The Version of Scholarly Record of this Article is published in INTELLIGENCE (2013), available online at: http://dx.doi.org/10.1016/j.intell.2012.10.004. Note that this article may not exactly replicate the final version published in INTELLIGENCE.

Cornoldi, C., Giofrè, D., & Martini, A. (2013). Problems in deriving Italian regional differences in intelligence from 2009 PISA data. *Intelligence, 41,* 25–33. doi:10.1016/j.intell.2012.10.004

Problems in deriving Italian Regional Differences in Intelligence from 2009 PISA data

Cesare Cornoldi*, David Giofrè*, & Angela Martini**

* Department of General Psychology, University of Padova, Italy;

** INVALSI, Frascati-Roma, Italy

Correspondence concerning this article should be addressed to:

David Giofrè,

Department of General Psychology,

University of Padova,

via Venezia, 8, 35131, Padova, Italy.

E-mail: david.giofre@gmail.com

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.

Keywords: International Assessment Programs, intelligence, Educational achievement, IQ regional differences, Group differences,

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 (Rinderman, 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 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 favour 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 significant 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 favourable 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 systems results, which, in the case of disadvantaged systems, may be substantially improved, even of 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."* (OECDb, 2010, 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-Vanhanen's database (2006), which offers useful preliminary information but also has many limitations (Hunt, 2011; Wicherts, Dolan, & van Der Maas, 2010).

As we have already argued, if PISA mainly assesses the quality of two school systems and the quality is dramatically different, then this difference may create confusion in the consideration of individual rather than school outcomes. Obviously, it can be argued (Lynn, 2010a) that the quality of the school system is related to the wealth of a region and that they are both the consequence rather than the cause of differences in achievement and intelligence. However, the opposite explanation is also legitimate. It can be argued that better school systems produce higher achievement levels because they usually provide a more favourable environment for fully achieving the students' potential. In fact, the quality of the school system (e.g., quality of teaching) has an important impact on academic success and academic achievement (Chetty et al., 2010; Rindermann & Ceci, 2009). In fact, some achievement effects can be attributed to factors other than intellectual gains, as stressed by Felice and Giugliano (2011) in their examination of the differences between North and South Italy. Checchi and Jappelli (2004), for instance, reported a quality score of public schools by regions (both as perceived by parents and as measured by indicators of school resources). They found a substantial regional variation and that the school quality is considerably lower in the South. By using aggregate indicators, they determined that the public schools (primary,

lower secondary, and upper secondary) in the South are lower in quality. For example, they reported that the proportion of students in double shifts due to school congestion is virtually nil in northern Italy, but may range between 6% and 14% in southern areas.

Recent data, independent of data based on achievement scores, confirms that northern Italy has a better functioning scholastic system (and in general has better public services, e.g., a better health care system) compared to southern Italy (Agasisti & Cordero, 2010). In particular, the difference among Italian regions in educational resources is relevant: Some regions suffer a shortage in resource quality (e.g., Sardinia) while others report values well above the OECD average (e.g., Lombardy) with a difference that can reach the value of .68 standardized points in the case of the WLE index (SCMATEDU from the PISA Questionnaire), which described the quality of educational resources. This index explained the 9.53% of the variance in the Science performance of Italian students (Agasisti, 2011; Agasisti & Cordero, 2010).

3. General Increase in Achievement of southern Italy Students

The observation of rapid changes in achievement and intelligence may help to clarify the impact of the quality of the school system on the level of achievement. On the basis of the Flynn effect (2009a; 2009b) and the assumption of a genetic basis of intelligence, one could predict that passing from one PISA administration to the following one, students' intellectual performances would slightly increase (but as the scores are standardized, they remain identical), and regional differences would be substantially preserved or would require long periods of time for minor changes. On the contrary, if we assume that results in achievement reflect the contingent and modifiable quality of a school system, we can explain rapid changes in achievement outcomes.

If we compare the scores obtained in different assessment procedures, we can see that

the size of the gap between different populations may be reduced very rapidly. For example, Nisbett et al. (2012) stated that the gap between whites and blacks diminished by 0.33 SD in recent years (but see, for a different conclusion, Rushton and Jensen, 2005). Similarly, the gap between northern and southern Italy children has become substantially smaller in the last few years (Table 1). In the last decade, the public national Institute INVALSI has moved from rough assessment procedures (which had offered the excessively positive description of southern Italy schools, used by Robinson et al. (2011) to more systematic and better controlled studies that showed the poor performances of southern Italian regions and motivated some of them to invest more resources in the education, and the growth in PISA performance can be attributed to this. Concerning the variation in performance from PISA 2006 to PISA 2009, Table 1 clearly shows that the southern Italy performance in PISA increased whereas the northern Italy performance remained similar.

