Rev. Psicol. (Arequipa. Univ. Católi. San Pablo) / Año 2026 / Vol 16 / N° 1 / pp. 173-182
ISSN 2306-0565 versión impresa / ISSN 2311-7397 versión on line

National IQs, Years of Education, and a Difficult Decision

CIs nacionales, años de educación y una decisión difícil

Federico León Eyzaguirre
Universidad San Ignacio de Loyola, Lima, Perú
https://orcid.org/0000-0002-2290-6837
Correspondencia: federicorleone@gmail.com

Abstract:
*In the study "Likely Electromagnetic Foundations of Gender Inequality" across 98 countries, published in Cross-Cultural Research, 2023, 57, 2-3, 239-263, the author of this note reported that, through apparent actions on a genetic polymorphism (ACP*B) and pathogen prevalence, as well as directly, ultraviolet radiation (UVR) impaired cognitive performance - with the final result of increasing gender inequality. A reader complained that the use of Becker's (2019) National IQs dataset, based on Lynn and Vanhanen's national IQs derived from cognitive ability tests, invalidated the study's results and conclusions. However, the Becker National IQs were not used alone but averaged with years of education per country. In the present note, it is demonstrated that reanalyzing the data with the use of years of education alone produces virtually the same outcomes reported in the original study on the basis of a combination of the two variables. That is, the use of Becker's IQs did not invalidate the study's conclusions. Nonetheless, the author retracted the article in order to stimulate improvements in the Becker collection.*

Keywords: Becker's IQs, years of education, reader's complaints, ultraviolet radiation, gender inequality.

Resumen:
*En el artículo "Posibles raíces electromagnéticas de la desigualdad de género" a través de 98 países, publicado en Cross-Cultural Research, 2023, 57, 2-3, 239-263, el autor de esta nota reportó que, aparentemente mediante acciones sobre un polimorfismo genético (ACP*B) y la prevalencia de patógenos, así como directamente, la radiación ultravioleta (UVR) afecta negativamente el rendimiento intelectual – con el resultado final de incrementar la desigualdad*

https://doi.org/10.36901/psicologia.v16i1.1954

de género. Un/una lector/a se quejó de que el uso de la base de datos de coeficientes intelectuales de Becker, basados en los de Lynn y Vanhanen (derivados de pruebas de habilidades cognitivas), invalidaba los resultados y conclusiones del estudio. Sin embargo, los coeficientes de Becker no se usaron solos, sino promediados con el número de años de educación. En la presente nota, se demuestra que el reanálisis de los datos usando solamente el número de años de educación por país produce virtualmente lo mismos resultados que fueron obtenidos de la combinación de años de educación y coeficientes intelectuales de la colección de Becker. Es decir, el uso de los coeficientes intelectuales no invalidó las conclusiones del estudio. Sin embargo, el autor retractó el artículo con el fin de estimular la mejoría de la colección de Becker.

Palabras clave: Cocientes intelectuales de Becker, años de educación*, quejas de una lectora, radiación ultravioleta, desigualdad de género.*

Background

In the study "Likely Electromagnetic Foundations of Gender Inequality" across 98 countries, published in *Cross-Cultural Research*, 2023, 57, 2-3, 239-263, the author of the present article reported that, through apparent actions on a genetic polymorphism (ACP*B) and pathogen prevalence, as well as directly, ultraviolet radiation (UVR) seemingly impaired cognitive performance - with the final result of increasing gender inequality. One of the variables used in the study was the 'NIQ_QNWSASGEO' column of Becker's (2019) National IQs dataset, based on a combination of Lynn and Vanhanen's (2002, 2012) national IQs derived from cognitive ability tests, scores from PISA, TIMSS, and PIRLS standardized assessments of educational achievements (expressed as IQ points), and assignment of neighbor countries' averaged scores to missing cases. The measures of intelligence utilized by Lynn and Vanhanen (2002, 2012), Raven's Progressive Matrices and Wechsler's Adult (or Children) Intelligence Scale, target, respectively, fluid and a mix of fluid and crystalized intelligence. The PISA, TIMSS, and PIRLS scores, which have countries as units of measurement, are regarded as measures of crystallized cognitive ability (Rindermann, 2007). Whereas Rindermann (2007) called the common construct "complex cognitive ability", the use of the more descriptive term "cognitive performance" is preferrable.

