

Retraction of “Declines in Religiosity Predict Increases in Violent Crime—but Not Among Countries With Relatively High Average IQ”

Psychological Science

1

© The Author(s) 2020

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0956797620941437

www.psychologicalscience.org/PS



The following article has been retracted at the request of the authors:

Clark, C. J., Winegard, B. M., Beardslee, J., Baumeister, R. F., & Shariff, A. F. (2020). Declines in religiosity predict increases in violent crime—but not among countries with relatively high average IQ. *Psychological Science*, 31, 170–183. doi:10.1177/0956797619897915

Clark and colleagues requested that this article be retracted out of concern that some of the measures used in the research were invalid. Specifically, they note that the National IQ data used in their analyses, largely based on Lynn and Vanhanen’s (2012) compilation, are plagued by lack of representativeness of the samples, questionable support for some of the measures, an excess of researcher degrees of freedom, and concern about the vulnerability of the data to bias. They also noted that the cross-national homicide data used in the research are unreliable, given that many countries included in the data set provided no actual data on homicides that had occurred. Instead, in these countries, homicide rates were estimated on the basis of other variables that may or may not be closely related to homicide rates. Importantly, some of the variables used to create the estimates were confounded with variables of interest in the research. When the authors reanalyzed the data without the imputed values, the reported effects were no longer apparent.

In the conclusion of their request for retraction, the authors reflected that although articles with certain types of errors may still be helpful to have in the literature,

they do not believe theirs falls into that category. They explicitly expressed concern that leaving the article in the literature could “prolong the use of Lynn & Vanhanen’s cross-national IQ measures.”

As Editor of *Psychological Science*, I have decided to honor the authors’ request and retract this article. I hope that this action on the part of the authors and the journal will encourage all researchers to exercise extreme care in selection and use of the data sets on which they base their analyses, conclusions, and interpretations. Critiques of Lynn and Vanhanen’s (2012) National IQ data were available in the literature prior to the publication of Clark et al. (2020). It is unfortunate that these critiques were not consulted, thereby potentially avoiding publication and the necessity for retraction.

—Patricia J. Bauer
Editor in Chief

Reference

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

RETRACTED: Declines in Religiosity Predict Increases in Violent Crime—but Not Among Countries With Relatively High Average IQ



Cory J. Clark¹, Bo M. Winegard², Jordan Beardslee³, Roy F. Baumeister^{4,5}, and Azim F. Shariff⁶

¹Department of Psychology, Durham University; ²Department of Psychology, Marietta College; ³Department of Psychological Science, University of California, Irvine; ⁴School of Psychology, University of Queensland; ⁵Department of Psychology, Florida State University; and ⁶Department of Psychology, University of British Columbia

Psychological Science
2020, Vol. 31(2) 170–183
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0956797619897915
www.psychologicalscience.org/PS


Abstract

Many scholars have argued that religion reduces violent behavior within human social groups. Here, we tested whether intelligence moderates this relationship. We hypothesized that religion would have greater utility for regulating violent behavior among societies with relatively lower average IQs than among societies with relatively more cognitively gifted citizens. Two studies supported this hypothesis. Study 1, a longitudinal analysis from 1945 to 2010 (with up to 176 countries and 1,046 observations), demonstrated that declines in religiosity were associated with increases in homicide rates—but only in countries with relatively low average IQs. Study 2, a multiverse analysis (171 models) using modern data (97–195 countries) and various controls, consistently confirmed that lower rates of religiosity were more strongly associated with higher homicide rates in countries with lower average IQ. These findings raise questions about how secularization might differentially affect groups of different mean cognitive ability.

Keywords

IQ, intelligence, self-control, religion, religiosity, crime, violence, open data, open materials, preregistered

Received 1/25/19; Revision accepted 10/20/19

Many of the world's great global religions offer inviolable moral rules and threats of supernatural punishment should those rules be violated. By appealing to basic human intuitions and motivations such as desires to conform to a powerful authority, to belong to an organized in-group, and to avoid punishment, religions may constrain and guide human behavior. Overall, religiosity predicts numerous positive life outcomes (McCullough & Carter, 2013); among these are moderate deterrent effects of religiosity on criminal behavior (Baier & Wright, 2001). Religion is associated with higher self-control, which facilitates prosocial behavior and decreases antisocial behavior (McCullough & Carter, 2013). However, the relationship between religiosity and moral behavior has been contested by scholars, and the size of this effect varies substantially, which suggests that there are moderators influencing the inconsistency of this relationship (Shariff, 2015).

