The generational intelligence gains are caused by decreasing variance in the lower half of the distribution: Supporting evidence for the nutrition hypothesis

Roberto Colom\textsuperscript{a,},* Josep M\textsuperscript{a} Lluis-Font\textsuperscript{b}, Antonio Andrés-Pueyo\textsuperscript{b}

\textsuperscript{a}Facultad de Psicología, Universidad Autónoma de Madrid, 28049 Madrid, Spain
\textsuperscript{b}Universidad de Barcelona, Spain

Received 13 October 2003; received in revised form 12 May 2004; accepted 20 July 2004
Available online 21 September 2004

Abstract

Generational intelligence gains are one intriguing finding in science. Nutrition and cognitive stimulation are among the most remarkable causes of the upward trend in intelligence. The nutrition hypothesis predicts a primary impact on the most deprived, producing disproportionate gains at low intelligence levels. The cognitive stimulation hypothesis predicts gains along the intelligence distribution. However, data from the entire distribution are rarely available. The present study compares a sample of children tested in 1970 with an equivalent sample tested 30 years later. Data for the entire distributions were available. The results are consistent with the nutrition hypothesis, because the gains were mainly concentrated in the lower and medium halves of the distribution and were negligible in the very top half of the distribution. Moreover, an impressive gradual decrease in the gains was observed from the lower half to the top half of the distribution.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Intelligence; Generational intelligence gains; Intelligence distribution; Nutrition; Cognitive stimulation

Generational intelligence upward trends are among the most intriguing findings in science (Neisser, 1998). Several studies in different countries has shown that intelligence have been increasing during the
XX Century. These increases have been reported for the United States (Flynn, 1984), Japan (Lynn, 1982; Lynn & Hampson, 1986), Africa (Daley, Whaley, Sigman, Espinosa, & Neumann, 2003), Britain (Lynn & Hampson, 1989), Spain (Colom, Andrés-Pueyo, & Juan-Espinosa, 1998; Colom, Juan-Espinosa, & García, 2001; Colom & García-López, 2003), Australia, New Zealand, and several countries in Continental Europe (Flynn, 1987).

The size of the gains varies with the intelligence measure used. Tests of visuo-spatial abilities show greater rates of increase than those of verbal abilities (Neisser et al., 1996). Several causes of the gains have been suggested: the increased use of speeded tests (Brand, 1996), random agents (Jensen, 1996), improvements in cognitive stimulation (Flynn, 1999a), and nutritional changes (Lynn, 1990) are some examples. However, the fact that major intelligence improvements are from the domain of abstract problem solving ability on cultural-free tests presumably diminishes the relevance of several causes with a cultural and educational bases.

There are a couple of questions especially germane to find a plausible cause of the gains: Does the effect operates across the whole ability distribution? Is the effect operating equally on those of low, medium, and high ability, so that the whole intelligence distribution is being improved at a constant rate? These questions are extremely important (Rodgers, 1999).

If intelligence increases, mean scores increase as well. But a change in the population standard deviation of intelligence across time could also produce the same results that are the basis for the generational intelligence gains. The gains reported by Flynn (1984, 1987) are consistent with increased variability in the upper half of the intelligence distribution, and/or decreased variability in the lower half. These changes in variability will cause the mean to increase.

A given improvement that affected all people’s intelligence positively and equally would be reflected in an overall mean intelligence change without changing the shape of the distribution. An alternative improvement that reduced variability in the lower half of the distribution by improving intelligence for underprivileged people would also show an overall mean IQ change, but the intelligence in the top half would not change at all. The overall shape of the distribution would change considerably. These two effects could produce the intelligence gains, but distinguishing them is extremely important to understand the causes of the gains. Nevertheless, to know the facts, researchers should attend to the overall structure of the intelligence distribution (Rodgers, 1999).

Previous evidence exists supporting the reduction of variability in the bottom half of the distribution. First, Teasdale and Owen (1989) studied a Danish representative sample. The gains were observed among lower intelligence levels, finding no evidence of gains at the top half of the distribution. Second, Lynn and Hampson (1986) reported a double rise in the lower half than in the top half of the intelligence distribution.

