In Italy, North-South Differences in Student Performance Are Mirrored by Differences in Polygenic Scores for Educational Attainment

2 authors, including:

Davide Piffer
Ulster Institute for Social Research
84 PUBLICATIONS **895** CITATIONS

In Italy, North-South Differences in Student Performance Are Mirrored by Differences in Polygenic Scores for Educational Attainment

Davide Piffer*, Richard Lynn
Ulster Institute for Social Research, London

* Corresponding author at pifferdavide@gmail.com

In Italy, there is a consistent north-south gradient in economic development, education, intelligence (IQ) and stature. The aim of this paper is to investigate the possible genetic cause of differences in educational attainment by using publicly available DNA data and recent genetic predictive models. Principal components analysis confirmed the presence of a high degree of population structure and clusters with little overlap between northern and southern Italy, corresponding to traditional macro-regions (Northern, Central, and Southern). This ancestry component correlates with the polygenic scores (PGS) for educational attainment (r = .2 to .3). The differences in the genetic scores for educational attainment and height follow the same north-south gradient and are of similar magnitude to the measured phenotypic differences (Cohen's d is 0.5 to 0.6).

Key Words: Italy, Inequality, Intelligence, Cognitive ability, Education, Genetics, Polygenic score

It has been shown that there are large north-south differences in intelligence in Italy (Lynn, 2010). In this study, the correlation for twelve Italian regions between IQ and latitude was .963. It was also shown that these regional IQs are strongly correlated with average incomes at r = .937, as well as with stature, infant mortality, literacy and educational attainment. It was proposed that the lower IQ in southern Italy may be attributable to genetic admixture with populations from the Near East and North Africa. Smirni (2020) reports an IQ of 92 for Sicily, inferred from the Raven Coloured Progressive Matrices (Raven et al., 2010) administered to children. Similar results were shown for fluid intelligence based on the PISA Creative Problem Solving test, which showed a difference of 9.2 IQ

points between north-western and southern Italy (Piffer & Lynn, 2014).

In a further study, it was reported that the percentage of the population with natural blond hair in the Italian regions was correlated at .81 with latitude and .74 with IQ. The percentage of blond hair was correlated .73 with the haplogroup xR1 allele marker of European ancestry and -.41 with the haplogroup E1b↑b allele marker of North African and Middle Eastern ancestry (Lynn, 2012). These results have been extended by Templer, who showed that regional IQs across the Italian regions were correlated positively with cephalic index (relative roundness of the skull) ($r = .88$), multiple sclerosis rate ($r = .85$) and schizophrenia rate ($r = .46$), and negatively with prevalence of black hair and brown eyes ($r = -.73$ and -0.61) (Templer, 2012). In addition, attempted murder rate was associated with lower IQ, lower cephalic index, dark eyes ($r = .94$), and black hair (Templer, 2013).

These findings were disputed by subsequent studies, which claimed that there was not enough evidence for IQ differences between northern and southern Italy and that these differences are not persistent in time and not attributable to genetic factors (Felice & Giugliano, 2011). Moreover, the use of school measures to derive IQ indices was criticized because they are dependent on regional differences in schooling quality (Cornoldi et al., 2010). The present study employs new genetic datasets and the results from the latest round of standardized testing of Italian students to address these criticisms. Moreover, it will restrict the analysis to education-related variables and their genetic correlates, without trying to measure differences in general intelligence or IQ.

Every year, Italy runs a standardized test of skills learned by pupils in school whose purpose is to measure the "ability to understand texts, logical faculties and the ability to solve new problems" (INVALSI, 2018). The test is divided in two sections: Italian, and Mathematics. The Italian test measures the knowledge and correct use of grammar and the ability to understand written texts taken from literature, non-fiction or everyday life, whereas "the Mathematics tests measure the ability to use mathematical knowledge to solve problems, real or otherwise, logical skills, interpretation of graphs, interpretation of phenomena with a quantitative dimension, modelling, or use in various scientific disciplines" (INVALSI, 2018). While we are not familiar with any studies linking INVALSI to traditional intelligence tests, it is well established that the association between scholastic/achievement tests and intelligence is about .80 when measurement error is removed (Zaboski et al., 2018). Thus, these tests are a very good proxy for intelligence differences between the regions, though they may also index school quality and non-cognitive personality traits to some degree.

The aim of this paper is to test the hypothesis that the differences in performance between northern and southern Italian students are genetic in origin and not purely due to non-genetically caused cultural or economic conditions.