As he wrote the article, the author paid attention to Ebbesen's (2020) criticism of the Lynn-Vanhanen IQ estimates – referring to lack of representativeness of some samples – and, having doubts on the precision of the Becker (2019) measurements, sought to strengthen the measure of cognitive performance by averaging his National IQs with the Human Development Report's actual number of years of education (UNDP, 2011, Table 1). "One year of education may add four IQ points (Hegelund *et al*.,

2020) and the sum was justified by the correlation observed between the two variables in the 98 countries of the study (r = .77, p < .001)" (León, 2023, p. 244). The averaging was a key decision for the acceptance of the manuscript by one of the journal reviewers.

The Challenge

A reader of the article claimed that 1. The National IQ dataset I utilized is derived from small and unrepresentative samples; hence, it may not be adequate for calculating nationally representative, globally comparable measures of cognitive performance. 2. That the adjustments made to the dataset do not improve the original data; therefore, my data set is not fit for use to conduct the analyses of the study and draw the conclusions made. 3. That the National IQ dataset used in the article contains biased assumptions around IQ and geographical location/race/ethnicity. 4. That the article's reliance on the National IQ dataset brings into question the validity of the conclusions.

Regarding point 1, the author of the study generally agrees with the reader's criticism of small and nonrepresentative samples from some of the countries targeted by Lynn and Vanhanen (2002, 2012). However, small sample size and weak representativeness do not bias a study; they impinge on its reliability and precision. This is why the author strengthened the measure of cognitive performance by averaging Becker's (2019) scores with years of education, based on very large and representative national samples (UNDP, 2011).

Regarding point 3, the author assumes the reader was following Ebbesen's (2020) view that "IQ scores are a poor measure of cognitive ability, prone to cultural biases in testing paradigms" (p. 1). More specifically, Ebbesen (2020) criticized the presence of extremely low average IQs in some countries and claimed racial bias considering the difference calculated between the average European IQ and that for Sub-Saharan African countries (−1.6 standard deviations). This difference, however, has a parallel in terms of years of education (UNDP, 2011), which amounts to 1.96 SDs in the dataset utilized by the author to compare 24 European countries with 21 Sub-Saharan African ones. However, the measurement of years of education is not racially biased, which contradicts Ebbesen's argument about the extent of the difference between Europeans and Sub-Saharan Africans.

