North-South Differences in Spain in IQ, Educational Attainment, per capita Income, Literacy, Life Expectancy and Employment

Richard Lynn*
University of Ulster, Coleraine, Northern Ireland

IQs are presented for fifteen regions of Spain showing a north-south gradient with IQs highest in the north and lowest in the south. The regional differences in IQ are significantly correlated with educational attainment, per capita income, literacy, employment and life expectancy, and are associated with the percentages of Near Eastern and North African genes in the population.

Key Words: Spain; IQ; Income; Educational attainment; Literacy; Employment; Literacy; Life expectancy.

Introduction

It has been reported in a number of studies that average IQs are lower in the Balkans in the south east of Europe and in southern Italy than in northern and central Europe and northern Italy (Lynn, 2006, 2010a, b; Lynn & Vanhanen, 2006; Meisenberg & Lynn, 2012). In relation to a British mean IQ of 100 (SD = 15), IQs throughout northern and central Europe and northern Italy are also approximately 100, except in Ireland where the IQ has been estimated at 95.7 (Meisenberg & Lynn, 2012). In Southeast Europe (the Balkans), IQs were found to be lower in Bulgaria (93), Greece (92), Romania (94), and Serbia (89). The lower IQ in the Balkans has been confirmed in an analysis of cognitive ability assessed by reading, math and science abilities by Rindermann (2007), who has added an IQ of 87 for Albania and 89 for Macedonia. In a further study, an IQ of 88 for Serbia has been reported by Rushton & Čvorović (2009). Further data confirming these results and giving IQs of 97 for Croatia, 94 for Greece, 88 for Serbia, and 87 for Bulgaria

* E-mail: Lynnr540@aol.com

have been reported by Lynn & Mikk (2009).

The explanation proposed for the lower IQ in the Balkans and in southern Italy is that the populations are a cline of partly European and partly Near Eastern origin (Lynn, 2006, p. 18; 2010a). Average IQs in the Near East (Lebanon, Syria, Jordan, Iraq and Iran) are given as 84 and in Turkey as 90 in Lynn & Vanhanen (2006). IQs in the Balkans are therefore intermediate between those of approximately 100 in northern and central Europe, and 84 in the Near East, as would be expected in a cline of the two ancestral populations.

The reason that the populations of the Balkans are a mixed European-Near Eastern cline is that the Balkans are separated from south west Asia by only about 1 mile of water (the Dardanelles) that has been easily and frequently crossed for millennia. Archaeological evidence indicates that farming peoples from present day Turkey migrated into the Balkans in Neolithic times (Semino et al, 1996, 2000). In more recent times the exchange of populations across the Dardanelles has been facilitated during the many centuries in which the Balkans and the Near East have been unified in single states. Between the 9th and 4th centuries BC the Greeks established colonies in Anatolia (present day Turkey). The Romans and later the Byzantines ruled the Balkans and south west Asia from around 100 BC to around 1300 AD. From 1354 the Ottoman Empire based in present day Turkey began to colonize the south east of the Balkans. By 1430 the Ottomans had conquered most of the Balkans and continued to expand their control until 1683, when they reached the outskirts of Vienna but were defeated and pushed back from the north, but they retained most of the Balkans until the late 19th century, when they lost this territory except for the region around Istanbul.

During these millennia people moved freely between the Balkans and the Near East. These people interbred, smoothening the European-Near Eastern cline. This has been shown by Cavalli-Sforza et al (1994) in their genetic descent tree, in which Greeks are shown to be more closely related to Iranians and other southwest Asian peoples than to Italians, Danes, and English. This genetic similarity is also apparent for intelligence, for which IQs in the Balkans are intermediate between those in northern and central Europe and the Near East.

This theory has been tested by examining IQs in the north and south of Italy (Lynn, 2010a). The south, including Sicily and Sardinia, has been colonized intermittently for millennia by peoples from the Near East and North Africa. Around 750 BC the Phoenicians (from present day Lebanon) and later the Carthaginians (from present day Tunisia in North Africa) colonized parts of Sicily. Later, "Arabs occupied Sicily, Sardinia and Southern Italy in the seventh to the ninth centuries" (Cavalli-Sforza et al 1994, p. 261). Sicily remained an Arab state for about 300 years until 1060 and "during and after the Arab conquest, large Arab immigration took place" (M.M.C., 1960, p. 608). These colonizations have produced a cline of partly European and partly Near Eastern/North African origin in southern Italy. This has been shown genetically by Cavalli-Sforza et al (1994, p. 277), who write of the population genetics of Italy that "northern Italy shows similarities with countries of central Europe, whereas central and southern Italy are more similar to Greece and other Mediterranean countries. This corresponds to the well-known differences in physical type (especially pigmentation and stature) between the northern and north-central Italians on the one side and southern Italians on the other". By "Mediterranean countries" Cavalli-Sforza et al mean the countries that border the Mediterranean including those of North Africa and the Near East. They note also that the Sardinians are genetically more closely related to the Greeks, Lebanese and North African Berbers than to central and northern Europeans (Cavalli-Sforza et al, 1994, pp. 78, 274).