This has become possible because molecular genetics has made huge progress over the last decade, spurred also by a reduction in genomic sequencing costs, with the collection of large numbers of genomes along with medical, psychological and biometric data for nationally representative samples. Thousands of researchers across the world have capitalized on this opportunity by developing sophisticated statistical methods to identify genetic variants associated with a wide range of human polygenic traits. For example, polygenic scores derived from genome-wide association studies of height employing large samples (~700K) can now explain around 25% of the variation between individuals (Yengo et al., 2018). A new GWAS currently in print, using data from over 5.4 million individuals, explains ~40% of individual differences in Europeans, and 10-20% in other ancestries (Yengo et al., 2022).

The first successful GWAS of educational attainment, published in 2013, identified 3 SNPs associated with educational attainment in individuals of European ancestry at a genome-wide significance level (Rietveld et al., 2013). In the following years, larger and larger samples enabled the identification of thousands of genetic associations (Okbay et al., 2022; Lee et al., 2018). In the meantime, attempts were made to use the genetic variants identified in these studies to identify patterns of divergent selection between ethnic groups (Berg & Coop, 2014; Dunkel et al., 2019; Guo et al., 2018; Piffer, 2013, 2015, 2019). Strong correlations (about .9) were found between average population IQ and polygenic scores (PGS). Specifically, genetic scores for educational attainment tended to be higher among Ashkenazi Jews and East Asians, followed by Europeans, then South Asians and Africans (Dunkel et al., 2019; Piffer, 2015, 2021).

Regarding Italy, evidence for selection at several genetic loci related to immunity and pigmentation has been found (Parolo et al., 2015) and a strong north-south cline in genetic variation has been shown (Cavalli-Sforza et al., 1994; Parolo et al., 2015; Raveane et al., 2019). Recently, pervasive signatures of climate-mediated positive selection at genes that control thermogenesis and adiposity were found in northern but not southern Italians, related to genes affecting insulin secretion and glucose homeostasis (Sazzini et al., 2020).

The present study, using the genetic variants identified by the latest and largest GWAS of educational attainment (EA4, Okbay et al., 2022) comprising over 3 million individuals, aims to test the prediction that northern Italians have higher polygenic scores for educational attainment than southern Italians. Central Italians are expected to have an intermediate score. Moreover, a secondary prediction is that the magnitude of the genetic differences is similar to that of the test score differences.

EA4 explains 12-16% of the variance in educational attainment, an

improvement of about 20% over the previous GWAS (Lee et al., 2018). However, the accuracy is reduced by half after controlling for parental PGS. That is, the direct effects of EA4 account for 6-8% of the variance in EA, pointing to the importance of genetic nurture (Okbay et al., 2022). Moreover, EA4 was found to be predictive of cognitive and academic performance (8.7% in verbal test scores and 12.3% in grade point average) (Okbay et al., 2022).

Our choice of test scores of academic performance (INVALSI) for population comparison, instead of IQ, is justified by the need to have a measure that is as close as possible to the PGS identified by the GWAS of educational attainment. Because academic performance is not a pure measure of intelligence, it is possible that other, non-cognitive factors contribute to differences in this outcome between individuals and between northern and southern Italians.

Methods

The polygenic score for educational attainment (henceforth EA4, to distinguish it from previous GWAS, see Piffer, 2021) was computed from the clumped set of 3,951 SNPs that met the GWAS significance threshold, and weighted by effect size using R (R Core Team, 2022). Allele frequencies for individuals were computed using PLINK 2.0 (Chang et al., 2015).

For height, the GWAS summary statistics from Yengo et al. (2018) were used, employing the betas from the analysis with the highest predictive accuracy (lasso+ctpr). These were pruned using a 250 KB window in steps of 5, with R^2 threshold = 0.1, leaving 33,168 SNPs. This was done using PLINK 2.0.

Two datasets of Italian genomes were used for this study. One comprised 591,217 SNPs genotyped with the Illumina Omni array for 129 Italian individuals (Raveane et al., 2019). This dataset was imputed using the Michigan Imputation Server. The imputed files were filtered for variants with $R^2 > .4$ using BCFtools (Danecek et al., 2021). The 8,618 SNPs in the GWAS summary file had 8,426 matches in the imputed files. After applying the p-value filter ($p < 5 * 10^{-8}$), 3,902 out of 3,951 SNPs (98.8%) were found.

The second dataset comprised 947 individuals sequenced at the whole-genome level. However, the publicly accessible files contain only aggregate frequencies for three Italian populations (Cocca et al., 2020): North-West (Val Borbera - VBI), North-East (Friuli Venezia Giulia - FVG) and South-East (Carlantino - CAR). These populations belong to the INGI (Italian Network of Genetic Isolates). A potential methodological problem is that these are isolated populations from rural areas, hence they are not representative of the respective Italian regions, particularly for education and intelligence, which are prone to selective migration, usually from rural to metropolitan areas (Abdellaoui et al., 2019).

