Intelligence 41 (2013) 25–33

Contents lists available at SciVerse ScienceDirect

Intelligence

Problems in deriving Italian regional differences in intelligence from 2009 PISA data

Cesare Cornoldi [a], David Giofrè [a,*], Angela Martini [b]

[a] Department of General Psychology, University of Padova, Italy
[b] INVALSI, Frascati-Roma, Italy

ARTICLE INFO

Article history:
Received 26 March 2012
Received in revised form 3 September 2012
Accepted 10 October 2012
Available online 15 November 2012

Keywords:
International assessment programs
Intelligence
Educational achievement
IQ regional differences
Group differences

ABSTRACT

Recent results of international assessment programs (e.g., PISA) have shown a large difference in high school students' performance between northern and southern Italy. On this basis, it has been argued that the discrepancy reflects differences in average intelligence of the inhabitants of regions and is associated with genetic factors (Lynn, 2010a, 2012). This paper provides evidence in contrast to this conclusion by arguing that the use of PISA data to make inferences about regional differences in intelligence is questionable, and in any case, both PISA and other recent surveys on achievement of North and South Italy students offer some results that do not support Lynn's conclusions. In particular, a 2006–2009 PISA data comparison shows a relevant decrease in the North–South difference in only three years, particularly evident in the case of a single region (Apulia). Other large surveys (including INVALSI-2011) offer different results; age differences suggest that schooling could have an important role.

1. Introduction

Even though cognitive ability and academic achievement are distinct constructs and specific cognitive factors are important to explain specific aspects of achievement—not only the general factor (Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012)—it is unquestionable that measures of reading comprehension and mathematical achievement offer good approximations of the individual's intelligence levels. In fact, the linguistic, reasoning, working memory and attentional processes that underlie reading and mathematical operations also underlie intellectual functioning (Deary, Strand, Smith, & Fernandes, 2007; Hunt, 2011). The relationship is also supported by empirical evidence: Studies have found a good correlation between achievement tests (like SAT and ACT) and a g-factor measure, and these results are consistent because correlations are high (typically between .6 and .7) (Coyle & Pillow, 2008; Frey & Detterman, 2004; Koenig, Frey, & Detterman, 2008). Therefore, using achievement measures to derive IQ estimations is appropriate. As a consequence, some researchers have studied regional differences in IQ by taking advantage of the outcomes of the international assessment projects that have administered the same achievement tests in different countries (Rindermann, 2007).

Along this line of research, the comparison of the IQ of youngsters living in northern versus southern Italy has been seriously studied by international scholars, and the results have also been discussed in the popular Italian media. In particular, an influential and discussed study by Lynn (2010a) examined achievement scores obtained by southern and northern Italy students in the PISA2006 (Project for the International Assessment of Achievement) of students aged 15 (OECD, 2007) and associated the low scores obtained by southern Italy students with low intelligence levels. The study produced a series of other studies offering opposing arguments. In particular, Cornoldi, Belacchi, Giofrè, Martini, and Tressoldi (2010) reconsidered the results of the PISA2006 survey, which had been the basis for Lynn's conclusion, and other achievement studies and argued that North–South differences were not as

* Corresponding author at: Department of General Psychology, University of Padova, via Venezia, 8, 35131, Padova, Italy.
E-mail address: david.giofre@gmail.com (D. Giofrè).

0160-2896/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.intell.2012.10.004

clear as Lynn assumed (2010a). Beraldo (2010) raised methodological concerns while Felice and Giugliano (2011) stressed the relevance of socio-cultural factors. However, Lynn disputed the points raised by these studies (2010b, 2012). In particular, Lynn (2012) examined the achievement data obtained in the most recent PISA survey (OECD, 2010a) and offered counterarguments in favor of his thesis. In a latter paper, in agreement with the large body of evidence (e.g., Dick et al., 2007) showing the genetic bases of intelligence, Lynn also considered genetic differences between people living in northern versus southern Italy, further stressing the assumptions that there are strong differences in intelligence between them and that these differences are inherited. The issue was also examined by Templer (2012) who offered important data showing that both biological and social variables differentiating northern and southern Italy may explain the differences found in achievement. In the meantime, other papers were published on these issues. D'Amico, Cardaci, Di Nuovo, and Naglieri (2012) showed that regional differences may disappear using other intelligence testing procedures, and Robinson, Saggino, and Tommasi (2011), on the basis of different sources of information (obtained from INVALSI; Istituto Nazionale per la VALutazione del Sistema di Istruzione e di Formazione; National Institute for the Assessment of the Instruction System), showed that the achievement of southern Italy students may even be higher than that of northern Italy students.