It can therefore be predicted that the IQ in northern Italy should be about the same as that in central and northern Europe, while the IQ in southern Italy should be about the same as that in the Balkans. It has been shown that this is so: the IQ in northern Italy is approximately 100, while in southern Italy the IQ is approximately 90, about the same as that in the Balkans (Lynn, 2010a). This north-south IQ gradient in Italy has been disputed by Beraldo (2010) and by Cornoldi et al (2010). These criticisms have been answered in Lynn (2010b), who shows that the north-south Italian IQ gradient has been found in Piagetian tests, Progressive Matrices tests, in PISA tests of reading comprehension, mathematic ability, and science understanding, and in two further data sets (MIT-Advanced and Invalsi) reported by Cornoldi et al (2010).

In the present paper a further test of this theory is made by examining whether the same north-south difference in IQ is present in Spain. There has been extensive immigration of North Africans into southern Spain over the course of millennia, as into the Balkans and southern Italy, and this has likewise resulted in a mixed European-North African population in the south. It can therefore be predicted that there should be a north-south gradient of intelligence in Spain similar to that in the Balkans and Italy, with higher IQ is higher in the north than in the south.

Before presenting data on this prediction it will be useful to sketch the history of the immigration of North Africans into Spain. The Iberian peninsula (present day Spain and Portugal) was first populated from the north by European peoples (Cavalli-Sforza et al, 1994, p. 259). North Africans from present day Tunisia (Carthaginians) and Near Eastern Phoenicians from present day Lebanon began to colonize the south around 550 BC when they established settlements on the coast of southern Spain (Aubet, 2001). They held these for approximately 350 years until the Romans defeated the Carthaginians (led by Hannibal) in the second Punic

War in 202 BC and took control of the Carthaginian settlements in southern Spain. The North African and Near Eastern peoples were not expelled but were assimilated with the indigenous population.

The next invasion of Spain by North Africans took place in 711 AD when the Umayad Caliph's army invaded and took possession of southern Spain. By 716 the North Africans had conquered the whole of the Iberian peninsula except for the northern regions of Asturias, the Basque country, Cantabria, Catalonia, Navarre and Rioja (Times, 2001).

The North Africans established their capital at Cordoba in southern Spain. They held most of Spain for approximately a century, but were expelled by the French under Charlemagne from the most northerly third of the country between 800 and 813. During the next six and a half centuries the European peoples of the north gradually regained control of the peninsula. By 1100 they had pushed the North Africans back into the southern half of the country, and by 1300 they had pushed them back further into the southern kingdom of Grenada, where they continued to speak Arabic and adhere to the Muslim religion (R.T.D., 1960).

The Spanish reconquered this last North African state in 1492 and gave the inhabitants the choice between expulsion or conversion to Christianity. Virtually all opted for conversion (R.T.D., 1960, p.122), so at the end of this period many ethnically North African and mixed-race people remained in southern Spain, adopted the Spanish language and religion, and became wholly assimilated with the European Spanish population.

During these occupations there was interbreeding between the North African colonists and the indigenous population, mainly between North African men and indigenous women. This introduced North African genes into the gene pool of the indigenous Spanish population.

The genetic impact of the occupation of southern Spain by North Africans for a period of almost eight centuries has been noted by Cavalli-Sforza et al (1994, pp. 292, 295), who show that the southern Spanish are more related genetically to peoples of the Near East and North Africa than to the Spanish of the central and northern regions. The Portuguese geneticists Pereira, Prata & Amorim (2000) have confirmed this observation: "historical sources document a deeper influence of North African Berber (as well as Arab) people in Central and particularly South Iberia, compared to North Iberia where the Muslim presence is recorded to have been more ephemeral and consequently to have made less cultural and demographic impact".

In further studies, the Spanish geneticists Gonzalez, Brehm & Perez (2003) have analyzed the genetics of the populations of Spain and North Africa and reached a similar conclusion: "our results are in agreement with the gene flow (19.5%) from northwest Africa to the Iberian Peninsula estimated in a recent study of variation in the autosomal CD4 locus and with the evidence of northwest African male input in Iberia calculated at around 20%, using the relative frequency of northwest African Y-chromosome-specific markers in Iberian samples". They report that North African genes are more common in the south. These results have been further confirmed by Zalloua et al (2008), who report genetic traces of male Phoenicians (from present day Lebanon and Tunis) in the populations of southern Spain.

We now examine the question of whether the immigration of North African and Arab peoples into southern Spain has led to a reduction in the IQ of the population. Some data on this have been published by Colom (2002, p. 231), who gives mean IQs for four regions of Spain derived from the standardization sample (n = 1,345) of the WAIS-III (Wechsler Adult Intelligence Scale Three). The four regions, IQs and numbers in the samples consist of (1) the north-east, comprising Catalonia, Valencia and the Balearic Islands (IQ = 102.1; n=359); (2) the north, comprising Asturias, Galicia, Navarre, Pais Vasco, Rioja and Castille-Leon (IQ = 100.5; n=348); (3) the center, comprising Madrid, Castille-La Mancha, Aragon and Extremurada (IQ = 101.2; n=299); and (4) the south, comprising Andalucia and Murcia (IQ = 96.8; n=363). There is little difference between the IQs in the north-east, north and center and none of these is statistically significant. The average of the three regions is 101.3 and there is a drop of 4.5 IQ points from this to an IQ of 96.8 in the south. The 4.5 IQ point difference between the south and the three north-east/north/center regions is statistically significant ($t = 4.75$, $p < .01$, two-tailed).