As an additional validation, the SNPs (N = 3,269) from a previous GWAS of educational attainment (EA3, Lee et al., 2018) were used. The MTAG summary statistics were used because they include different cognitive abilities besides education (cognitive performance, highest math class completed, and math ability). Results are reported in the supplementary files.

Statistical analyses were carried out using R version 4.1.3. The R code and supplementary figures can be found at:
https://figshare.com/articles/dataset/In_Italy_North-South_differences_in
_student_performance_are_mirrored_by_differences_in_polygenic_scores_for_
educational_attainment_/19755274/1.

Results

Population structure in the Italian peninsula

A principal component analysis of 129 Italian genomes was performed with PLINK 2.0. The first principal component shows clusters, with little overlap between northern and southern Italians on the first component (Figure 1). Considerable overlap is present between central and, to a larger extent, southern Italians.
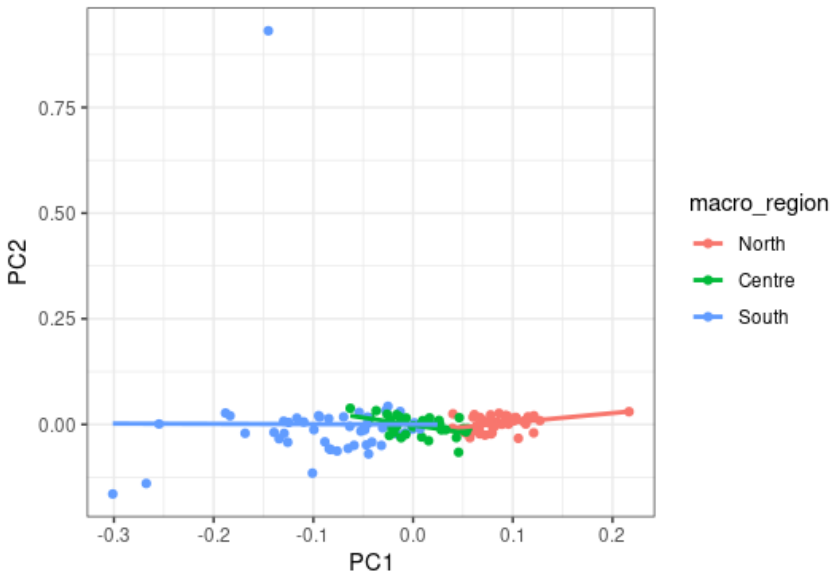

**Figure 1.** *Principal components analysis of Italian genomes.*

Polygenic scores follow a north-south gradient

The genomes originating from different Italian regions were divided into three

groups: North, Central and South and the polygenic scores of educational attainment for each individual were computed. The results show a clear north-south gradient, with central Italians occupying an intermediate position (Figure 2). A north-south gradient was apparent also using the INGI samples (Figure 3), although the Friuli V. G. Sample (*NE_Italy*) had an intermediate score between the NW Italy and the Southern Italy sample.
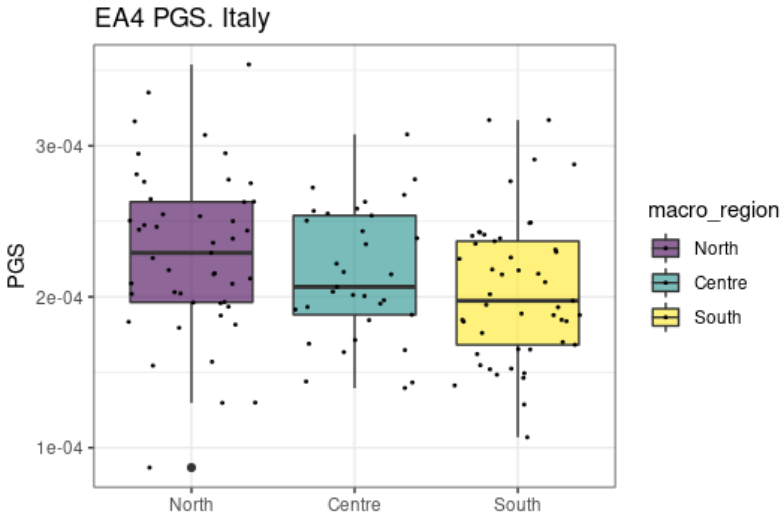
Figure 2. *Educational attainment (EA4) polygenic score (PGS) for Northern, Central and Southern Italy (imputed array data).*
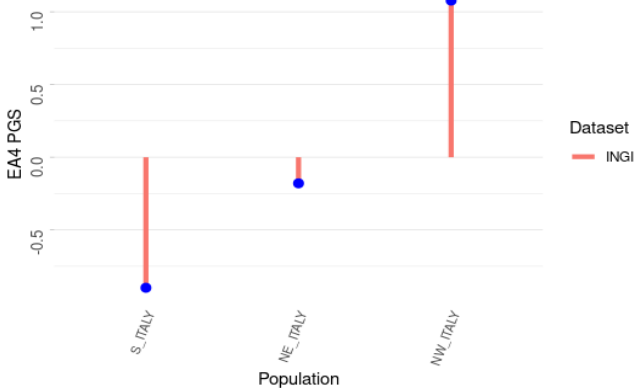
Figure 3. *Educational attainment (EA4) polygenic score (PGS), INGI genomes (whole genome sequencing data).*