Table 1 about here

This pattern is consistent across all the PISA areas (i.e., reading, math, and science) and is significant. Dividing the difference between the performances of different PISA administrations by 100 (the population's standard deviation) produces a measure of the growth expressed in SDs. The overall improvement in the last years for the south and islands is substantial. Data collected by OECD in 2009 (OECD, 2010b) allows a comparison for the different regions of Italy for PISA 2006 and 2009. As result, it is possible to establish the change in performance of the Italian regions participating in both PISA 2006 and PISA 2009. Table 1 presents a summary of the changes for the different areas of Italy. As can be seen, the Northwest also changed, but the change was more dramatic in South Italy. We derived a comparison from the original data that focused on the contrast between the most northern and most southern Italy regions for which both 2006 and 2009 data were available by treating different regions as subjects. For northern Italy, we included seven regions (Friuli, Trentino,

Lombardy, Veneto, Piedmont, Emilia-Romagna, and Liguria); for southern Italy, we included five regions (Basilicata, Campania, Apulia, Sardinia, and Sicily). The area of Bolzano was excluded even if it further supported the hypothesis of a reduction of the North-South gap because its substantial drop is due to the local professional schools (mainly attended by low achievement students) not being considered in the 2006 survey.

The pattern of results is robust across different observations. All the effects are significant, and the effect size is high (Cohen, 1988). In fact, 2 × 2 ANOVAs (year [Pisa 2006 and Pisa 2009]) × geographical area [North, South]) on the scores showed a significant interaction between year and geographical area, with large effect sizes and important decreases in the differences in score , either for the overall score, η_p^2= .45 (Figure 1), or for the specific scores in reading, η_p^2= .54, mathematics, η_p^2= .44, and science, η_p^2= .37. The decreases in the score differences between North and South, computed on the basis of the National 2009 standard deviations (100), were of .22, .18, .21, and .19 SDs respectively for the overall score, reading, mathematics and science. Furthermore, Bonferroni's post-hoc comparisons showed for the northern Italy regions no significant difference between PISA 2006 and Pisa 2009 in the four scores; on the contrary, for the southern Italy regions, there was a significant improvement in all the four scores of .27, .21, .30, and .21 SDs respectively (*p* ranging between .036 for Science and .003 for Mathematics).

Figure 1 about here

Using the PISA databank, we could also explore whether relevant school factors changed in different ways between North and South Italy between 2006 and 2009. To this purpose, we considered the following indexes: SC14Q04=Shortage qualified teachers; SC14Q07=Shortage science lab equipment; SC14Q08=Shortage instruct material; SC14Q09=Shortage computers; SC14Q10=Shortage Internet; SC14Q11=Shortage computer software; SC14Q12=Shortage library materials; SC14Q13=Shortage audio-visual;

IRATCOMP= index of availability of computers; and SCMATEDU=Quality of educational resources. However, we did not find in this short period changes in the differences between North and South that could explain the change in the achievement differences but rather only some paradoxical effects due to increases in the complaints about the availability of educational resources by southern schools that could be also interpreted as signals of an increased sensibility to the importance of them.

The presence of non-native students may affect the performance (typically of about 8 points in the PISA Italian sample), and this could have lowered the scores of northern Italy where migrants are more frequent and also may have been partly responsible for the variations of PISA scores in the last years. However, using the PISA databank, we could conclude that this effect should have been similar in 2006 and 2009 as the percentages of native students did not strongly vary. In fact, in 2006, the percentages of native students of the PISA sample were 92.13 for the North and 95.61at the South. In 2009, the percentages only changed slightly and were 90.53 and 97.58, respectively, a variation that could only explain a variation of less than .3 points in the PISA scores (the migrants who moved to the North could be a particularly intelligent group [Lynn, 2006]).

The improvement of southern Italy students observed by PISA 2009 is not isolated as it had already been anticipated by another international survey concerning literacy (i.e., Progress in International Reading Literacy Study (PIRLS); Mullis, Martin, Kennedy, & Foy, 2007) (Table 2) comparing fourth grade students in 2001 and 2006, respectively (INVALSI, 2008a, 2008b). The comparison shows large differences between northern and southern Italy regions (about 27 points) in 2001, and smaller differences in 2006 (about 9 points). Nevertheless, the reduction of the North-South Italy gap was not evident in another assessment that considered participants within an age range similar to PISA (Trends in International Mathematics and Science Study (TIMSS); Martin, et al., 2008; Mullis et al.,

2008). Table 3 offers an overall view of the standardized differences observed between northern and southern Italy students, showing that the variations in the different studies are so high to legitimate radically different conclusions, suggesting that further evidence is needed to reach unquestionable conclusions.