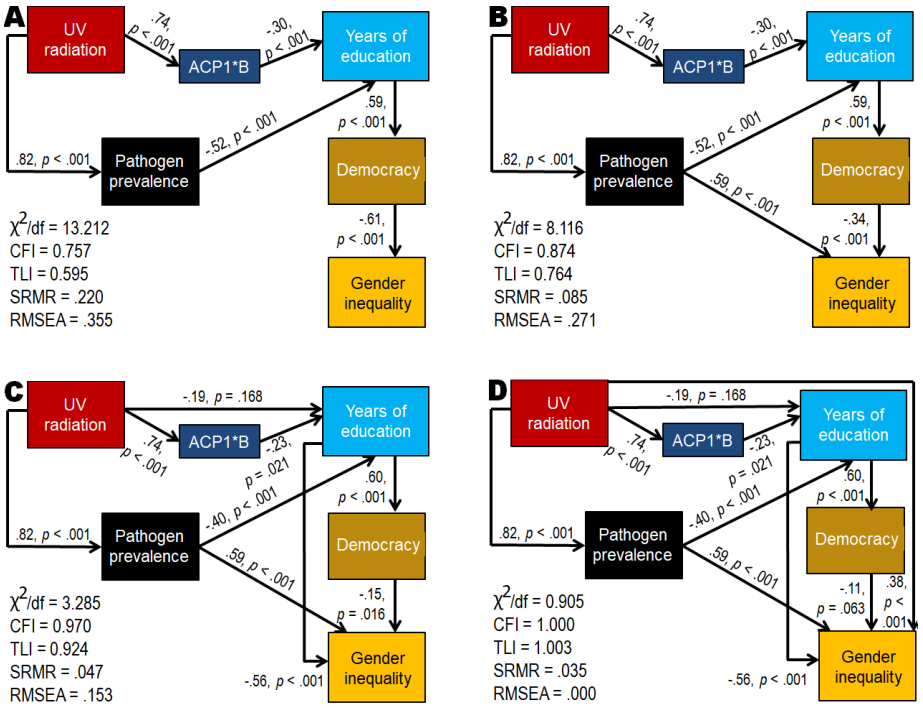
Post-Publication Analyses

To evaluate whether the averaging of Becker's (2019) National IQs variable with UNDP's years of education biased the measurement of cognitive performance (point 2) and whether the validity of the conclusions of the study was impaired by the measurement operations (point 4), I recalculated the values in my original figures 4-7 discarding Becker's (2019) data, that is, estimating cognitive performance only on the basis of years of education. Preparatory work leading to the evidence sustaining my study conclusions was shown on Fig. 4 of the published article. The recalculated results of original Fig. 4 are depicted on Fig. 1 here. The

mediation models originally evaluated showed improved indicators of fit to the data from Fig. 4A to 4B to 4C to 4D and this was replicated by the models depicted on Figs. 1A through 1D here. It can be seen on Figs. 1A and 1B that, whereas the indicators of model fit increased moderately, the path coefficients remained virtually identical. However, in the case of the results depicted on figures 1C and 1D, the direct UVR → years of education coefficient did not attain significance and similar was the case of the democracy → gender inequality path on Fig. 1D.

Figure. 1.

Recalculation of the values shown on original Fig. 4 of "Likely Electromagnetic Foundations of Gender Inequality" using only years of education as estimator of cognitive performance

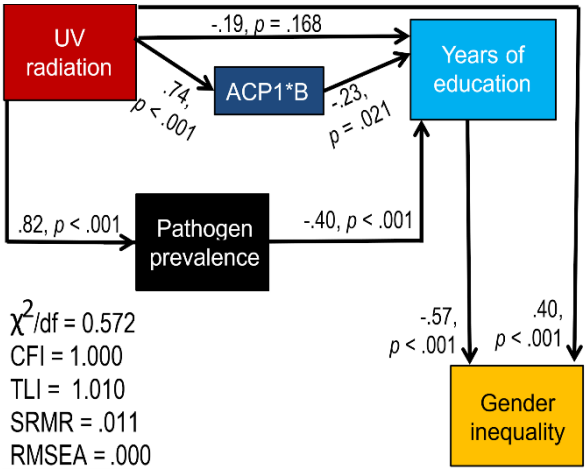


A better mediation model obtained in the original study was that of original Fig. 5, which required the exclusion of democracy and the addition of a direct UVR → gender inequality path. The exclusion of democracy is also justified by the results shown on Figs. 1C and 1D below. The results of recalculating the model in original Fig. 5 using only years of education as estimator of cognitive performance are depicted on Fig. 2 here. With the exception of the non-significant

UVR → years of education linkage, the model emerged quite similar to that in original Fig. 5. Exclusion of the non-significant linkage slightly impaired model fit ($\chi 2$ = 0.911, CFI = 1.000, TLI = 1.002, SRMR = .015, RMSEA = .000), but left the path coefficients virtually invariant in terms of significance (although pathogen prevalence → years of education increased negatively to -.52).

Figure 2.

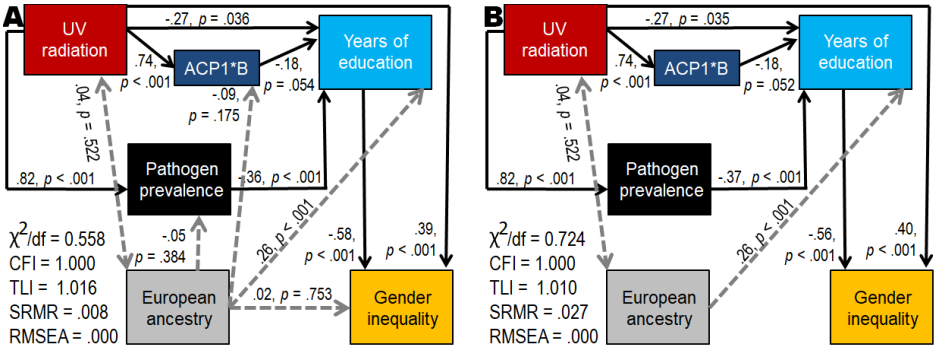
Recalculation of the values shown on original Fig. 5 of "Likely Electromagnetic Foundations of Gender Inequality" using only years of education as estimator of cognitive performance



