Intelligence 57 (2016) 66–72

Contents lists available at ScienceDirect

Intelligence

# Lasting effects of quality of schooling: Evidence from PISA and PIAAC

Jan-Eric Gustafsson

*University of Gothenburg, Sweden*
*University of Oslo, Norway*

ARTICLE INFO

*Article history:*
Received 31 January 2016
Received in revised form 4 May 2016
Accepted 10 May 2016
Available online 17 May 2016

*Keywords:*
Quality of schooling
Age cohorts
Adult skills
PISA
PIAAC

ABSTRACT

The aim was to investigate to what extent quality of compulsory schooling is reflected in adult literacy and numeracy performance levels. Data from five administrations of the PISA survey between 2000 and 2012 for 20 countries were analyzed, along with data from corresponding age cohorts for the same set of countries participating in the cross-sectional PIAAC survey. For each country the PISA data was used to estimate linear achievement trends for literacy, numeracy and science to indicate change in quality of schooling and for the PIAAC data mean differences were computed between a younger and an older age group. The PIAAC performance differences were strongly related to the PISA achievement trends ($r = 0.70$), and relations held up when controls were introduced for level of education attained ($\beta = 0.55$) and for general social and cultural development of the country ($\beta = 0.48$). It is concluded that quality of schooling has lasting impact on adult literacy and numeracy performance levels.



## 1. Introduction

International comparative studies of educational achievement such as TIMSS (e.g., Beaton et al., 1996), PIRLS (e.g., Mullis, Martin, Gonzalez, & Kennedy, 2003) and PISA (e.g., OECD, 2001) have demonstrated large, and sometimes unexpected, differences in relative levels of achievement among different educational systems (Mullis & Martin, 2007). Moreover, substantial country differences in development of achievement over time have also been found. Hanushek, Peterson, and Woessmann (2012) linked the results from TIMSS, PIRLS and PISA and estimated the rate of change in achievement between 1995 and 2009 for 49 countries. Eleven countries showed annual growth, while student performance declined in nine countries. Hanushek et al. (2012) also demonstrated the changes to be unrelated to initial level of performance, thus neither supporting the hypothesis that low-performing countries catch up with higher performing countries, nor the hypothesis that high-performing school systems can build upon their past achievements.

However, little is known about the extent to which such achievement trends in compulsory school have lasting effects on adult levels of performance. The main aim of the present study was, therefore, to investigate if change over time in levels of skills at the end of compulsory schooling can be empirically connected to age cohort differences in skill levels among adults. This was made possible by the release of data from the Programme for the International Assessment of Adult Competencies (PIAAC) study (OECD, 2013), which is an international comparative study which investigates literacy, numeracy and problem solving skills in representative samples of the adult population. The PIAAC study is in many ways similar to PISA, but rather than sampling only 15-year-olds as is done in PISA, it covers the age range from 16 to 65 years. The first round of PIAAC was conducted in 2012, with 23 participating countries, almost all of which had also participated in PISA. Five rounds of PISA were conducted between 2000 and 2012. The participants in PISA 2000 were 27 years old when PIAAC was conducted, the PISA 2003 participants were 24 years old, and so on. To test the hypothesis that achievement trends in schooling are expressed as cohort effects in adult age, changes in PISA scores over time were related to age cohort differences in PIAAC.

### 1.1. Long-term effects of education

The amount of research on long term effects of quality and quantity of schooling is limited. However, a study of the effects of changes in length of schooling in six European countries found that longer compulsory schooling causally improved cognitive performance up to four decades later (Schneeweis, Skirbekk, & Winter-Ebmer, 2014). This study thus suggests that there are long-run effects of increased quantity of education. Some recent studies have also shown that there are long-run effects of improved quality of education. Chetty et al. (2010) used follow-up data from age 27 for students participating in the large class-size experiment STAR, and found, among other things, that students in small classes were more likely to attend college, and that students who had a more experienced teacher in kindergarten had higher earnings. Similar results were obtained in a Swedish study of both short- and long-term effects of class size in primary school (Fredriksson,

*E-mail address:* Jan-Eric.Gustafsson@ped.gu.se.

http://dx.doi.org/10.1016/j.intell.2016.05.004
0160-2896/

Öckert, & Oosterbeek, 2013). The study not only showed effects on ability and achievement at ages 13 and 16, but smaller classes also had positive effects on completed education, wages, and earnings at adult age. In a longitudinal study of mathematically talented adolescents Wai, Lubinski, Benbow, and Steiger (2010) investigated long-term effects of pre-college educational opportunities to study advanced science, technology, engineering or mathematics (STEM), for example in the form of special academic training, advanced subject matter placement, or taking Advanced Placement courses or college courses. The quasi-experimental study showed that those who had obtained more such opportunities achieved better real-world accomplishments in the STEM field in adult age, in the form of, among other things, PhDs, publications, and occupations.