Tables 2 and 3 about here

4. The Case of Apulia

Apulia is an example of how contextual factors may dramatically affect PISA outcomes. In the Apulia region, there was an enormous growth in the performance in the last PISA survey (Table 4). The overall change from PISA 2006 to PISA 2009 for the Apulia is impressive (about 48 points), and if the PISA can be used as a measure of intelligence, it is equivalent to about 7 IQ points. Moreover, Apulia's performance is now more similar to the northern Italy regions than to the southern ones (Table 4). This finding cannot be explained on a genetic basis in such a short period of time nor as a simple case of regression towards the mean, as it was specific of Apulia and also partly predictable. However the result cannot be explained with the improvements in the quality of the school system, which could not easily produce such substantial changes in such a short period.

Table 4 about here

In our view, the impressive improvement of Apulia can be attributed to a specific program that the region initiated using the European Union's Social Fund for the development of poor areas and the improvement of the quality of the schools. Four Italian regions obtained funds (Campania, Calabria, Apulia, and Sicily), and in the case of Apulia, a large portion was spent on improving achievement and the ability to take achievement tests (by training teachers, providing additional resources, and sensitizing on the importance of national examinations) (Rubinacci, 2011). In fact, in 2006, Lucrezia Stellacci was nominated as the

new Director for the Regional School System (Ufficio Scolastico Regionale) of Apulia and decided to use the European funds to improve not only the regional school system (quality of teaching, buildings, labs) but also to expand assessment practices in Apulia schools and develop test taking skills in its students (Rubinacci, 2011). The dramatic improvement in the achievement of Apulia students seems therefore to be due to improvements both in the school system and in the test taking skills (Martini, 2011).

The importance of test taking skills and the difficulties that southern Italy students have with group administration of written tests is further supported by the curious observation of Cornoldi et al. (2010) that some regional differences currently present in the results of group assessments may disappear if assessment is individual and interactive. Consequently, we have re-examined the mean scores reported by Cornoldi et al. (2010) for the reading and mathematics tests administered in groups versus individually. We have transformed them to z-scores based on normative data and then computed the difference in the mean z-score obtained by North and South Italy. We have found that a difference between North and South in the group average scores present both in 9th- and 10th-graders (.32 and .50, respectively) completely disappears in the case of the average scores obtained at the individual testing (.02 and -.08, respectively). It is true that group testing concerned tasks (reading comprehension and mathematical reasoning) that have a higher relationship to IQ than do the tasks that were individually administered (reading decoding and calculation), but—based on the assumption of strong regional differences—the differences had to be reduced and not eliminated.

5. Outcomes of the 2011 INVALSI Survey

The data collected by the official public Italian Institute charged with gathering data on achievement of Italian students INVALSI (Istitituto Nazionale per la VALutazione del

Sistema di Istruzione e di Formazione; National Institute for the Assessment of the Instruction System) appear particularly critical and authoritative and are typically used also in international studies. For example, Lynn (2012) relied on recent INVALSI (2011) data to find further evidence on Italian regional differences.

INVALSI data is consistent and reliable and open to international scholars. The 2011 survey included about 40,000 students for each grade that was involved (2nd, 5th, 6th, 8th, and 10th) who were representative of the Italian population. Moreover, data is corrected to control for cheating, and an external INVALSI examiner is present during the examination (with the exception of the 8th grade, in which the test is taken in the context of a diploma examination). Considering the final report for 2011 data (INVALSI, 2011), Lynn (2012) found further support for the North-South differences, showing that there is a significant difference between northern and southern Italy regions. However, the official INVALSI report only offers mean values and interval confidences. Therefore, due to the large sample size, it is possible that a significant effect is detected even if it is very small and does not reflect large differences between groups (Cohen, 1988). For this reason, we took advantage of the original INVALSI databank. We computed the SDs and were able to calculate the size of the difference (expressed in terms of mean differences by the Italian standard deviations) (Table 5; we excluded the 8th grade because there was no an external examiner during the evaluation). From Table 5, we can see that if we compare the mean scores of northern and southern Italy regions, North-South differences are relatively small and lower than in the PISA scores, ranging between -0.02 and 0.32 (M = 0.15).

Table 5 about here

6. Age Changes

Despite the fact that intuitively experience and schooling should have had a greater

impact after many years of exposure than after fewer years and then have a higher influence on older children, inherited differences emerge more clearly in older children than in younger children (Petrill et al., 2004). Therefore, the hypothesis of an inherited nature of Italian regional differences finds further support from the observation that achievement scores obtained by southern Italy children decrease with development. In fact, international assessment projects have even reported no regional differences in Italy for 4th graders, as occurred for PIRLS (INVALSI, 2008a), but substantial differences for older children. Furthermore, according to TIMSS data (INVALSI, 2008b), the southern Italy regions have a drop in performance of about 40 points (roughly .4 SD) from the 4th to the 8th grade. A similar pattern is apparently present for the recent INVALSI survey (Table 5).