We now consider whether this result can be confirmed by more extensive data for fifteen Spanish regions. Four predictions are made. The first is that there is a north-south gradient in intelligence measured by attainment in reading comprehension, mathematic ability, and science understanding obtained by 15 year olds in the 2006 PISA (Program for International Student Assessment) and the 2009 PISA studies. These tests are regarded as measures of intelligence because they all are components of general intelligence in Carroll's (1993) taxonomy and as argued in detail in Lynn (2010a, 2010b). For this reason they are designated PISA IQs.

The second prediction is that the north-south gradient in intelligence will be positively correlated with a number of economic and demographic phenomena that are typically associated with IQ at the individual and group level, namely per capita income, literacy, employment, and life expectancy. The third prediction is that the north-south gradient in intelligence should be positively correlated with the number of years that North African peoples occupied the regions in historical times. The rationale for this is that the longer the time of North African occupation, the greater the interbreeding with indigenous populations and the amount of admixture of European and North African peoples. The fourth prediction is that there should be genetic evidence for a greater proportion of North African genes in the south than in the north.

1 Method

Data for scores on reading comprehension, mathematic ability, and science ability of 15 year olds in ten regions of Spain have been obtained from the 2006 PISA (Program for International Student Assessment) studies given in OECD (2007) and for 15 regions of Spain in the 2009 PISA given in OECD (2009). Data for per capita income expressed as a percentage of that in Madrid in 1961 and 2001, and for life expectancy in 1998 are taken from Pastor et al (2010). Data for the percentages literate and percentages of working age employed in 1991 are calculated from the 1991 census and are given by Benach & Yasui (1999). The number of years that North African peoples occupied the regions are given in Times (2001). Data for latitude have been obtained from an atlas.

The PISA studies are designed primarily to provide valid comparisons between national populations and, in some countries, between regional populations for reading comprehension, science, and mathematics attainment. In the Spanish assessments of 2006 and 2009, 13 different booklets were administered comprising different combinations of the total item pool. The purpose of this design was to minimize testing time, whilst at the same time basing assessments upon a wide range of items. The underlying methodology utilized item response theory, specifically a generalization of the Rasch model: the mixed coefficients multinomial logit model (Adams, Wilson & Wang, 1997). This provided item difficulty scores which in principle gave comparable scale estimates across different students using different items. To ensure lack of bias, both after trialling and during calibration, items were removed if they showed evidence of differential functioning either nationally or cross nationally. Population estimates were in the form of plausible values which present several methodological advantages in comparison with classical Item Response Theory (IRT) estimates such as Maximum Likelihood Estimates or Weighted Maximum Likelihood Estimates. These plausible values return unbiased estimates of population parameters, such as the mean, standard deviation or decomposition of the variance. In addition, weighting procedures were employed in order to ensure population representativeness of the sample estimates (OECD, 2010).

PISA 2006 and 2009 used state of the art methodology to provide optimal estimates of student performance in reading, mathematics and science at the regional and national levels. As such, most of our analysis is based on PISA scores, in order to take advantage of this sophisticated methodology. However, despite this it may be considered that some cross check on the validity of this methodology is desirable. In consequence, we applied tests of measurement invariance for categorical data to one of the thirteen booklets, chosen to provide the best coverage of the complete test. Necessarily, this is a partial check on the quality of the data not least because it was only applied to 1,893 from the total sample of 25,887 participants. This analysis also provided a check on the validity of treating these scores as measures of *g*.

To examine the genetic impact of the occupation of Spain by North African peoples, the gene frequencies of two markers for North African ancestry are examined These are the Taql, p1 2f2-8-kb allele for which data are provided by Semino et al (1996), and the Y-chromosome haplogroup E (Hg E) for which data provided by Semino et al (2004).

2 Results

The descriptive statistics for the data are given in Table 1. The first column identifies the regions, the second column gives the years that the regions were occupied by

Table 1. Descriptive statistics for intelligence and related variables in the regions of Spain.