Magnitude of the differences

The magnitude of the differences in polygenic scores between North and South, as represented by Cohen's *d*, was 0.49 (95% C.I. = 0.08 to 0.90). An outlier was identified using the R function (rstatix package) "identify_outliers" (Kassambara, 2021). The effect size increased after removing the outlier (Cohen's *d* = 0.57, 95% C.I. = 0.15 to 0.99). One-way ANOVA was performed to test if the three group means were significantly different ($F_{(2,126)} = 3.1$, $p = .049$). After removal of the outlier, statistical significance increased ($F_{(2,125)} = 4.15$, $p = .018$).

The Abruzzo region is genetically and geographically a grey area between southern and central Italy (Sazzini et al., 2016; Raveane et al., 2019), hence the analysis was re-run after grouping Abruzzo with other central Italian regions. With this grouping, one-way ANOVA found similarly significant differences ($F_{(2,126)} = 3.37$, $p = 0.037$ and $F_{(2,125)} = 4.46$, $p = 0.013$), before and after removing the outlier, respectively. The effect size for the North-South difference increased to Cohen's *d* = 0.55 (95% C.I. = 0.11 to 0.98) (suppl. Figure 1). Without the outlier, Cohen's *d* increased to 0.64 (95% C.I. = 0.20 to 1.07).

Student performance

Since the polygenic score for educational attainment is obtained from adult GWAS samples that have completed their studies, the highest grade (10th) for INVALSI 2018 was used as the closest proxy of adult educational attainment. Cohen's *d* was 0.60 for math and 0.49 for language skills (Italian).

Table 1a. *INVALSI summary statistics for Italian regions (natives only), 2018; mean scores ± standard deviation are shown, and Cohen's d comparing northern and southern regions.*

Math 10th grade South		Math 10th grade North			
Region	Mean ± SD	Region	Mean ± SD	Difference	*d*
Abruzzo	201.7 ± 40	Valle d'Aosta	205.1 ± 35		
Molise	196.7 ± 40	Liguria	209.4 ± 39		
Campania	186.6 ± 38	Lombardia	217.8 ± 39		
Puglia	191.7 ± 37	Trentino	221.3 ± 37		
Basilicata	197.9 ± 37	Veneto	217.1 ± 37		
Calabria	176.6 ± 35	Friuli V.G.	216.2 ± 38		
Mean	191.9 ± 37.8		214.5 ± 37.5	22.6	0.60

Table 1b. *INVALSI summary statistics for Italian regions (natives only), 2018; mean scores ± standard deviation are shown, and Cohen's d comparing northern and southern regions.*

Italian 10th grade South		Italian 10th grade North			
Region	**Mean ± SD**	**Region**	**Mean ± SD**	**Difference**	***d***
Abruzzo	202.2 ± 39	Valle d'Aosta	210.2 ± 33		
Molise	196.3 ± 42	Liguria	208.9 ± 37		
Campania	191.0 ± 43	Lombardia	216.2 ± 35		
Puglia	194.8 ± 38	Trentino	216.6 ± 33		
Basilicata	198.7 ± 38	Veneto	215.5 ± 36		
Calabria	183.9 ± 42	Friuli V.G.	210.9 ± 35		
Mean	194.5 ± 40.3		213.1 ± 34.8	18.5	0.49