To evaluate the robustness of the model in Fig. 5 of the published study vis-à-vis the effects of European ancestry, I produced original Fig. 6. Fig. 3 below presents results of a replication utilizing years of education. Like in original Fig. 6, the results shown on Fig. 3 here indicated robustness of the model. Moreover, in this case, every element, including the UVR → years of education linkage, emerged identical or very similar to the values reported in the original study.

Figure 3.

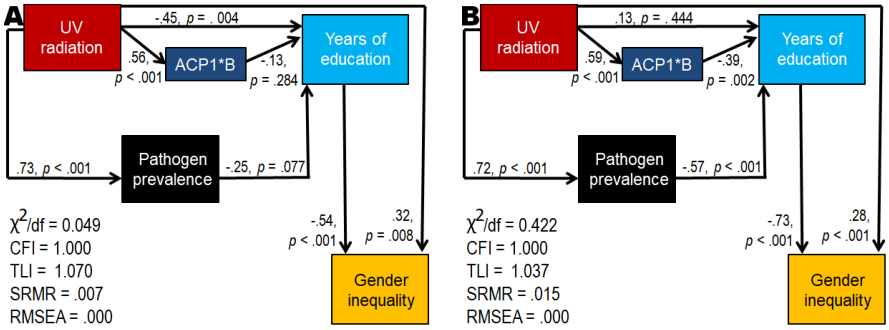
Recalculation of the values shown on original Fig. 6 of "Likely Electromagnetic Foundations of Gender Inequality" using only years of education as estimator of cognitive performance. (A) All paths from European ancestry to endogenous variables are included. (B) Only the significant path is included. Two-headed arrows represent correlations



In the original study, the author wanted to check aspects of the external validity of the findings. To this end, he tested the model on original Fig. 5 separately in Eurasia and the remainder of the world. Whereas sample size was 98 in all the previous analyses, it was reduced to 51 in Eurasia and 47 in the remainder of the world. It can be seen on Fig. 4 here that the use of years of education produced results very similar to those of the original analyses. The model emerged only slightly less fit to the data than in the original analyses.

Figure 4.

Recalculation of the values shown on original Fig. 6 of "Likely Electromagnetic Foundations of Gender Inequality" using only years of education as estimator of cognitive performance. (A) All paths from European ancestry to endogenous variables are included. (B) Only the significant path is included.



It is concluded from these analyses that the averaging of Becker's (2019) National IQs variable with years of education did not bias the measurement of cognitive performance in León (2023) (reader's issue 2) and that the validity of the conclusions of the study was not impaired by the measurement operations (reader's issue 4).

Discussion

The results of the re-analyses of data, by demonstrating that basically the same relationships with third variables are obtained by years education alone or combined with Becker's (2019) National IQs, strongly suggests that both variables measure the same construct: cognitive performance. Years of education is a righteous indicator of cognitive performance at the national level considering that, to approve the courses taken in a year of studies, the student must demonstrate that he/she learned the verbal and quantitative contents of the subjects of the curriculum. This allows one to infer that the population of Australia, with 12 years of formal education on average, has a cognitive performance above that of the population of Mali, with only two years. Whether improvements of actual years of education by national governments are possible is an empirical question.