Region	Years North African Occupation	Years North African Occupation	Latitude	PISA 2006 Reading	PISA 2006 Math	PISA 2006 Science	PISA 2009 Reading	PISA 2009 Math	PISA 2009 Science	Income 1961	Income 2001	Life Expectancy 1998	Employment 1991 %	Literate 1991 %
Andalucia	711-1492	781	37	445	463	474	461	462	469	43.7	56.1	96.5	70.1	91.3
Aragon	740-814	74	41	483	513	513	495	506	505	68.1	80.5	98.8	90.7	97.8
Asturias	740-814	40	43	477	497	508	490	494	502	69.6	64.4	97.2	84.0	99.0
Balearic Isles	711-1130	419	40	-	-	-	457	464	461	83.1	93.8	97.1	83.9	97.6
Basque Region	0	0	43	487	501	520	494	510	495	83.1	93.5	98.4	82.9	98.9
Cantabria	0	0	43	475	502	509	488	495	500	76.8	73.6	98.6	83.6	99.9
Castile & Leon	740-910	170	42	478	515	520	503	514	516	51.7	69.4	99.9	85.3	98.7
Castile Mancha	711-1060	349	39	-	-	-	-	-	-	43.1	59.7	99.0	86.8	97.0

Catalonia	0	0	42	477	488	491	498	496	497	101.4	89.6	98.5	88.7	98.0
Ceuta	711-2011	1300	36	-	-	-	423	424	426	-	-	-	69.9	93.0
Extre-madura	711-1220	529	39	-	-	-	-	-	-	36.7	47.6	98.0	69.7	92.3
Galicia	711-814	74	43	479	494	505	486	489	506	44.1	59.2	98.2	85.8	96.6
Madrid	711-1060	349	40	-	-	-	503	496	508	100.0	100.0	100.0	87.6	97.9
Murcia	711-1240	529	38	-	-	-	480	478	484	48.7	64.4	96.7	83.8	94.7
Melilla	711-2011	1300	36	-	-	-	399	409	404	-	-	-	70.9	91.8
Navarre	0	0	42	481	515	511	497	511	509	75.7	94.5	99.6	85.8	99.0
Rioja	0	0	42	492	526	520	498	504	509	74.3	86.2	98.5	88.8	98.9
Valencia	711-1240	529	40	-	-	-	-	-	-	73.0	72..1	97.2	83.7	97.1

Table 2. Correlations for intelligence and related variables in the regions of Spain.

		Years African Occu-pation	PISA 2006 Reading	PISA 2006 Math	PISA 2006 Science	PISA 2009 Reading	PISA 2009 Math	PISA 2009 Science	Income 1961	Income 2001	Life Expect-ancy 1998	Latitude	Employ-ment 1991
PISA 2006 Reading		-.911**											
	N	10											
PISA 2006 Math		-.716**	.858**										
	N	10	10										
PISA 2006 Science		-.739**	.876**	.905**									
	N	10	10	10									
PISA 2009 Reading		-.918**	.872**	.832**	.790**								
	N	15	10	10	10								
PISA 2009 Math		-.944**	.857**	.874**	.888**	.979**							
	N	15	10	10	10	15							
PISA 2009 Science		-.922**	.811**	.880**	.826**	.987**	.973**						
	N	15	10	10	10	15	15						

		Years African Occu-pation	PISA 2006 Reading	PISA 2006 Math	PISA 2006 Science	PISA 2009 Reading	PISA 2009 Math	PISA 2009 Science	Income 1961	Income 2001	Life Expect-ancy 1998	Latitude	Employ-ment 1991
Income 1961		-.450*	.476	.240	.182	.345	.259	.102					
	N	16	10	10	10	13	13	13					
Income 2001		-.456*	.643*	.535	.446	.329	.351	.113	.894**				
	N	16	10	10	10	13	13	13	16				
Life Exp-ectancy 1998		-.538*	.623*	.765**	.676*	.816**	.803**	.794**	.282	.419			
	N	16	10	10	10	13	13	13	16	16			
Latitude		-.940**	.807**	.582*	.728**	.812**	.858**	.834**	.422	.371	.426*		
	N	18	10	10	10	15	15	15	16	16	16		
Latitude		-.940**	.807**	.582*	.728**	.812**	.858**	.834**	.422	.371	.426*		
	N	18	10	10	10	15	15	15	16	16	16		
Employ-ment 1991		-.805**	.872**	.767**	.660*	.850**	.842**	.837**	.529*	.591**	.493*	.716**	
	N	18	10	10	10	15	15	15	16	16	16	18	
Literacy 1991		-.851**	.849**	.782**	.820**	.799**	.841**	.792**	.635**	.642**	.537*	.873**	.862**
	N	18	10	10	10	15	15	15	16	16	16	18	18

* Correlation significant at the 0.05 level; ** Correlation significant at the 0.01 level

North Africans, and the third column gives the number of years that the regions were occupied. Column four gives the approximate latitude of the regions. Columns five through seven give scores on reading comprehension, mathematic ability, and science understanding for 10 regions in the 2006 PISA study. Columns eight through ten give scores on reading comprehension, mathematic ability, and science understanding for 15 regions in the 2009 PISA study. Columns 11 and 12 give regional per capita income in 1961 and 2001. Column 13 gives life expectancy in 1998. Column 14 gives percentages of working age employed in 1991. Column 15 gives the percentages that were literate in 1991. Data for the Canary Islands are not included because these were colonised by Spaniards after the end of the occupation of parts of Spain by North Africans and are therefore not relevant to the hypothesis being investigated. The correlation matrix for the variables is given in Table 2.