In sum, the case of regional differences in Italy offers elements for the general discussion on ethnic differences in intelligence and its heritability versus modifiability by education. In fact, according to some authors (e.g., Ceci, 1991; Ceci & Williams, 1997), education and other environmental factors have substantial effects on IQ and academic achievement, and increments in school attendance convey substantial increments in intelligence. For example, a recent study indicates that two extra years of schooling beyond seventh grade have relevant effects on IQ above and beyond the Flynn effect, and the effect is substantial for students who are 19 years old (Brinch & Galloway, 2012). Nevertheless, since the appearance (1966) of the famous Coleman report, other researchers emphasized the role of IQ in self-selection into educational levels and provided support for the limited malleability of IQ by schooling and/or training (Herrnstein & Murray, 1994). Similarly, Lynn (2010a, 2012) argued that people from southern Italy have lower incomes and school levels because they are less intelligent and thus are less able to create favorable socioeconomic conditions for themselves.

At that point, the different theses could seem unfalsifiable and further studies comparing North and South Italy unproductive. Nevertheless, we think that reconsidering this point may have general implications for the debate on ethnic differences in intelligence (Hunt, 2011) and on the use of international data on achievement and thus can take advantage of the specific Italian case, for which more than a single source of evidence is available. In this paper, on the basis of the Italian data, we will show that i) it is risky to use PISA data to make inferences about the population's intelligence; ii) PISA 2009 data, if deeply analyzed and compared with the PISA 2006 data, offers a different picture than that derived by an overall North–South comparison; and iii) the outcomes from different sources of information about the achievement of Italian children offer different descriptions of the competencies of northern and southern Italy students.

2. Limitations of the PISA data for the international debate on intelligence

The PISA project is designed to evaluate education systems by testing the skills and knowledge of 15-year-old students in participating countries/economies. It has been argued that these measures are reliable and a good proxy of intelligence (e.g., Rindermann, 2007, 2008). Therefore, the use of PISA data may be ambiguous because it may be made both for assessing the efficiency of teaching and for deriving general ability measures. However, it must be taken into account that PISA studies originated for the need of educational assessment across countries and there is only clear evidence supporting this use. In fact, evidence supports the use of PISA in the context of national comparisons. For example, the results of PISA are highly correlated with the results of other achievement examinations (e.g., Trends in International Mathematics and Science Study [TIMSS], or Progress in International Reading Literacy Study [PIRLS]) (INVALSI, 2008a, 2008b).

PISA results also correlate with measures of intelligence (Lynn & Meisenberg, 2010; Rindermann, 2007). However, this evidence is open to criticisms. For example, according to Wicherts and Wilhelm (2007), this conclusion was based on aggregated-level analyses of correlations between means and cannot necessarily be interpreted at the level of individuals. In fact, in the case of PISA, data were collected to obtain information not about individual intellectual abilities but about groups. Furthermore data concerned academic achievement measures that, in a homogeneous population, may be highly related with ability measures, but in different populations and school systems may reflect educational system results, which, in the case of disadvantaged systems, may be improved, even of more than 1 standard deviation (Clarke, Snowling, Truelove, & Hulme, 2010) when appropriate teaching is introduced. The same goals reported in PISA documents specify that PISA is mainly intended to measure a contingent and modifiable efficiency of school systems: *"The design of PISA does not just allow for a comparison of the relative standing of countries in terms of their learning outcomes; it also enables each country to monitor changes in those outcomes over time. Such changes indicate how successful education systems have been in developing the knowledge and skills of 15-year-olds."* (OECD, 2010b, p.13)

The fact that the main goal of PISA is to assess the efficiency of the school system, not to make comparisons across individuals, is confirmed by the decision that participants must receive different tests. This is justified on the basis of the item response theory, but it makes comparisons difficult.

The outcomes of different programs assessing achievement seem only moderately correlated, and the correlations may be lower when intelligence and achievement scores are correlated (Baumert, Lüdtke, Trautwein, & Brunner, 2009; Kaufman et al., 2012). Therefore, examining the sources used for deriving the intelligence scores—which were correlated with achievement—is crucial. To our knowledge, these scores were mainly taken from Lynn and Vanhanen's database