The important point is that mean changes and variance changes can be occurring simultaneously. The distributional nature of these changes lies exactly at the nucleus of the understanding of what the generational intelligence gains are and what they mean. Rodgers (1999) wrote: “the issue of whether the effect is a fundamental shift on the mean of the overall distribution, or in the variance from a subset of the distribution is still open to future research (…) we will only begin to understand the effect when we interpret it at the distribution level” (pp. 350–351).

The present study was designed precisely with this main goal in mind. The data were collected by the second author in the city of Barcelona. The first sample was assessed in 1970 and the second sample was assessed in 1999. Both samples were highly equivalent (comparable) and data from the same IQ test were available for the analyses (see below).
Although there are several proposed causal explanations of the upward trend in intelligence (attitudes, schooling, genetic factors, random agents, biological factors, and so forth) we focus on two of the most frequently cited, namely, the nutritional and the cognitive stimulation hypotheses.

By “nutritional” we really mean that the increase in intelligence must be viewed as just one among a number of beneficial effects of a nurturing environment on physical development. Increased height and life span are some examples. The same causes are hypothesized: improved health, decreased rate of infant disease, or better nutrition. Both intelligence and height are supposed to be influenced by improvements in health care and nutrition. Therefore, the nutrition hypothesis predicts that improved nutrition would impact primarily on the most deprived and produce disproportionate gains at low intelligence levels (Flynn, in press). Those gains are expected especially for visuo-spatial abilities (see Lynn, 1990). This hypothesis predicts larger IQ rises in the lower half of the distribution.

By default, the hypothesis based in improved cognitive stimulation predicts a change for the entire distribution: “more leisure to enjoy cognitive challenges ranging from chess to video games. More professional work roles demanding lateral thinking. People were more disposed to invest more mental energy into problem solving for its own sake or at least problem solving on the spot” (Flynn, in press). There are no strong reasons to predict that those cognitive challenges impact primarily on the most deprived. People along the intelligence distribution would be affected by the greater availability of TV, radio, CDs, DVDs, computers, the Internet, and so forth.

1. Method

1.1. Participants

The fist sample comprised 459 boys and was tested in 1970. The second sample comprised 275 boys and was tested in 1999. Both samples were gathered using the same criteria. The students were from Barcelona. The boys for both samples were 7 years old.

Although no specific measure of SES was available, it was estimated according to the boys’ neighbourhood living standards, relatively easy to identify at the city of Barcelona. Low living standards were assigned to poor residence areas (so-called “sleeping cities”). Those residence areas are located around big factories and manufacture buildings. High living standards were assigned to rich residence areas (some examples were the districts of “Bonanova”, “San Gervasio”, and “Pedralves”). Finally, medium living standards were assigned to city areas between those of high and low affluence. Approximately 25% of the boys were of high SES, 25% of low SES, and 50% of medium SES.

1.2. Measures and procedures

The same IQ test was used to measure both samples: the Pressey’s Graphic Test (Calle, 1972; Ollacarizzqua, 1969). The test includes 80 items expressed in a visuo-spatial format. The items are divided into four blocks: verbal instructions, reasoning, classification, and spatial perception. Only the total score was available.

Because no specific information about the construct validity of the Pressey’s Graphic Test was available on the test manual, we tested 143 boys through the Pressey’s Test and the Culture Fair Intelligence Test (TEA, 1997). The computed Pearson correlation between the boys’ scores obtained in
both tests was +0.62 (p<0.000). Thus, the Pressey’s Test can be considered a reasonable proxy measure of fluid intelligence.

2. Results

Table 1 shows the descriptive statistics for both samples. The difference in $d$ units (the difference between the two means divided by the standard deviation) and IQ points is also shown. The mean difference is equivalent to 9.7 IQ points. Interestingly enough, the observed gain fits pretty nicely the rule of a gain equivalent to 3 IQ points per decade (Neisser et al., 1996). However, the interesting point concerns what has happened in the entire distributions. Table 2 shows the equivalence in a percentile scale of the total scores obtained by the 1970 and 1999 samples. Percentiles 1, 5, 15, 25, 35, 45, 65, 85, 95, and 99 are reported for illustrative purposes only. The results in Table 2 show that the gains decrease gradually from the bottom half to the top half of the distribution. The raw score difference between both samples falls from 9 to 1 point. It is noticeable that a monotonic decrease is observed.