We next investigated measurement invariance. To do this, we first divided the regions into south, comprising Andalucia, Murcia, Ceuta, Melilla, and Extremadura, and the north comprising the remaining regions. This division provided a sufficiently large sample for the valid investigation of measurement invariance. This sample was administered 12 maths items, 29 reading items and 17 science items. In a total of 3.37% of cases items were either not graded or not administered. In conformity with current best practice we used multiple imputation as implemented in Mplus to estimate values for this missing data, additionally for a small number (65) of unclassified individuals we imputed the value for the regions. Testing of invariance of mainly dichotomous items requires a deviation in practice as compared with that usually applied to continuous scales, because identification requires that all item thresholds are held invariant. So the order of testing is first a model which combines configural invariance with item thresholds which are fixed across groups. At the second stage, factor loadings are held invariant, which effectively provides a simultaneous test of metric and scalar invariance. For convenience we will refer to these stages as tests of augmented configural and scalar invariance respectively. These tests were initially applied to the first-order factor structure with three latent variables corresponding to mathematics, reading and science. Then a second-order model was estimated with the latent variables for mathematics, reading and science all loading on a single second-order factor, which loadings were free to vary across groups. Finally two models were tested in which the higher-order factor loadings were held invariant: in one the first-order factor means were free across region, and in the second the higher order mean was free. In order to estimate models we used diagonally weighted least squares since this provides unbiased estimates in relatively small samples (Flora & Curran, 2004). The fit statistics which are the average for five imputed data sets (standard deviations in brackets) are shown in Table 3.

Following the usual convention, the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root-mean-square-error-of-approximation (RMSEA) showed a close fit for all models. Examining increments in the CFI shows that the scalar invariant first-order model fit better than the augmented configurally invariant model. The second-order model which was configurally invariant and allowed the first-order factor means to vary across region had identical fit to its first-order equivalent, since the second-order factor was just identified. By all indices, the differences between the model in which the mean of the second-order factor was free to vary across region and the model in which the means for mathematics, reading and science were free were negligible. We may conclude that for this data the PISA tests are invariant across region, and given the highly similar fit, we prefer the more parsimonious model in which the general factor is free to vary across regions.

Table 3. Fit statistics for tests of first- and second-order invariance.

Model	X^2	*df*	RMSEA	CFI	TLI
First-order, invariant τ[a]	3896.5	2971	0.018	0.978	0.977
First-order, invariant τ, Δ	3891.5 (55.1)	3024	0.017 (0.001)	0.979 (0.001)	0.978 (0.001)
Second-order	As above	As above	As above	As above	As above
Second order, invariant Δ, *M*1-3 free	3886.4 (55.5)	3026	0.017 (0.001)	0.980 (0.001)	0.979 (0.001)
Second order, invariant Δ, *M*4 free	3890.9 (55.9)	3028	0.017 (0.001)	0.980 (0.001)	0.979 (0.001)

Note.! = latent intercept," = factor loadings, M1 = mathematics mean, M2 = reading mean, M3 = science mean, M4 = g

[a]Only one set of data converged.

Table 4. Standardized factor loadings for the second-order model.

Item[a]	Maths	Reading	Science	Factor	*g*
1	0.49			Mathematics	0.92
2	0.57			Reading	0.94
3	0.73			Science	0.96
4	0.65				
5	0.67				
6	0.55				
7	0.64				
8	0.42				
9	0.69				
10	0.41				
11	0.42				
12	0.67				
13		0.66			
14		0.48			
15		0.68			
16		0.55			
17		0.61			
18		0.33			
19		0.54			
20		0.63			
21		0.45			
22		0.61			
23		0.48			
24		0.59			
25		0.70			
26		0.41			
27		0.51			
28		0.61			

Item[a]	Maths	Reading	Science	Factor	g
29		0.75			
30		0.43			
31		0.57			
32		0.77			
33		0.29			
34		0.44			
35		0.40			
36		0.62			
37		0.48			
38		0.64			
39		0.62			
40			0.60		
41			0.55		
42			0.69		
43			0.49		
44			0.55		
45			0.66		
46			0.65		
47			0.68		
48			0.50		
49			0.49		
50			0.54		
51			0.64		
52			0.59		
53			0.39		
54			0.70		
55			0.57		
56			0.75		

[a]Item numbers follow the order of items in book 8 for the Spanish 2009 PISA data, with items 11 and 17 for reading removed since factor loadings were close to zero.

Table 4 shows the standardized factor loadings for the second-order scalar invariant model. The loadings of mathematics, reading and science on the second-order factor at 0.92, 0.94 and 0.96 respectively, are large and suggest that the individual tests do not measure much more than the general factor. Given these loadings, and the studies which suggest that the sum of PISA scores provides a good measure of *g* across nations (e.g. Rindermann, 2007), we interpret this general factor, by extension, as a measure of *g* across regions. Within this model the d-score difference for *g* was -0.574, which is equivalent to 8.6 IQ points.

Table 5. Frequencies of the Taql, p1 2f2-8-kb allele and the Y-chromosome haplogroup E (Hg E) in the Near East and North Africa, Italy, southern and northern Spain, and Central Europe.