But there are some inconsistencies in the data. First, if the drop in the achievement test was due to genetic factors, then the drop observed by the TIMSS project would be limited to South Italy only. Instead, there was also a drop in performance in the North in some achievement tasks—about 20 points for math and 30 points for science (INVALSI, 2008b). Second, data on the increased North-South Italy gap is partly contradictory. INVALSI (2011) data show a large difference in the 10th grade, but the size of the difference is also very large between the 5th and 6th grades (Table 5). This result is inconsistent with the assumption of a gradual increase with age in the role of genetic factors but is consistent with the shared opinion that in Italy, especially South Italy, the quality of the secondary school for the 6th to 8th grades is poor (Ferrer-Esteban, 2011). In fact, in Italy, the transition from the 5th to the 6th grade corresponds to a dramatic change in the type of school, and the new educational system is poor in many respects (e.g., teachers without any specific pedagogical and psychological training, higher emphasis on the learning of notions than on general abilities). This is in line with the fact that Italian teachers of this school system (6th to 8th grades) are characterized by a low self-perceived competence and by a high mean age

(around 52), the highest in the OCSE countries (Ferrer-Esteban, 2011).

7. Discussion

The hypothesis that relevant differences in intelligence between different populations can be found and that these differences are inherited has a long history in psychology and has found fresh support on the basis of international assessment projects that have administered exactly the same procedures to different populations. In particular, the PISA program has offered an impressive amount of data that can be used to respond to crucial questions about human abilities. Lynn's (2010a; 2012) focus on the comparison between North and South Italy had the advantages of comparing two populations that were administered exactly the same tasks in the same language and with the same school legislation. The two populations are apparently significantly different not only on a historical but also on a biological basis as South Italy is close to African communities and immigrations and North Italy has close exchanges with Central Europe. Therefore, the comparison between North and South Italy represents not only an important issue per se but also a crucial case for examining the more general issue on the existence of ethnic differences in intelligence and the possibility of using international assessment projects to assess not only the quality of school systems as the explicit goals of these projects state, but also intelligence.

In this paper, we have offered arguments in favour of two main points. First, the use of PISA data to make inferences about regional differences in achievement and intelligence levels raises a series of problems. Second, even if we use PISA data, the North-South Italy differences are not as clear as they may seem at first glance. Concerning the first point, in our opinion, PISA data must be used cautiously because it is collected in a particular way that permits the assessment of how students of different school systems are able to perform on achievement tests but does not offer direct and reliable measures of the abilities of single

individuals. The cases of Apulia and of group versus individual testing show that particular conditions may affect the levels of performance. The partially different outcomes of different assessment projects (ranging from no difference in individual assessment to a difference of more than half a standardized point, roughly corresponding to a difference in IQ of 7–8 points in the PISA scores) further support the hypothesis that PISA data is also affected by contingent factors. Furthermore, even when basing conclusions on PISA scores, the deductions that we can derive do not favor strong genetic differences between North and South Italy.

With the present paper we did not intend to question the general issue of heritability of intelligence, as, in our view, the evidence supporting the genetic bases of intelligence is robust and unquestionable, but we wanted to put in evidence the fact that measures related with intelligence may be affected by a series of other factors. As Hunt (2012) commented, national indicators of intelligence are markers of national differences in the ability to use the cognitive artifacts (i.e., physical instruments or styles of reasoning that amplify our ability to think); further, variations in national capabilities to use cognitive artifacts can be attributed to differences in the extent to which different nations provide techniques and institutions for the development of individual cognition. In particular, our criticism concerns the conclusion that regional differences in academic measures related with intelligence are due to genetic factors. Further, a genetic hypothesis cannot explain why in a limited number of years, the North–South difference dropped substantially. A genetic hypothesis may also have difficulty in explaining some variations in age in achievement differences, which are impressively related to the shift from primary to secondary school, a type of school system that in Italy appears particularly weak.

We are aware that the nature versus nurture controversy in intelligence cannot find a complete solution because the arguments in favor of a position can be reversed and used in

favor of the opposite position, and the present evidence is not decisive against Lynn's theory. However, we think that the mediating role of the environment should be better emphasized when examining Italian regional differences, as also suggested for other populations (Barsky, Semin, & Malykh, 2011; Molenaar, van der Sluis, Boomsma, & Dolan, 2011; Rodic et al., 2011). Therefore, to examine Italian regional differences in intelligence, if they really exist, it should be necessary to make assumptions concerning the degree to which there are additional "achievement specific" population differences that may bias the estimate of the mean IQs.

References

Agasisti, T. (2011). The efficiency of Italian secondary schools and the potential role of competition: a data envelopment analysis using OECD-PISA2006 data. *Education Economics*, (931809330), 1-25. doi:10.1080/09645292.2010.511840

Agasisti, T. & Cordero, J. M. (2010). Which factors affect students' performances?Exploring differences among Regions in Italy and Spain by using PISA2006 data and multilevel modelling. *Paper prepared for the First Lisbon Research Workshop on Economics and Econometrics of Education. Lisbon, Portugal.*

Barsky, P., Semin, A., &Malykh, S. (2011, December). *The heritability of verbal and nonverbal intelligence in Russian adolescents.* Paper presented at the 12th annual conference of ISIR, Limassol, Cyprus.