Fig. 1 shows the distribution for both samples. The horizontal line represents the interval raw scores, while the vertical line represents the percentage of boys falling within the raw score interval for each sample.

Fig. 1 displays the fact that the gains are concentrated in the bottom and medium halves of the distribution. Furthermore, those gains decrease gradually from the bottom to the top half of the distribution.

Table 1
Mean, S.D., and N for the samples tested in 1970 and 1999

<table>
<thead>
<tr>
<th>1970 Sample</th>
<th>1999 Sample</th>
<th>$d$</th>
<th>IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>S.D.</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>50.53</td>
<td>7.14</td>
<td>459</td>
<td>54.8</td>
</tr>
</tbody>
</table>

The mean difference is also expressed in $d$ units and IQ points.

Table 2
Correspondence of raw scores obtained in 1970 and 1999 in a percentile scale

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>1970 Raw scores</th>
<th>1999 Raw scores</th>
<th>Raw scores’ difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>39</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>37</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>15</td>
<td>42</td>
<td>49</td>
<td>7</td>
</tr>
<tr>
<td>25</td>
<td>45</td>
<td>51</td>
<td>6</td>
</tr>
<tr>
<td>35</td>
<td>48</td>
<td>53</td>
<td>5</td>
</tr>
<tr>
<td>45</td>
<td>50</td>
<td>54</td>
<td>4</td>
</tr>
<tr>
<td>65</td>
<td>53</td>
<td>57</td>
<td>4</td>
</tr>
<tr>
<td>85</td>
<td>58</td>
<td>61</td>
<td>3</td>
</tr>
<tr>
<td>95</td>
<td>62</td>
<td>64</td>
<td>2</td>
</tr>
<tr>
<td>99</td>
<td>66</td>
<td>67</td>
<td>1</td>
</tr>
</tbody>
</table>

The difference between raw scores is shown in the last column.
Arthur Jensen points in his review of the present paper that it would be interesting to get the means and S.D.s of the four subtests of the Pressey’s Test, separately for the 1970 and 1999 samples, as well as the corresponding correlation matrices. However, as it was previously noticed, only the total score was available. Furthermore, Jensen suggests a factor analysis of the correlation matrices among the subtests for the 1970 and 1999 samples to know if the IQ gains involve the \( g \) factor. This is clearly both desirable and interesting, but there is not enough information in the present data sets to test Jensen’s suggestion.

3. Discussion

The present study shows that intelligence gains are mainly concentrated in the bottom and medium halves of the distribution. Moreover, the gains are progressively smaller as we move toward the upper half of the distribution. The tendency is impressive, as Table 2 and Fig. 1 show.

Flynn (in press) has proposed that the nutritional hypothesis can be supported only if a primary impact at the low intelligence levels is observed. However, Flynn embraces a sociological explanation, because in his own words “post-1950 improvements in obstetric and neonatal care have probably had no net-effect”.

We think that Flynn’s view would be germane for some data sets (see below). However, in addition to the results observed in the present study, there are two studies with the entire distribution that have reported results consistent with the prediction made from the nutrition hypothesis: Teasdale and Owen (1989) and Lynn and Hampson (1986). Taken together, the three studies show that intelligence gains are mainly (although not exclusively) concentrated among lower intelligence levels, with no evidence of large gains at the highest level, as the nutrition hypothesis predicts.

Colom et al.’s (1998) report is consistent with this evidence. They observed an upward trend in intelligence in Spain much greater on the Standard Progressive Matrices (SPM) than on the Advanced Progressive Matrices (APM) within a time interval of 28 years (19.2 and 6.75 IQ points, respectively). Thus, the gains were much higher for the SPM, which suggested that the cause of the increase probably had a greater impact in the low and medium segments of the intelligence distribution.
Furthermore, Colom and García-López (2003) found that Spanish high school graduates show a smaller increase than high school students in the Culture Fair Intelligence Test within a time interval of 20 years (4 and 6 IQ points, respectively). They suggested that whatever the cause of the rise in intelligence, it has its greatest effect on those at the lower and medium range of the distribution.