Location	8-kb allele frequency	Hg E frequency
Lebanon	43.7	19.0
Tunisia	34.1	55.2
Turkey	33.0	13.8
Algeria	28.3	65.6
Greece	27.3	23.8
Italy: south	26.4	23.6
Italy: north	14.1	10.7
Spain: south	5.9	10.0
Spain: north	1.7	6.1
France	3.8	-
Netherlands	3.5	0

The gene frequencies of the two markers for North African ancestry, the Taql, p1 2f2-8-kb allele and the Y-chromosome haplogroup E (Hg E), are given in Table 5. It shows the frequencies of these two markers are high in a number of populations in North Africa and the Near East, lower in southern Italy and southern Spain, lower still in northern Italy and northern Spain, and lowest in Central Europe.

3 Discussion

The hypotheses to be examined were that there are regional differences in Spain in intelligence and in a number of typically related economic and demographic phenomena (per capita income, literacy, employment and life expectancy), and that these regional differences are associated with the number of years that the regions were occupied by North Africans and hence with the amount of genetic admixture with North African peoples, and that there should be genetic evidence for a greater proportion of North African genes in the south than in the north. All these hypotheses have been confirmed.

More specifically, there are eight points of interest in the results. One, the mean IQ measured by the Spanish standardization sample of the WAIS-R was 101.3 in the north-east, north and center, which North Africans never occupied or occupied for a relatively short period, while in the south, which North Africans occupied for many centuries, the mean IQ was significantly lower at 96.8.

Two, these IQ differences between the north-east/north-center and the south were corroborated by results for ten regions obtained in the reading, mathematical, and science abilities of 15 year olds assessed in the PISA 2006 study, and were further corroborated by more extensive results from fifteen regions obtained in the reading, mathematical and science abilities of 15 year olds assessed in the PISA 2009 study. All of the six measures of cognitive ability measured in the two PISA studies ("PISA IQs") were significantly correlated with latitude with correlations ranging between .582 and .858 showing lower PISA IQs in the more southerly latitudes. The validity of PISA data as measures of intelligence at the population level has been shown by Lynn & Meisenberg (2010), who found that these are perfectly correlated at r = 1.0 across 108 nations.

Three, the regional differences in PISA IQs were positively correlated with per capita incomes. For incomes in 1961, the correlations range between .476 and .102, and for incomes in 2001, the correlations range between .643 and .113. All of 12 correlations are positive, and a sign test gives the probability of this as 0.00024, showing that the positive relationship between per capita income and PISA scores is statistically significant.

These positive correlations between population IQs and per capita incomes are predictable because of the numerous studies showing that IQ predicts income at the level of individuals (e.g. Jencks, 1972; Herrnstein & Murray, 1994; Irwing & Lynn, 2006) and at the level of populations, e.g. in the United States (Davenport & Remmers, 1950; McDaniel, 2006) and in the British Isles, France and Italy (Lynn, 1979, 1980, 2010a). It is proposed that at the population level the association between IQ and per capita income arises through a positive feed-back loop in which the population IQ is a determinant of per capita income, and per capita income is a determinant of the population IQ. Thus, the population's IQ is both a cause and a result of its per capita income. The population's IQ is a cause of its per capita income because individuals and populations with high IQs are able to work more efficiently than those with low IQs and consequently command higher incomes. The population's IQ is also a result its per capita income because prosperous populations provide a better environment (better nutrition, health care and education) for the development of their children's intelligence.

Four, the regional differences in PISA IQs were positively and significantly correlated with life expectancy with correlations ranging between .623 and .816. This result is predictable from studies at the individual level showing positive correlations between IQ and longevity, e.g. Batty, Deary & Gottfredson (2007), Batty, Deary & Macintyre

(2007), Batty et al (2008), Corley et al (2009), Gottfredson & Deary (2004), and Hemmingsson (2009).

Five, the regional differences in PISA IQs were positively and significantly correlated (range = .660 to .872) with the percentage of the population that was employed in 1991, when approximately 13 per cent more of the population were employed in the north (83 percent or more) than in the south (e.g. 70 percent in Andalucia and Extremadura). This result is predictable from studies at the individual level showing positive correlations between IQ and employment, e.g. low IQs in the unemployed have been reported in the United States by Toppen (1971) and Herrnstein & Murray (1994, p. 373-5), and in Northern Ireland by Lynn et al (1984). An association at the population level between IQ and unemployment (r = 0.82) has been reported across regions of the British Isles (Lynn, 1979). The likely reason for this is that those with low IQs more frequently lack the skills required to secure employment.

Six, the regional differences in PISA IQs were positively and significantly correlated (range = .782 to .849) with the percentage of the population that was literate in 1991, when literacy was 98 per cent or higher in the north but 92 percent in Andalucia and 91 percent in Extremadura. These differences are predictable from the populations' IQs, because in low IQ populations there are greater numbers who have difficulty in acquiring literacy.