Baumert, J., Lüdtke, O., Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement?. *Educational Research, 4*, 165-176. doi:10.1016/j.edurev.2009.04.002

Beraldo, S. (2010). Do differences in IQ predict Italian north–south differences in income? A methodological critique to Lynn. *Intelligence, 38*, 456-461. doi:10.1016/j.intell.2010.06.007

Brinch, C. N., & Galloway, T. A. (2012). Schooling in adolescence raises IQ scores. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 425-30. doi:10.1073/pnas.1106077109

Ceci, S. J. (1991). How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology*, *27*, 703-722. doi:10.1037/0012-1649.27.5.703

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, *52*, 1051-1058. doi:10.1037/0003-066X.52.10.1051

Checchi, D., & Jappelli, T. (2004). School Choice and Quality. CEPR Discussion Paper No. 4748. Retrieved from: http://ssrn.com/abstract=654564

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. H., &Yagan, D. (2010). *How does your kindergarten classroom affect your earnings?* Evidence from Project Star. Cambridge, MA: National Bureau of Economic Research.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence ErlbaumAssociates

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, F., Mood, A. M., Weinfeld, F. D., et al. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office

Cornoldi, C., Belacchi, C., Giofrè, D., Martini, A., & Tressoldi, P. (2010). The mean Southern Italian children IQ is not particularly low: A reply to R. Lynn (2010). *Intelligence, 38,* 462-470. doi:10.1016/j.intell.2010.06.003

Coyle, T. R., &Pillow, D. R. (2008). SAT and ACT predict college GPA after removing g. *Intelligence, 36,* 719-729. doi:10.1016/j.intell.2008.05.001

D'Amico, A., Cardaci, M., Di Nuovo, S., & Naglieri, J. A. (2012). Differences in achievement not in intelligence in the north and south of Italy: Comments on. *Learning and Individual Differences, 22,* 128-132. doi:10.1016/j.lindif.2011.11.011

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35,* 13-21. doi:10.1016/j.intell.2006.02.001

Dick, D. M., Aliev, F., Kramer, J., Wang, J. C., Hinrichs, A., Bertelsen, S., Kuperman, S., et al. (2007). Association of CHRM2 with IQ: converging evidence for a gene influencing intelligence. *Behavior Genetics, 37,* 265-72. doi:10.1007/s10519-006-9131-2

Felice, E., & Giugliano, F. (2011). Myth and reality: A response to Lynn on the determinants of Italy's North–South imbalances. *Intelligence, 39,* 1-6. doi:10.1016/j.intell.2010.09.004

Ferrer-Esteban, G. (2011). *Beyond the traditional territorial divide in the Italian Education System. Effects of system management factors on performance in lower secondary school*. Retrieved from: www.fga.it

Flynn, J. R. (2009a). Requiem for nutrition as the cause of IQ gains: Raven's gains in Britain 1938–2008. *Economics and Human Biology, 7,* 18–27. doi:10.1016/j.ehb.2009.01.009

Flynn, J. R. (2009b). The WAIS-III and WAIS-IV: Daubert motions favor the certainly false over the approximately true. *Applied Neuropsychology, 16,* 1–7. doi:10.1080/09084280902864360

Frey, M. C., & Detterman, D. K. (2004). Scholastic Assessment or g? The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychological science, 15,* 373-8. doi:10.1111/j.0956-7976.2004.00687.x

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York, NY: Free Press.

Hunt, E. (2011). *Human Intelligence*. New York, NY: Cambridge University Press.

Hunt, E. (2012). What makes nations intelligent? Perspectives on Psychological Science, 7, 284–306, http://dx.doi.org/10.1177/1745691612442905.

INVALSI (2008a). *Ricerca Internazionale IEA PIRLS 2006. La lettura nella scuola primaria*. Rapporto Nazionale. *[PIRLS 2006. Litteracy in primaryeducation. National repeport]*. Roma, Italy: Armando

INVALSI (2008b). *Rapporto Italia TIMSS 2007. Prime valutazioni sugli apprendimenti degli studenti italiani. [First assessments on Italian students learning]*. Retrieved from www.invalsi.it/ric-int/timss2007/rapporti.php

INVALSI (2011). *Servizio Nazionale Di Valutazione 2010-11 [National evaluation service 2010-11]*. Retrieved from www.invalsi.it/invalsi/istituto.php?page=rapporti.