The literature's negative treatment of the Lynn and Vanhanen (2002, 2006, 2012) datasets is methodologically justified, but its virulence can be understood as a reaction to Lynn's (1991) Cold Winters Theory, which leaves little hope of intellectual progress to the populations of Sub-Saharan Africa. Lynn's views originate in the work of Jensen (1969) on the racial bases of intelligence, which placed Blacks as genetically endowed with less intelligence than Whites. Lynn (1991) explained the superior intelligence of Caucasoids and Mongoloids compared to Negroids considering that, as they arrived in regions close to the North Pole dozens of millennia ago, needed to secure a food supply by hunting large animals and to keep warm in winter by making fires, clothing, and shelter. "Survival in these difficult conditions acted as a selection pressure favoring enhanced intelligence" (p. 109). Sub-Saharan Africans did not have these opportunities.

On the other hand, the valid averaging of Becker's (2019) National IQs with country's average number of years of education demonstrated in the present analyses suggests that national cognitive performance can be enhanced by increasing the number of years of education. This reasoning not only is based on common sense and the evidence provided by a single study (Hegelund *et al*., 2020); it is also consistent with a number of studies and reviews about IQ heritability that demonstrate the environmental malleability of human intelligence (Sauce & Matzel, 2018; Uchiyama *et al*., 2022).

Conclusion

The reader's commentary does not invalidate the conclusions of the article published in *Cross-Cultural Research*. On the other hand, there are better approaches

to the reporting of country IQs than Becker's (2019). Rindermann (2007) demonstrated that the standardized assessment of educational attainments is quite similar to the Raven and WAIS intelligence measurements in terms of national results. Becker (2019) combined them, but the PISA national scores alone represent a better alternative; their validity is not affected by the sampling deficiencies of Becker's (2019) national IQs (Caro & Kyriakides, 2019) and their usefulness has been demonstrated (Erdogdu & Erdogdu, 2025; OECD, 2023). Given this situation, the reader who complained about the publication in *Cross-Cultural Research* was right concerning one point. Researchers can abandon the use of Becker' (2019) country IQs and work with such alternatives as country's number of years of education or PISA scores. The UNDP report has the advantage of providing scores for a considerably greater number of countries. This strategy could be maintained until Becker's (2019) collection is improved through the replacement of deficient samples (e.g., Somalia) by more rigorous ones. The author retracted the article with the aim of stimulating this process, not as a response to the conclusions of a Sage panel which produced further criticism of the original article.

Financing

This research received no grant for any funding agency in the public, comercial or not-for-profit sectors.

Conflict of interest

The author report there no competing interest to declare.

References

Caro, D., & Kyriakides, L. (2019). Assessment design and quality of inferences in PISA: Limitations and recommendations for improvement. *Assessment in Education: Principles, Policy & Practice*, *26*(4), 363-368.

Ebbesen, C. L. (2020). Flawed estimates of cognitive ability in Clark et al. Psychological Science, 2020. *PsyArXiv*. https://doi.org/10 .31234/osf.io/tzr8c

Erdogdu, F., & Erdogdu, E. (2025). Beyond rankings: a multidimensional analysis of factors shaping student achievement in the PISA 2022 dataset. *Educational Review*, *77*(5), 1585-1626.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review, 39*, 1-123

León, F. R. (2023). Likely electromagnetic foundations of gender inequality. *Cross-Cultural Research*, *57*(2-3), 239-263.

Lynn, R. (1991). The evolution of racial differences in intelligence. *Mankind Quarterly*, *32*(1), 109.

Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Bloomsbury Publishing USA.

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

Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences*. Ulster Institute for Social Research.

Organization for Economic Cooperation and Development (OECD) (2023). *PISA results 2022*. OECD.

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: Published for the European Association of Personality Psychology*, *21*(5), 667-706.

Sauce, B., & Matzel, L. D. (2018). The paradox of intelligence: Heritability and malleability coexist in hidden gene-environment interplay. *Psychological Bulletin*, *144*(1), 26.

Uchiyama, R., Spicer, R., & Muthukrishna, M. (2022). Cultural evolution of genetic heritability. *Behavioral and Brain Sciences*, *45*, e152.

UNDP (2011). *Human Development Report 2011: Sustainability and equity: A better future for all*. United Nations Development Programme.

Recibido: 25 de setiembre de 2025
Revisado: 18 de noviembre de 2025
Aceptado: 6 de febrero de 2026