Seven, the hypothesis that the north-south IQ gradient in Spain is associated with the percentage of North African genes in the population was confirmed by the high and significant correlations of regional IQs with the length of time the regions were occupied by North Africans (rs between .716 and .944) and with latitude (rs between .582 and .858). These correlations are predictable because the longer the years of occupation, the greater the interbreeding with indigenous peoples. The hypothesis was further confirmed by the frequencies of two genetic markers for North African ancestry. The first of these is the Taql, p1 2f2-8-kb allele whose frequencies in a number of populations are shown in Table 5. It will be seen that this is an eastern Mediterranean allele with frequencies of between 28.3 and 43.7 percent in the Near East and North Africa, of 27.3 per cent in Greece and 26.4 in southern Italy, both of which experienced considerable immigration from the Near East and North Africa in historical times. In Spain the frequency is 5.9 percent in the south, but only 1.7 per cent in the north. The allele has a low frequency in Central Europe represented by France (3.8 percent) and the Netherlands (3.5 percent).

The second genetic marker for North African ancestry is the Y-chromosome haplogroup E (Hg E). Table 5 gives its frequency in a number of populations in North Africa, the Near East, southern Italy, northern Italy, southern Spain, northern Spain, and central Europe. It will be seen that this is a "Berber" haplogroup with high frequency in North Africa represented by Tunisia (55.2 percent) and Algeria (65.6 percent). It has a lower frequency in southern Italy (23.6 percent) and a still a lower frequency in northern Italy (10.7 percent), reflecting the greater immigration of North Africans into southern Italy. In Spain, the haplogroup has a higher frequency in the south (10.0 percent) than in the north (6.1 percent), resulting from the greater immigration of North Africans into southern Spain. This haplogroup is rare or absent in Central Europe represented by the Netherlands.

Eight, the data presented here for Spain are predictable from those with similar results in the Balkans and in Italy showing higher IQs in more northern latitudes, and higher values for a number of economic, demographic and sociological phenomena including educational attainment, per capita income, literacy, employment, and life expectancy. In all three locations these higher IQs and higher socio-economic values are associated with lower frequencies of alleles from the Near East and North Africa.

References

Aubet, M.E.
(2001). The Phoenicians and the West: Politics, Colonies and Trade. Cambridge: Cambridge University Press.

Batty, G.D., Deary, I.J. & Gottfredson, L.S.
(2007). Premorbid (early life) IQ and later mortality risk: a systematic review. *Annals of Epidemiology* 17: 278-288.

Batty, G.D., Deary, I.J. & Macintyre, S.
(2007). Childhood IQ in relation to risk factors for premature mortality in middle aged persons: the Aberdeen children of the 1950s study. *Journal of Epidemiology and Community Health* 61: 241-247.

Batty, G.D., Shipley, M.J., Mortensen, L.H. & Deary, I.J.
(2008). IQ in late adolescence/early adulthood, risk factors in middle aged and later all-cause mortality in men: the Vietnam experience study. *Journal of Epidemiology and Community Health* 62: 522-531.

Benach, J. & Yasui, Y.
(1999). Geographical patterns of excess mortality in Spain explained by two indices of deprivation. *Journal of Epidemiology and Community Health* 53: 423-431.

Beraldo, S.
(2010). Do differences in IQ predict Italian North–South differences in income? A methodological critique to Lynn. *Intelligence* 38: 456! 461.

Capelli, C., Brisighelli, F., Scarnicci, F., Arredi, B., Caglia, A., Vetrugno, G., Tofanelli, F., Onofri, V., Tagliabracci, A., Paoli G. & Pascali, V.L.
(2007). Y Chromosome genetic variation in the Italian peninsula is clinal and supports an admixture model for the Mesolithic-Neolithic encounter. *Molecular Phylogenetics and Evolution 44*: 228-239.

Capelli, C., Redhead, N., Romano, V., Cali, F., Lefranc, G., Delague, V., Megarbane, A., Felice, A.E., Pascali, V.L., Neophytou, P.I., Poulli, Z., Novelletto, A., Malaspina, P., Terrenato, L., Berebbi, A., Fellous, M., Thomas, M.G. & Goldstein, D.B.
(2006). Population structure in the Mediterranean basin: a Y chromosome perspective. *Annals of Human Genetics 70:* 207-225.

Carroll, J.B.
(1993). *Human Cognitive Abilities.* Cambridge: Cambridge University Press.

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

Colom, R.
(2002). *En los limites de la inteligencia.* Madrid: Ediciones Piramide.

Corley, J., Crang, J.A. & Deary, I. J.
(2009). Childhood IQ and in-service mortality in Scottish army personnel in World War II. *Intelligence* 37: 238-242.

Cornoldi, C., Belacchi, C., Giofre, D., Martini, A. & Tressoldi, P.
(2010). The mean Southern Italian children IQ is not particularly low: a reply to R. Lynn. *Intelligence* 38: 462! 470.

Davenport, K.S. & Remmers, H.H.
(1950). Factors in state characteristics related to average A-12 V-12 test scores. *Journal of Educational Psychology* 41: 110-115.

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.

Flora, D.B. & Curran, P.J.
(2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods* 4: 466-491.