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence*. *40*, 123-138. doi:10.1016/j.intell.2012.01.009

Koenig, K., Frey, M., &Detterman, D. (2008). ACT and general cognitive ability. *Intelligence, 36,* 153-160. doi:10.1016/j.intell.2007.03.005

Lynn, R. (2006). *Race Differences in Intelligence: An Evolutionary Analysis*. Augusta, GA, US: Washington Summit Publishers

Lynn, R. (2010a). In Italy, north–south differences in IQ predict differences in income, education, infant mortality, stature, and literacy. *Intelligence, 38,* 93-100. doi:10.1016/j.intell.2009.07.004

Lynn, R. (2010b). IQ differences between the north and south of Italy: A reply to Beraldo and Cornoldi, Belacchi, Giofre, Martini, and Tressoldi. *Intelligence, 38,* 1-5. doi:10.1016/j.intell.2010.07.001

Lynn, R. (2012). Intelligence IQs in Italy are higher in the north: A reply to Felice and Giugliano. *Intelligence, 40,* 255-259. doi:10.1016/j.intell.2012.02.007

Lynn, R., & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence, 38,* 353-360. doi:10.1016/j.intell.2010.04.007

Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Augusta, GA: Washington Summit Publishers.

Martin, M.O., Mullis, I.V.S., Foy, P., Olson, J.F., Erberber, E., Preuschoff, C. et al. (2008a). *TIMSS 2007 International Science Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martini, A. (2011). OCSE-PISA 2009: è vero che stiamo migliorando? [OCSE-PISA: is it true that we are improving?]. *Rivistadell'Istruzione, 2,* 63-70.

Molenaar, D., van der Sluis, S., Boomsma, D. I., & Dolan, C. V. (2011, December). *Genotype by environment interactionsi in cognitive ability tested in 14 different studies*. Paper presented at the 12th annual conference of ISIR, Limassol, Cyprus.

Mullis, I. V. S., Martin, M. O., Foy, P. Olson, J. F., Preuschoff, C., Erberber, E., et al. (2008). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in*

International Mathematics and Science Study at the Fourth and Eighth Grades. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M. & Foy P. (2007). *IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *The American psychologist, 67,* 130-59. doi:10.1037/a0026699

OECD (Organization for Economic Cooperation & Development). (2007). PISA 2006 competencies for tomorrow's world. Paris (France): OECD.

OECD (Organization for Economic Cooperation & Development) (2010a).*PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science (Volume I).* doi:10.1787/9789264091450-en

OECD (Organization for Economic Cooperation & Development) (2010b).*PISA 2009 Results: Learning Trends: Changes in Student Performance Since 2000 (Volume V).*doi:10.1787/9789264091580-en

Petrill, S. A., Lipton, P. A., Hewitt, J. K., Plomin, R., Cherny, S.S., Corley, R., & DeFries, J.C. (2004). Genetic and environmental contributions to general cognitive ability through the first 16 years of life. *Developmental Psychology, 40,* 805-812.

Rindermann, H. (2007). The g-factor of international cognitive ability comparisons: the homogeneity of results in PISA, TIMSS, PIRLS and IQ-tests across nations. *European Journal of Personality, 21,* 667-706. doi:10.1002/per.634

Rindermann, H. (2008). Relevance of education and intelligence at the national level for the economic welfare of people. *Intelligence, 36,* 127-142. doi:10.1016/j.intell.2007.02.002

Rindermann, H., & Ceci, S. J. (2009). Educational Policy and Country Outcomes in International Cognitive Competence Studies. *Perspectives on Psychological Science, 4,* 551-577. doi:10.1111/j.1745-6924.2009.01165.x

Robinson, D., Saggino, A., & Tommasi, M. (2011). The case against Lynn's doctrine that population IQ determines levels of socio-economic development and public health status. *Journal of Public Mental Health, 10,* 178-189. doi:10.1108/17465721111175056

Rodic, M., Tikhomirova, T., Zohu, X., Wei, W., Malykh, S., Ismatulina, V., Sabirova, E., Tosto, M., &Kovas, Y. (2011, December). *What really counts?: A cross-cultural perspective on cognitive underpinnings of individual differences in mathematics.* Paper presented at the 12th annual conference of ISIR, Limassol, Cyprus.

Rubinacci, A. (2011). *La Puglia brillante esempio di governo [Apulia an example of good governance]*.Tuttoscuola, n. 509, 15-24.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235–294. doi:10.1037/1076-8971.11.2.235