Correlation between individual ancestry and educational attainment PGS
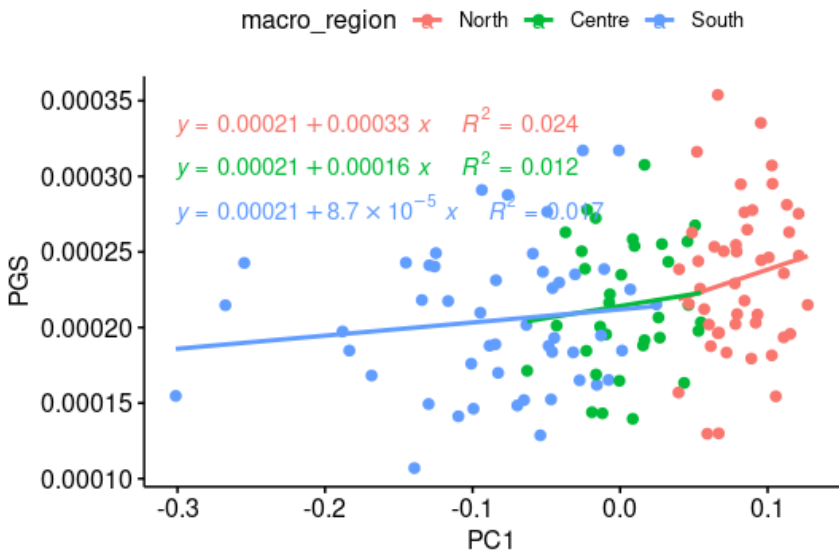
Figure 4. *Correlation between EA4 and PC1 by macro-region.*

Italy has a very strong population structure, as demonstrated by previous studies and shown in the principal component plot in Figure 1. However, even within the same regions there is individual variation in the amount of genetic admixture with different ancestral groups. This reflects internal migration, mixing, and residual historical variation. The score on the first principal component,

representing the north-south gradient, for each individual, was correlated to the educational attainment polygenic score (PGS). There was a positive correlation ($r = .20$, $p = .02$) between the first principal component (PC1) and the EA4 PGS, implying that samples with greater northern Italian (or lower southern Italian) admixture had higher EA4 PGS. Removal of the outlier increased the correlation from $r = .20$ to $r = .27$. Sub-group analysis showed that there was a linear relationship between PC1 and EA4 for each macro-region (Figure 4). There was homogeneity of regression slopes as the interaction term was not statistically significant, $F(2, 122) = 0.3$, $p = 0.738$.

Comparing regional PGS after accounting for population structure

ANCOVA was run to test the model that polygenic scores are significantly different between groups after accounting for population stratification (first principal component, PC1). After adjusting for PC1, there was no statistically significant difference in EA4 between the three macro-regions ($F(2,124) = 0.234$, $p = 0.791$).

Ranking within the European genetic landscape

EA4 PGS was computed for the European populations in the gnomAD dataset (non-Finnish European or "NFE") and for the Bergamo (Northern Italy, Lombardy) HGDP sample (Bergstrom et al., 2020).
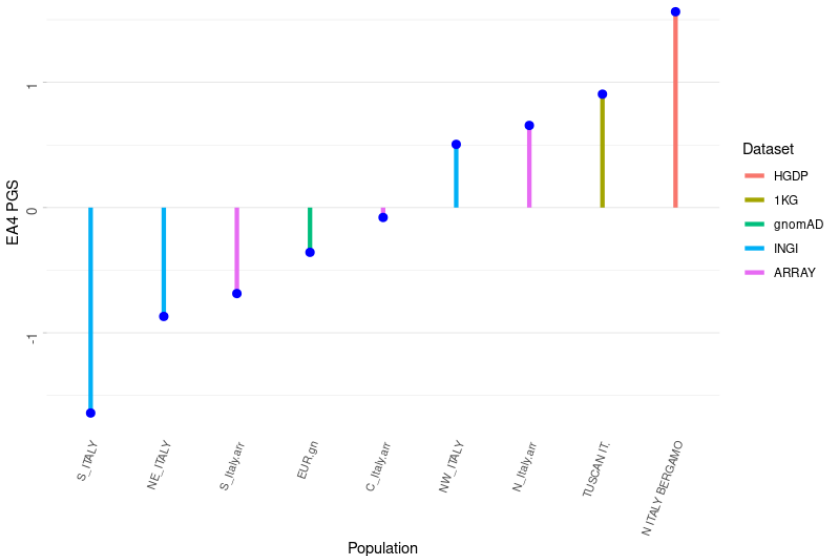

Figure 5. *Educational attainment (EA4) polygenic score (PGS) of European populations.*

The sample was not included in the main analysis because of its small size (N = 12). However, its EA4 PGS was the highest of all samples. It can be seen from Figure 5 that Northern Italy's EA4 PGS is higher than the average European PGS.