During the last half of the 20th century, there have been considerable improvements in living standards in the economically developed nations. These have allowed the populations to improve their nutrition and health care standards. If nutrition and health care are primary candidates to explain intelligence gains, then they could be expected to have an impact on other physically rooted factors, like height or brain size (Lynn, 1990). First, the secular increase in height over the last half of the 20th century in the Netherlands is equivalent to 1 S.D. (Roede & Van Wieringen, 1985). Second, the secular increase in brain size in Britain over the second half of the 20th century is equivalent to 1 S.D. (Whitehead & Paul, 1988). Third, the secular increase in intelligence over the last half of the 20th century is equivalent to 1 S.D. (Neisser et al., 1996). Fourth, there is a positive association between height and intelligence of approximately 0.20 (Jensen, 1998; Wilson et al., 1986) and between brain size and intelligence of approximately 0.40 (Rushton & Ankey, 1996; Vernon, Wickett, Bazana, & Stelmack, 2000).

The increases in height and brain size are mainly due to the improvements that have taken place in nutrition and health care. Several studies have shown that poor nutrition reduces brain size (see Winick, Rosso, & Waterlow, 1970). The positive association of height with brain size and intelligence could be the common effect of nutrition (and health care). The causal effect of nutrition on intelligence could operate through the neural development of the brain and could affect brain size and height, without brain size being causal to intelligence (Lynn, 1990).

However, it must be emphasized that although brain size and IQ have increased by about 1 S.D. over the years 1950–2000, suggesting that the same factor (nutrition) has caused increases of the same size, it could be argued that as brain size is correlated with IQ at 0.4, a 1 S.D. increase in brain size would produce 0.4 S.D. increase in IQ. Thus, other factors must have been operating, although the nature of these factors is still unknown.

Lynn (1990) gives several pieces of empirical evidence that demonstrate that visuo-spatial abilities are more sensitive to nutritional effects than verbal abilities. Remember that it is widely recognised that verbal abilities have been rising more slowly than visuo-spatial abilities (Neisser et al., 1996).

The most promising alternative theory to the nutrition hypothesis is that the increase in intelligence is derived from the improvements in cognitive stimulation. This would come from parents who are better educated, from smaller families, greater availability of radio, TV, and educational toys, and presumed improvements in education. However, improvements in cognitive stimulation would be expected to act especially (although not exclusively) on verbal abilities which are learned in families and schools (see Lynn, 1990). But this is not the case. The differential rates of increase of verbal and visuo-spatial abilities are not consistent with the main prediction from cognitive stimulation theory, but follow from the nutrition hypothesis. Furthermore, increases in intelligence have occurred among very young children (Hanson, Smith, & Hume, 1985). The increase at early ages suggests neurological development supporting the nutrition hypothesis once again.

The parallel increases that have taken place in height and brain size, the positive association between height, brain size, and intelligence, and the evidence that poor nutrition reduces height, brain size, and intelligence, are consistent with the nutrition hypothesis of the increases in intelligence.

Nevertheless, we must acknowledge that there is enough room to consider evidence not consistent with the nutritional hypothesis. Flynn (1998) states that “if enhanced nutrition is a factor, IQ gains over
time should come disproportionately from those below average IQs. Denmark fits that pattern but most nations do not. A good sign that IQ gains extend to every IQ level is that score variance remains unchanged over time, or diminishes only because of clear ceiling effects”.

However, it is important to consider the nature of the IQ measure. Thus, for instance, Flynn (1985) has studied the distribution for WISC to WISC-R. He found gains at all intelligence levels. But it is usually recognized that the Wechsler scales display a crystallized bias (“the Wechsler batteries...yield a verbal g”, Lynn, 1999, p. 9). Flynn (1999b) himself acknowledges this fact, stating that crystallised batteries are biased against generational intelligence gains.