More indirect lines of evidence also suggest long-term impact of levels of achievement at the end of compulsory schooling. Thus, economists have shown that country differences in school achievement correlate with country differences in economic growth (Hanushek & Woessmann, 2012). To the extent that this is a causal effect of quality and quantity of schooling it must be assumed that increased skills caused by improved schooling last to adult age. Hanushek and Woessmann (2012) conducted several analyses to investigate the tenability of a causal interpretation. In these analyses they demonstrated that the growth effects could be related to institutional characteristics of the school systems, that skills acquired by immigrants in their home countries affected earnings in the U.S. to the same extent as skills acquired in the U.S. by immigrants from the same country of origin, and, finally, that countries that improved their achievement over time also improved economic growth. On the basis of these results Hanushek and Woessmann (2012) concluded that economic returns come only from educational policies that improve student achievement and not from increasing the length of schooling without improving achievement. Unless the school achievement effects are long-lasting it also is difficult to imagine a mechanism through which they could have impact on economic growth. This, in turn, implies that changes in school achievement over time can be expected to be reflected in age cohort differences in levels of skills in the adult population.

Further support for the causal interpretation is obtained from a study reported by Rindermann (2008) who fitted cross-lagged path model to country-level longitudinal data spanning the period 1970 to 2000. The results showed that education and cognitive abilities were more important for economic wealth than economic wealth was for education and cognitive abilities.

The "Flynn effect" (e.g., Flynn, 1987, 2007) refers to increases in population levels of intelligence over time and it typically is assessed as changes over time in the level of performance of a particular age cohort, such as 18-year-olds. Several possible determinants of the Flynn effect have been proposed, such as schooling, changes in infectious disease prevalence, familiarity with cognitive performance tests, introduction of mass media, better hygiene, better nutrition and health, and smaller families (Neisser, 1998; Williams, 2013). However, few clear-cut conclusions about the effects of such factors have yet been established, even though increases in the amount of schooling have been identified as a major factor in this context too (see, e.g., Flynn, 2007; Skirbekk, Stonawski, Bonsang, & Staudinger, 2013).

Rosdahl (2014) reported results from a Danish longitudinal study, in which a subsample of PISA 2000 participants was followed up in PIAAC. Two main questions were investigated. One concerned to what extent reading literacy performance at age 15 predicts educational and occupational success at age 27, and the other to what extent there are changes in reading literacy skills between age 15 and age 27.

A strong association between PISA scores and choice of academically oriented education was found. Low scores were associated with a higher risk of not having a vocational education at age 27, and of receiving societal economic support. There also was a strong association between reading literacy scores in PISA and PIAAC, even though there also were changes in relative skill levels in PISA and PIAAC. Level of education was associated with an improved relative position, while those who had obtained economic support had a lower relative position in PIAAC than in PISA.

Rosdahl (2014) also observed a performance increase in the PIAAC data from age 16 to around 30 years of age. Such an increase can be due to an age effect or to a cohort effect, or to a combination of both. However, since the Danish PISA results have remained more or less stable between PISA 2000 and PISA 2012 Rosdahl (2014) concluded that the increase mainly is due to a positive effect of age on literacy. In a series of regression analyses Rosdahl (2014) showed that the strongest determinant of literacy scores was the number of years of education, and when both years of education and age were included in the analysis, the age effect was negative.

In summary, the Rosdahl (2014) results show that the PISA reading test predicts educational and occupational outcomes up to at least age 27, and that the PISA literacy level at age 15 is highly related to the PIAAC literacy level at age 27. The results also show that the educational level attained is a determinant of PIAAC literacy performance over and above the PISA literacy performance. However, given that the PISA and PIAAC results were measured on different scales, Rosdahl (2014) could not investigate changes in absolute levels of performance between the two measurement occasions.

### *1.2. Cognitive abilities at individual and national levels*

The present study takes advantage of the two international studies PISA and PIAAC and analyzes data from these at the country level. This raises the question what is actually measured in these surveys of educational achievement and adult skills. Rindermann (2007) compiled country-level results for studies of reading, mathematics and science (primarily PIRLS, PISA and TIMSS) conducted between 1991 and 2003, along with country-level estimates of IQ. Correlations between results achieved in the different domains, including IQ, were high, many values reaching above 0.80. A one-dimensional factor-model accounted for most of the variance and factor loadings were high: close to unity for the PISA tests, 0.97–0.99 for TIMSS, and 0.96 for IQ. While these results support the reliability of the international assessments, they also raise questions about the lack of differentiation between tests of school achievement and general cognitive ability.

Brunner (2008) used data from a German national extension of the PISA 2000 study, in which two measures of fluid intelligence (Gf) (Cattell, 1987) also were included to test the hypothesis that that a single domain-general g-factor would account for the relations among four mathematics subtests and three reading subtests. However, the one-factor model was rejected in favor of a nested-factor model comprising a g-factor identified by the two Gf tests along with two residual factors representing domain-specific mathematics and reading performance. These results thus show that at individual level a general cognitive factor is not sufficient to explain individual differences in mathematics and reading achievement.

Härnqvist, Gustafsson, Muthén, and Nelson (1994) estimated two-level CFA models at individual- and class-levels for a battery of cognitive tests administered to students in grades 4 to 9. At both levels general factors were identified, but the characteristics of the factors were different. At the individual level the general factor was defined by Gf-tests, and the model was in line with findings from previous investigations of the structure of cognitive abilities (Gustafsson, 1984, 1988). At the class-level, the general factor was defined by tests measuring crystallized intelligence (Gc) (Cattell, 1987), suggesting that different factors lie behind the structure at the aggregated level. Härnqvist et al. (1994) proposed three different mechanisms to account for this: (1) selection into schools on the basis of demographic factors; (2) differences between class-rooms with respect to quality of education; and (3) self-selection to more or less academically oriented programs in the upper grades. There still is little research available to determine the relative importance of these different factors. There are, however, results