Height

Statistical data for human stature for Italian regions is publicly available and consists of representative samples (conscripts, males aged about 18-20). Morever, the GWAS of height have reached a good predictive accuracy, which enables us to predict individual and group-level values (Piffer, 2021). Because height in Italy follows a north-south gradient similar to that of education and IQ, we used height as a control phenotype to test the hypothesis that the polygenic score mirrors the observed physical differences between northern and southern Italians.

A north/south gradient was found for the height PGS (Figure 6). Northeastern Italians from Friuli had scores comparable to northern Europeans (Finnish, Americans of Central and Northwestern European descent, British), other northern Italians (NW Italy, N Italy). Tuscans had slightly lower scores, and Southern Italians had the lowest PGS.
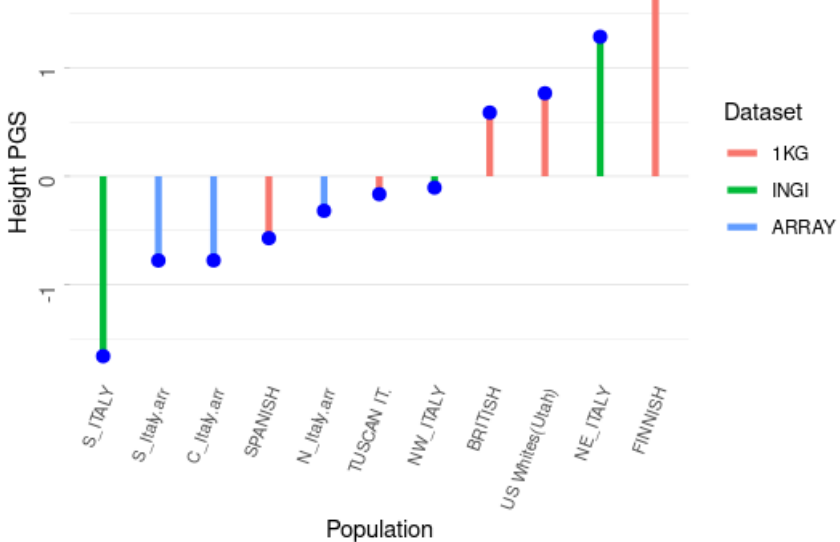
Figure 6. *Height polygenic scores (PGS) of European populations.*

Discussion

The differences in students' performance between northern and southern Italians are mirrored by differences in the frequencies of alleles that are associated with educational attainment, and similar in magnitude (Cohen's *d* was 0.49 to 0.64 depending on grouping criteria and outlier removal). The limited nature of the genomic dataset with individual data (N = 129) did not allow higher spatial resolution, but it is likely that the polygenic score difference between the highest and lowest scoring regions is of a higher magnitude.

The results were replicated using a larger WGS dataset (N = 947), although there was a large difference between the North-Eastern and the North-Western sample, similar to the difference between their average and the southern sample. This irregularity is probably caused by the nature of the samples, which come from genetically isolated, remote villages. Nonetheless, the overall pattern of higher Northern Italian polygenic scores was preserved. Central Italy's polygenic score occupied an intermediate position between northern and southern samples, in line with results from student performance (Piffer & Lynn, 2014) and their ancestry profile (Figure 1).

The height PGS showed a similar distribution. Remarkably, the Friuli-V.G. (NE Italy) sample had very high PGS similar to northern Europeans. This was expected based on this region consistently obtaining the highest average stature among Italian regions over a century of military conscripts data, with the latest average height (1980 cohort) of 178 cm (Corsini, 2008).

The data analyzed in this paper do not enable us to fit different models with or without non-cognitive factors to account for the observed differences. Educational attainment is a complex phenotype that is influenced by intelligence but also a broad array of non-cognitive factors. Thus, in the absence of large GWAS of IQ or general intelligence, it is not possible to estimate the relative contribution of cognitive and non-cognitive factors to the observed population differences in educational performance.

The results obtained with the SNPs from the previous EA GWAS were less clearcut. The north-south gap was evident in the INGI dataset (suppl. Figure 2), which showed the same gradient as the EA4 PGS (NW > NE > S). However, in the array data set (suppl. Fig. 3), northern Italians occupied an intermediate position between southern and central Italian (Centre > North > South). This is likely a result of the lower accuracy of the EA3 PGS, because central Italians are genetically equally distant to southern and northern Italians (Sazzini et al., 2020), reflecting the smooth genetic gradient observed across Italy. This interpretation is bolstered by the INVALSI and PISA scores, which show central Italian regions occupying an intermediate rank between the north and the south.

