<td>-0.61sd</td>
<td>+0.50</td>
<td>+0.44</td>
<td>-0.74sd</td>
<td>-1.1sd</td>
<td></td>
</tr>
</tbody>
</table>

Here the pattern of results are the same as for Pendulum, except that they are larger: the proportion at the 3A (early formal) and above had dropped from 20.35 to 5.2%. It is almost as though some environmental pressure were favouring concrete generalisation (2B*) thinking at the expense of formal thinking.

Before looking at the results for Volume & Heaviness, it is necessary to remind ourselves that the above two tests assess only ‘here-and-now’ thinking: the ability to manipulate the given data. Volume & Heaviness is quite different.

**Volume & Heaviness: 11/12 YEAR-OLDS**

<table>
<thead>
<tr>
<th>Year</th>
<th>&lt;2A</th>
<th>2A</th>
<th>2A/2B</th>
<th>2B</th>
<th>2B*</th>
<th>3A</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>4.2</td>
<td>7.9</td>
<td>23.6</td>
<td>30.2</td>
<td>25.8</td>
<td>8.3</td>
<td>5.15</td>
</tr>
<tr>
<td>2003/4</td>
<td>10.0</td>
<td>24.1</td>
<td>28.4</td>
<td>26.9</td>
<td>9.4</td>
<td>0.98</td>
<td>4.281</td>
</tr>
</tbody>
</table>

---

1 Piagetian scale is 2A = 3; 2A/2B = 4; 2B = 5; 2B* = 6; 3A = 7
This test was designed in 1976 to test every level of thinking from 2A (Early Concrete) to 3A (Early formal), that is from the average six-year-old upwards. Only 4 of the 15 items asked for here-and-now thinking about the relation between volume and weight (density). But the other 10 require no here-and-now work: they simply ask whether the children possess the conservations that normally are achieved between 4/5 and 8 years of age. Clearly as many as 34.1% of the 11/12 year-olds in 2003 are at the level of only the average 6/7 year-old (2A) on the conservations of quantity, weight and volume, compared with as few as 12.1 % in 1976. Something is diverting children away from the ordinary play experience that would have given them those conservations. In this figure of the boy/girl differences there is more evidence on the changes:

![Graph showing differences between boy and girl performance over years]

Between 1976 and 1994 there was no change whatsoever on the test’s standardisation at the 50th percentile. And there was still a boy/girl difference of 0.5 SD. Starting with 1995, both deteriorated, but boys faster than girls, so that by 2002 they had both become the same. When in 1976 the data were published it was suggested that the difference was due how they used their hours of spare time between the ages of 4 and 8. By 2002 this had become the same, and was continuing down at the same rate. So whatever be the environmental cause it had to (a) have begun in 1995, and (b) got larger each year, and (c) probably continued to get larger.

The same differential applied to the test Equilibrium in the Balance²: in 2006 a difference of only 0.02SD compared with 0.45 SD in 1976, both in favour of boys. Since this is a formal operational test the probable age-range of the effect on the boy/girl difference would be 9 years and upwards. This adds a note (d) to the last paragraph: the environmental cause would have been equally effective on children between 4 and 8, as between those 9 years old and upwards.

² There were never any sex-differences on Pendulum, I think because the abstract reasoning involved did not relate to any previous concrete experiences.
In Flynn and Shayer (2017, submitted) it was suggested that the only cause consistent with conditions (a) to (d) was already the subject of much sociological research. By the late 80s, for the older children computer games were already occupying the time of some, and around 1990 the positive Flynn effect began to drop. In 1995 two major games machines flooded the market, and apart from that, from the late 80s children were also beginning to spend more of their time watching television. Then there were simple mobile-phones, and more recently smart-phones. they even text during school lessons. 40 hours and upwards each week.

**How does this affect our work?**

I think this is for you to discuss. Compared with 1976, children leaving primary school now, and entering secondary, are nearer one Piagetian level lower in their cognitive development.

On page 5 are the original CSMS survey, alongside the results from Parkside School in Cambridge who were taught the original CASE. Someone who entered secondary school just at the 2B level would be expected to be just at the 2B* level at the end of Y8. But in that year-group in Parkside they would be just at the 3A level.

One suggestion I can make is that the same environmental effects that caused the drop in cognitive development since 1995 are still operative. It seems that they are exerting a positive pressure towards the concrete generalization (2B*) level: pushing some up, but actively pulling down others attempting any move towards the formal. Most, if not all computer games ask for this level of processing. So do most activities of texting. In-the-minute action can utilize quite a network of relevant aspects successfully but it can't model the connections between them. So what do you do? Do you still teach as usual (cursing), or do you put something in your lesson plans that would fight the environmental pressure?

Professor Michael Shayer, King’s College University, London

See also:


Results from Parkside School in Cambridge

Cognitive Development and Intervention:
based on CSMS survey data 1975-77

PIAGETIAN LEVEL

AGE (years)

Intervention period

Formal

Concrete

Upper 5%

Upper 10%

Upper 25%

Average

Lower 25%

Lower 10%

Lower 5%