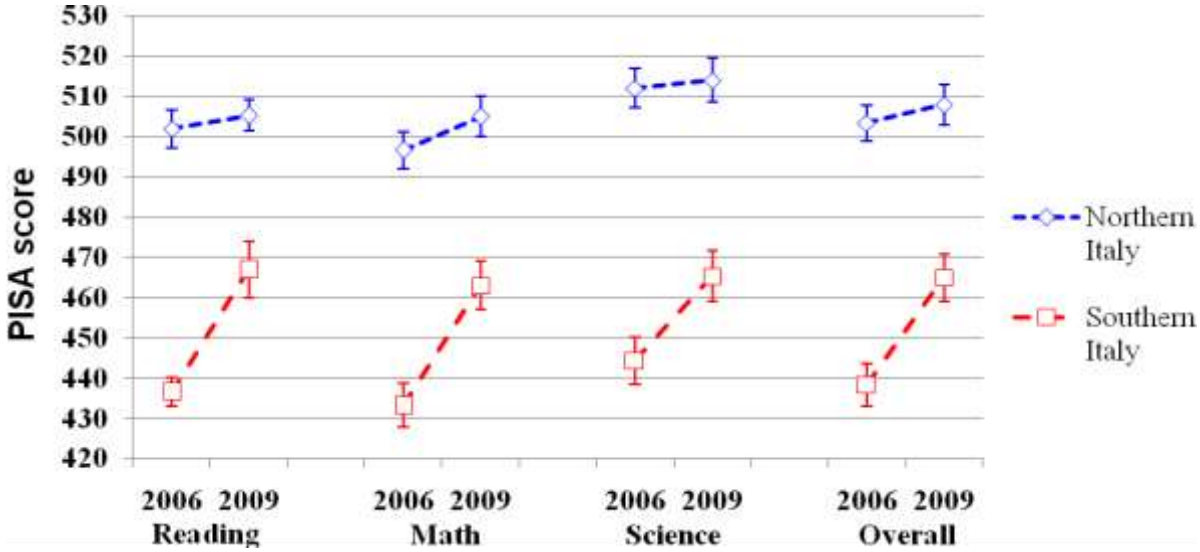
Rushton, J. P., & Jensen, A. R. (2010). Race and IQ: A theory-based review of the research in Richard Nisbett's intelligence and how to get it. *The Open Psychology Journal, 3*, 9–35, http://dx.doi.org/10.2174/1874350101003010009

Templer, D.I. (2012), Biological correlates of northern–southern Italy differences in IQ, *Intelligence* ,doi:10.1016/j.intell.2012.04.005

Wicherts, J. M., Dolan, C. V., & van DerMaas, H. L. J. (2010). A systematic literature review of the average IQ of sub-Saharan Africans. *Intelligence, 38,* 1-20. doi:10.1016/j.intell.2009.05.002

Wicherts, J.M., & Wilhelm, O. (2007). What is the national g-factor? *European Journal of Personality, 21,* 763-765.

Figure 1. Mean PISA scores obtained in 2006 and 2009 by northern and southern Italian Regions.



Note. Data are provided by INVALSI. The overall score is calculated by the arithmetic mean of reading, math, and science. North Italy: Trento, Lombardy, Veneto, Piedmont, Emilia-Romagna, Liguria; South Italy: Basilicata, Campania, Puglia, Sardinia, Sicily.

Table 1

Mean PISA 2006 and 2009scores in Italy for macro-area (INVALSI elaborations) and specific and overall mean differences (MD).

	Reading			Math			Science			
Macro Area	Pisa 2006	Pisa 2009	MD	Pisa 2006	Pisa 2009	MD	Pisa 2006	Pisa 2009	MD	Overall MD
North West	494 (4.7)	511 (3.9)	17 (7.3)	487 (4.3)	507 (4.0)	20 (6.0)	501 (4.1)	516 (4.0)	15 (6.2)	17.3
North East	506 (3.2)	504 (2.8)	-2 (5.9)	505 (3.1)	507 (2.9)	1 (4.5)	520 (2.8)	515 (2.8)	-5 (4.7)	-2
Center	482 (8.9)	488 (2.6)	5 (10.1)	467 (8.1)	483 (3.2)	16 (8.8)	486 (8.0)	491 (3.0)	5 (8.9)	8.7
South	443 (3.8)	468 (3.9)	26 (6.8)	440 (5.2)	465 (4.8)	25 (7.2)	448 (3.7)	466 (4.2)	19 (6.2)	22
South and islands	425 (6.9)	456 (4.8)	30 (9.3)	417 (5.2)	451 (5.1)	34 (7.4)	432 (4.6)	454 (4.8)	22 (7.1)	28.7
Italy	469 (2.4)	486 (1.6)	18 (5.0)	462 (2.3)	483 (1.9)	21 (3.2)	475 (2.0)	489 (1.8)	13 (3.7)	17.3

Note. Standard error in parentheses. MD = Mean difference PISA 2009–2006. North West: Val D'Aosta, Piedmont, Liguria, and Lombardy; North East: Bolzano, Trento, Veneto, Friuli, Emilia-Romagna; Central Italy: Tuscany, Umbria, Marche, Lazio; South: Abruzzi, Molise, Campania, and Apulia; South and Islands: Basilicata, Calabria, Sicilia, Sardinia.

Table 2

Mean PIRLS scores and score change between 2001 and 2006.