The correlation between population structure (first principal component) and

EA4 was homogeneous and in the same direction within the three macro-regions (Figure 4). After adjusting for the first principal component, representing ancestry-related population structure, EA4 did not significantly differ between the groups.

The lack of an effect of macro-region independent of ancestry on EA4 could suggest that the polygenic score differences are due to older geographic patterns of ancestry rather than recent migrations driven by education, contrary to what has been shown to be the case in the UK, where regional differences in the EA PGS were not fully accounted for by ancestry. In fact, recent migrations explained the differences in education in the UK (Abdellaoui, 2019).

An interpretation of this finding is that in Italy, natural selection acted over many generations to produce different population means for education-related traits (intelligence, impulsivity, openness to experience, etc.), thus becoming associated with old ancestry components. Conversely, in the presence of recent massive selective migration, with large numbers of migrants whose average levels of cognitive traits are different from those who stay behind, the association between ancestry and EA PGS would break down (Abdellaoui, 2019)

In fact, the individuals in our study had all their grandparents born in the same region, hence the data set does not represent the effects of the massive migrations that occurred over the last century, mainly from the countryside to the city and from the poorer South to the industrialized North, on the polygenic scores of people living in different regions. The present study lacked the power to detect subtle differences in polygenic scores independent of ancestry, due to the small sample size and the low spatial resolution.

Another possible interpretation of the findings is that the GWAS failed to remove the effects of population stratification on the phenotype (Sohail et al., 2019). However, the EA4 GWAS was carried out on individuals of White British and European American ancestry living in the UK and in the US. Therefore, population structures in the discovery populations are different from those present in Italy, and therefore less likely to bias the Italian results.

References

Abdellaoui, A., Hugh-Jones, D., Yengo, L., Kemper, K.E., Nivard, M.G., Veul, L., ... & Visscher, P.M. (2019). Genetic correlates of social stratification in Great Britain. *Nature Human Behaviour* 3: 1332-1342. doi: 10.1038/s41562-019-0757-5. Epub 2019 Oct 21. PMID: 31636407.

Berg J.J. & Coop, G. (2014). A population genetic signal of polygenic adaptation. *PLoS Genetics* 10(8): e1004412. https://doi.org/10.1371/journal.pgen.1004412

Bergström, A., McCarthy, S.A., Hui, R., Almarri, M.A., Ayub, Q., Danecek, P., ... & Tyler-Smith, C. (2020). Insights into human genetic variation and population history from 929 diverse genomes. *Science* 367(6484): eaay5012. doi:10.1126/science.aay5012.

Cavalli-Sforza, L.L., Menozzi, P. & Piazza, A. (1994). *The History and Geography of Human Genes*. Princeton NJ: Princeton Univ. Press.

Chang, C.C., Chow, C.C., Tellier, L.C.A.M., Vattikuti, S., Purcell, S.M. & Lee, J.J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience* 4.

Chung, W., Chen, J., Turman, C., Lindstrom, S., Zhu, Z., Loh, P.-R., . . . & Liang, L. (2019). Efficient cross-trait penalized regression increases prediction accuracy in large cohorts using secondary phenotypes. *Nature Communications* 10: 569. doi: 10.1038/s41467-019-08535-0

Cocca, M., Barbieri, C., Concas, M.P., Robino, A., Brumat, M., Gandin, I., ... & Mezzavilla, M. (2020). A bird's-eye view of Italian genomic variation through whole-genome sequencing. *European Journal of Human Genetics* 28: 435-444. https://doi.org/10.1038/s41431-019-0551-x

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. https://doi.org/10.1016/j.intell.2010.06.003.

Corsini, C.A. (2008). *Statura, salute e migrazioni: le leve militari italiane*. Forum Edizioni. ISBN: 978-8884204448

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., ... & Li, H. (2021). Twelve years of SAMtools and BCFtools. *Gigascience* 10(2): giab008

Dunkel, C.S., Woodley of Menie, M.A., Pallesen, J. & Kirkegaard, E.O.W. (2019). Polygenic scores mediate the Jewish phenotypic advantage in educational attainment and cognitive ability compared with Catholics and Lutherans. *Evolutionary Behavioral Sciences* 13: 366-375. https://doi.org/10.1037/ebs0000158

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. https://doi.org/10.1016/j.intell.2010.09.004.