Rushton (1999) analyzed the relationship between the g factor and upward trends in intelligence after the Wechsler scale, failing to find a positive correlation. Colom et al. (2001) replied Rushton’s study, but they did find a positive correlation between the g factor and intelligence gains after the DAT battery. Thus, there is no correlation between a crystallized g and generational gains, but the correlation between a fluid g and those generational gains is positive.

Furthermore, Colom, Abad, García, and Juan-Espinosa (2002) found that there is no association between the scientific construct of general intelligence (g) and the differences in IQ assessed by the Wechsler scale (WAIS-III). Wechsler Full Scale IQ does not directly measure g across the full range of the population distribution of intelligence.

Recently, Dolan, Colom, Abad, Wicherts, and Hessen (2004) replied this theoretically informative finding using the method of covariance and mean structure analysis (Dolan, 2000). They found that the g factor was not the dominant predictor of educational attainment, which suggests that IQ measures do not always reflect mainly g.

Given these considerations, it can be suggested that the cognitive stimulation hypothesis could not be tested directly against the nutritional hypothesis using IQ measures showing a crystallized bias. It is reasonable to state that high intelligence individuals are sensitive to cognitive stimulations acting on crystallised measures. However, fluid measures would be less sensitive to those educational factors.

The present paper shows that the gains are mainly concentrated in the low and medium halves of the intelligence distribution. It is imperative to remember that the items from the Pressey’s Test are based on a visuo-spatial format and that this test correlates reasonably well with the Culture Fair Intelligence Test. Therefore, this IQ measure can be considered a proxy estimate of fluid intelligence. The observed results are consistent with the nutritional hypothesis, but less so with the cognitive stimulation hypothesis.

In his review of the present paper, Pete Legree asks if increases in head size and height are concentrated among the smaller headed and the shorter. Unfortunately, we failed to find evidence to answer this question.1 Nevertheless, Colom and García-López (2003) reported height gains in Spain, Belgium, the Netherlands, and Norway that can be seen as consistent with the nutritional hypothesis. The height gain in Spain between 1960 and 1990 was 7.3 cm, in Belgium 4.4 cm, in the Netherlands 5.2 cm, and in Norway 2.5 cm. In 1960, Spanish conscripts had a lower mean height than those from these European countries: 6.1 lower than Belgium, 9.5 cm lower than the Netherlands, and 10.7 cm lower than Norway. However, in 1990, Spanish conscripts still had a lower mean height than those European

---

1 We contacted Stephen Buka, from the National Collaborative Perinatal Project. Only means, standard deviations, and univariate statistics for head circumference at 1 year of life were made available to us (from 1959 to 1966). Searching for height data, we contacted Gloria Quiroga, from the Department of Economics at the Universidad de Cantabria (Spain). Once again, they only report means and S.D.s. The available evidence indicates that height has increased 10 cm on average in Spain from 1955 to 1999, whereas height has increased 3 cm on average from 1893 to 1954.
countries, but the difference decreased: 3.2 cm lower than the Belgium, 7.4 cm lower than the Netherlands, and 5.9 lower than Norway. Therefore, a larger height gain for the shortest Spanish conscripts was observed.

In summary, the evidence reported in the present paper seems to be more consistent with the nutritional hypothesis than with the cognitive stimulation hypothesis. However, this is not to say that nutrition is the only causal factor. Surely, nutrition, health care, schooling, child rearing practices, and still other unidentified factors, are all involved. Anyway, it should be kept in mind that the important point is what of these agents is more important to account for the upward trend in intelligence. Rodgers (1999) has stated that research on the generational intelligence gains can help us understand the underlying causes of individual differences in intelligence. Nutrition and health care seem to be more important determinants of intelligence than is usually supposed. It has been assumed that cognitive stimulation is the major environmental determinant of intelligence, but the available results are disappointing (Neisser et al., 1996). Baumeister and Bacharach (2000) have written: “we insist that it is irresponsible to ignore scientific and clinical realities of the variable causative aspects of intellectual and learning differences, including those of biological origin” (p. 171, emphasis added). We think that the results reported in the present paper support the explanatory power of biological factors like nutrition and health care.

Acknowledgements

We would like to thank Prof. Richard Lynn and Prof. Mª Angeles Quiroga for their helpful comments to a draft of the present paper.

References