Macro Area	Mean 2001 (SE)	Mean 2006 (SE)	MD
North West	560.22 (4.13)	555.48 (4.20)	-4.74
North East	546.19 (6.24)	555.44 (6.96)	9.25
Center	548.31 (3.89)	557.49 (4.80)	9.18
South	527.83 (5.27)	545.97 (7.23)	18.14
South and Islands	525.38 (4.47)	546.13 (7.36)	20.65
Italy	541 (2.40)	551 (2.90)	11 (3.8)

Note. INVALSI elaborations. Standard error in parentheses. MD = Mean difference PIRLS 2006–2001. North West: Val D'Aosta, Piedmont, Liguria, and Lombardy; North East: Bolzano, Trento, Veneto, Friuli, Emilia-Romagna; Central Italy: Tuscany, Umbria, Marche, Lazio; South: Abruzzi, Molise, Campania, and Apulia; South and Islands: Basilicata, Calabria, Sicily, Sardinia.

Table 3

Overall view of the differences (in standard deviations) in achievement between northern and southern Italy in the international surveys.

	Reading							Math							Science						
Study	'00	'01	'03	'06	'07	'09	'11	'00	'01	'03	'06	'07	'09	'11	'00	'01	'03	'06	'07	'09	'11
PISA																					
15 years old	0.67		0.75	0.66		0.46		0.68		0.85	0.68		0.49		0.74		0.90	0.71		0.56	
PIRLS																					
4th grade		0.27		0.09																	
TIMSS																					
4th grade										0.22		0.21					0.38		0.29		
8th grade										0.42		0.42					0.57		0.49		
INVALSI																					
2nd grade							0.21							0.02							
5th grade							0.11							0.19							
6th grade							0.29							0.45							
10th grade							0.39							0.42							

Note. The mean difference was calculated by averaging the means of North- West/East together then subtracting the averaged means of the South and South and Islands macro-area. The mean difference was then divided by 100 in the case of PISA, PIRLS, and TIMSS, and by the National standard deviation in the case of INVALSI.

Table 4

Mean PISA scores change between 2006 and 2009 in Apulia (INVALSI elaborations) and specific and overall mean differences (MD).

	Reading			Math			Science			
Macro Area	Pisa 2006	Pisa 2009	MD	Pisa 2006	Pisa 2009	MD	Pisa 2006	Pisa 2009	MD	Overall MD
Apulia	440 (6.7)	489 (5.0)	49 (9.3)	435 (4.8)	488 (6.9)	53 (8.6)	447 (4.3)	490 (6.3)	43 (8.0)	48.3

Note. INVALSI elaborations. Standard error in parentheses. MD = Mean difference PISA 2009–2006.

Table 5

Mean scores obtained by children of different areas of Italy at the 2011 assessment of language (READ) and mathematical (math) achievement by Invalsi (Sds in parentheses) and mean differences in Sds between north and south regions.

	2nd grade		5th grade		6th grade		10th grade	
	READ	MATH	READ	MATH	READ	MATH	READ	MATH
North West	70.8 (18.41)	60.6 (17.04)	73.6 (13.03)	69.6 (15.83)	64.7 (16.33)	49.9 (18.48)	73.3 (14.29)	51.4 (17.37)
North East	70.3 (18.54)	60.0 (17.49)	73.3 (13.46)	69.9 (15.82)	63.9 (16.85)	50.8 (18.83)	73.0 (15.05)	52.3 (16.93)
Center	70.8 (18.15)	60.9 (17.74)	74.3 (12.81)	69.0 (15.79)	64.2 (16.56)	48.0 (18.35)	68.9 (16.71)	46.6 (17.39)
South	67.8 (19.56)	60.7 (19.22)	72.8 (14.35)	67.6 (18.07)	60.8 (18.01)	43.3 (18.22)	68.5 (16.27)	46.3 (18.50)
South and Islands	65.4 (20.84)	59.3 (19.91)	71.2 (14.77)	65.4 (18.06)	57.7 (18.27)	40.5 (17.76)	65.3 (16.98)	42.5 (16.66)
Italy	69.2 (19.17)	60.3 (18.24)	73.1 (13.71)	68.4 (16.80)	62.4 (17.38)	46.6 (18.74)	69.8 (16.14)	47.9 (17.80)
North-South Differences in SD	0.13	-0.03	0.05	0.15	0.21	0.32	0.33	0.32

Note. READ = Reading. North West: Val D'Aosta, Piedmont, Liguria, and Lombardy; North East: Bolzano, Trento, Veneto, Friuli, Emilia-Romagna; Central Italy: Tuscany, Umbria, Marche, Lazio; South: Abruzzi, Molise, Campania, and Apulia; South and Islands: Basilicata, Calabria, Sicily, Sardinia. North: Val D'Aosta, Piedmont, Liguria, Lombardy, Trento, Veneto, Friuli, Emilia-Romagna, Tuscany, Umbria, and Marche; South: Lazio, Abruzzi, Molise, Campania, Apulia, Basilicata, Calabria, Sicily, Sardinia.