Guo, J., Wu, Y., Zhu, Z., Zheng, Z., Trzaskowski, M., Zeng, J., ... & Yang, J. (2018). Global genetic differentiation of complex traits shaped by natural selection in humans. *Nature Communications* 9: 1865. https://doi.org/10.1038/s41467-018-04191-y

INVALSI (2018). *The INVALSI tests according to INVALSI*. https://invalsi-areaprove.cineca.it/docs/2018/INVALSI_tests_according_to_INVALSI.pdf

Kassambara, A. (2021). *rstatix: Pipe-friendly framework for basic statistical tests*. R package version 0.7.0. https://CRAN.R-project.org/package=rstatix

Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., ... & Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics* 50: 1112–1121. doi: 10.1038/s41588-018-0147-3

Lynn, R. (2010). North-south differences in IQ in Italy. *Intelligence* 38: 90-100.

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

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S.M., … & Young, A.I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics* 54: 437-449.

Parolo, S., Lisa, A., Gentilini, D., Di Blasio, A.M., Barlera, S., Nicolis, E.B., ... & Bione, S. (2015). Characterization of the biological processes shaping the genetic structure of the Italian population. *BMC Genetics* 16: 132 (2015). https://doi.org/10.1186/s12863-015-0293-x

Piffer, D. (2013). Factor analysis of population allele frequencies as a simple, novel method of detecting signals of recent polygenic selection: The example of educational attainment and IQ. *Mankind Quarterly* 54: 168-200. doi: 10.46469/mq.2013.54.2.3

Piffer, D. (2015). A review of intelligence GWAS hits: Their relationship to country IQ and the issue of spatial autocorrelation. *Intelligence* 53: 43-50. doi: 10.1016/j.intell.2015.08.008

Piffer, D. (2019). Evidence for recent polygenic selection on educational attainment and intelligence inferred from GWAS hits: A replication of previous findings using recent data. *Psych* 1(1): 55-75. doi: 10.3390/psych1010005

Piffer, D. (2021). Divergent selection on height and cognitive ability: Evidence from Fst and polygenic scores. *OpenPsych*. Doi: 10.26775/OP.2021.04.03

Piffer, D. & Lynn, R. (2014). New evidence for differences in fluid intelligence between north and south Italy and against school resources as an explanation for the north–south IQ differential. *Intelligence* 46: 246-249.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raveane, A., Aneli, S., Montinaro, F., Athanasiadis, G., Barlera, S., Birolo, G., ... & Capelli, C. (2019). Population structure of modern-day Italians reveals patterns of ancient and archaic ancestries in Southern Europe. *Science Advances* 5(9): eaaw3492. Doi: https://doi.org/10.1126/sciadv.aaw3492

Raven, J., Raven, J.C. & Court, J.H. (2010). *Standard Progressive Matrices*. Oxford: Oxford Psychologists Press.

Rietveld, C.A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N.W., ... & Preisig, M. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340: 1467-1471. doi: 10.1126/science.1235488.

Sazzini, M., Gnecchi Ruscone, G., Giuliani, C., … & Luiselli, D. (2016). Complex interplay between neutral and adaptive evolution shaped differential genomic background and disease susceptibility along the Italian peninsula. Sci Rep 6, 32513. https://doi.org/10.1038/srep32513

Sazzini, M., Abondio, P., Sarno, S., Gnecchi-Ruscone, G.A., Ragno, M., Giuliani, C., ... & Garagnani, P. (2020). Genomic history of the Italian population recapitulates key evolutionary dynamics of both Continental and Southern Europeans. *BMC Biology* 18: 51. https://doi.org/10.1186/s12915-020-00778-4

Smirni, D. (2020). The Raven's Coloured Progressive Matrices in healthy children: A qualitative approach. *Brain Sciences* 10: 877-889.

Sohail, M., Maier, R.M., Ganna, A., Bloemendal, A., Martin, A.R., Turchin, M.C., ... & Sunyaev, S.R. (2019). Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* 8: e39702. DOI: 10.7554/eLife.39702

Templer, D.I. (2012). Biological correlates of northern-southern Italy differences in IQ. *Intelligence* 40: 511-517. http://dx.doi.org/10.1016/j.intell.2012.04.005

Templer, D.I. (2013). Biological and cognitive determinants of murder and attempted murder in the Italian regions. *Mankind Quarterly* 54: 26-48.

Yengo, L., Vedantam, S., Marouli, E., Sidorenko, J., Bartell, E., Sakaue, S., ... & Understanding Society Scientific Group (2022). A saturated map of common genetic variants associated with human height from 5.4 million individuals of diverse ancestries. *bioRxiv*. Doi: https://doi.org/10.1101/2022.01.07.475305

Zaboski II, B.A., Kranzler, J.H. & Gage, N.A. (2018). Meta-analysis of the relationship between academic achievement and broad abilities of the Cattell-Horn-Carroll theory. *Journal of School Psychology* 71: 42-56.