Gonzalez, A.M, Brehm, A, Perez, J.A, Maca-Meyer, N, Flores, C. & Cabrera V.M.
(2003). Mitochondrial DNA affinities at the Atlantic Fringe of Europe. *American Journal of Physical Anthropology* 120: 391-404.

Gottfredson, L.S. & Deary, I.J.
(2004). Intelligence predicts health and longevity, but why? *Current Directions in Psychological Science* 13: 1–4.

Hemmingsson, T.
(2009). Cognitive ability in adolescence and mortality in middle age: a prospective life course study. *Journal of Epidemiology and Community Health* 63: 697-702.

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

Irwing, P. & Lynn, R.
(2006). The relation between childhood IQ and income in middle age. *Journal of Social, Political and Economic Studies* 31: 191-196.

Jencks, C.
(1972). *Inequality.* London: Penguin.

Lynn, R.
(1979). The social ecology of intelligence in the British Isles. *British Journal of Social and Clinical Psychology* 18: 1-12.

Lynn, R.
(1980). The social ecology of intelligence in France. *British Journal of Social and Clinical Psychology* 19: 325-331.

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

Lynn, R.
(2010a). In Italy, north-south differences in IQ predict differences in income, education and infant mortality. *Intelligence* 38: 93-100.

Lynn, R.
(2010b). IQ differences between the north and south of Italy: a reply to Beraldo and Cornoldi, Belacchi, Giofre, Martini & Tressoldi. *Intelligence* 38: 451-455.

Lynn, R., Hampson, S. & Magee, M.
(1984). Home background, intelligence, personality and education as predictors of unemployment in young people. *Personality and Individual Differences* 5: 549-558.

Lynn, R. & Meisenberg, G.
(2010). National IQs validated for 108 nations. *Intelligence* 38: 353-360.

Lynn, R. & Mikk, J.
(2009). National IQs predict educational attainment in math, reading and science across 56 nations. *Intelligence* 37: 305-310.

Lynn, R. & Vanhanen, T.
(2006). *IQ and Global Inequality.* Athens, GA: Washington Summit Publishers.

M.M.C.
(1960). *Sicily.* Encyclopaedia Britannica, Vol.20. Chicago: William Benton.

McDaniel, M.A.
(2006). Estimating state IQ: measurement challenges and preliminary correlates. *Intelligence* 34: 607-619.

Meisenberg, G. & Lynn, R.
(2011). Intelligence: a measure of human capital in nations. *Journal of Social, Political and Economic Studies* 36: 421-454.

Montmollin, M.
(1958). Le niveau intellectuel des recrues du contingent. *Population* 13: 259-268.

OECD
(2007). *Program for International Student Assessment.* Paris: OECD.

OECD
(2009). *Pisa 2006 Technical Report.* Paris: OECD.

OECD
(2010). *Program for International Student Assessment.* Paris: OECD.

Pastor, J.M., Pons, E. & Serrano, L.
(2010). Regional inequality in Spain: permanent income versus current income. *Annals of Regional Science* 44: 121-145.

Pereira, L., Prata, M.J. & Amorim, A.
(2000). Diversity of mtDNA lineages in Portugal: not a genetic edge of European variation. *Annals of Human Genetics* 64: 491-506.

R.T.D.
(1960). *Spain.* Encyclopedia Britannica. Chicago: William Benton.

Rindermann, H.
(2007). The g-factor of international cognitive ability comparisons. *European Journal of Personality* 21: 667-706.

Rushton, J.P. & Čvorović, J.
(2009). Data on the Raven's Standard Progressive Matrices from four Serbian samples. *Personality and Individual Differences* 46: 483-486.

Semino, O., Magri, C., Benuzzi, G., Lin, A.A., Al-Zahery, N., a Battaglia, V., Maccioni, L., Triantaphyllidis, C., Shen, P., Oefner, P.J., Zhivotovsky, L.A., King, R., Torroni, A., Cavalli-Sforza, L.L., Underhill, P.A. & Santachiara-Benerecetti, A.S.
(2004). Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the Neolithization of Europe and later migratory events in the Mediterranean area. *American Journal of Human Genetics* 74: 1023-1034.

Semino, O., Passarino, G., Brega, A., Fellous,M. & Santachiara-Benerecetti, A.S.
(1996). A view of the Neolithic demic diffusion in Europe through two Y chromosome-specific markers. *American Journal of Human Genetics* 59: 964–968.

Semino, O., Passarino, G., Oefner, P.J., Lin, A.A., Arbuzova, S., Beckman, L.E., De Benedictis, G., Francalacci, P., Kouvatsi, A. & Limborska, S.
(2000). The genetic legacy of Paleolithic *Homo sapiens sapiens* in extant Europeans: a Y chromosome perspective. *Science* 290: 1155–1159.

The Times
(2001). *The Times History of Europe.* London: Harper-Collins.

Toppen, J.T.
(1971). Unemployment: economic or psychological? *Psychological Reports* 28: 111-122.

Zalloua, P.A., Platt, D.E., El Sibai, M. & Khalife, J.
(2008). Identifying genetic traces of historical expansions: Phoenician footprints in the Mediterranean. *American Journal of Human Genetics* 83: 633–642.