Like religion, higher intelligence and self-control (which are positively related, according to Boisvert, Stadler, Vaske, Wright, & Nelson, 2013; see also Zuckerman, Silberman, & Hall, 2013) are associated with lower rates of antisocial behavior and crime (Boutwell et al., 2015; Moffitt, 1993; Moffitt et al., 2011). Although the (likely multiple) reasons for these relationships remain obscure, higher intelligence and self-control afford citizens some unique capacities to function in large, complicated social environments that require sophisticated cooperation and coordination. Religious belief has declined among advanced industrialized societies with highly educated and intelligent populaces (Inglehart &

Corresponding Author:

Cory J. Clark, Durham University, Department of Psychology, Upper Mountjoy, South Rd., Durham DH1 3LE, United Kingdom
E-mail: cory.j.clark@durham.ac.uk

Norris, 2003; Lynn & Vanhanen, 2002, 2006), which suggests that religion may be less uniquely useful for people with relatively higher cognitive ability and self-control. These individuals may be better able than others to structure their lives around abstract moral principles (e.g., utilitarianism; Piazza & Sousa, 2014) and to resist immediate temptations to attain longer-term rewards. Furthermore, groups composed of such people may be better able to create and sustain secular institutions (e.g., democracies, rule of law) that constrain behavior (Kanyama, 2014), foster a sense of fairness, and maintain the trust required for cooperation and economic prosperity (Fukuyama, 1995). Therefore, social groups composed of citizens with relatively high cognitive ability and high self-control may not benefit much from the vivid moral lessons of religion, whereas social groups composed of citizens relatively lower in cognitive ability and self-control may benefit from the particularly powerful and intuitive interdictions and admonishments of many religious narratives. This would not indicate that certain groups need religion more than others; there are many cultural routes to regulating and enforcing norms of cooperation and peace. However, religiosity may be differentially advantageous for populations of different mean cognitive ability, and thus a decline in religiosity may have a different effect on these groups.

We tested the hypothesis that intelligence moderates the relationship between religiosity and moral and immoral behavior. We hypothesized that religion would have greater utility for regulating violent behavior among societies with lower average IQs than among societies with more cognitively gifted citizens. We focused on intelligence (rather than self-control) for two reasons: First, intelligence is associated with lower religiosity (Zuckerman et al., 2013), which suggests that religion may provide less service to people of high intelligence in the modern world; and second (and more practically), intelligence scores are widely available across many countries (whereas self-control data are not), which allows for cross-national analyses. Although IQ, especially when measured cross-nationally, is controversial, myriad analyses suggest that it has high construct validity, even in non-Western countries (Hunt, 2011). For example, educational attainment correlates strongly with both cross-national measures of IQ scores and IQ estimates derived from surrounding regions ($r_s > .90$; Lynn & Meisenberg, 2010). Like all psychometric constructs, IQ is not perfect, and the quality of cross-national data varies. Still, analyses with estimates of cross-national IQs have proven fruitful and have spurred novel theorizing about numerous important group-level outcomes (Rindermann & Thompson, 2011). Of course, all human societies are populated by very intelligent people. In the present research, lower intelligence is

merely a relative description, and it would be more precise to characterize our results as reflecting different degrees of high intelligence across different societies.

We used an easily quantified form of violent behavior—homicide rates—to examine our hypothesis. (Other crime rates are less reliable because of cross-national differences in how they are defined, detected, and recorded; Neopolitan, 1996.) We predicted that lower religiosity would be associated with higher homicide rates among societies with relatively low average intelligence but would have a weaker or nonexistent relationship in societies with relatively high average intelligence. We tested this first using longitudinal data (Study 1) and then again using available modern data and various controls (Study 2).

Study 1

Method

Study 1 examined the within-country association between religiosity and homicide rates over time (i.e., whether these two variables change in tandem over time) and whether the nature of this association varied on the basis of the country's average IQ. All countries and time points for which the relevant data could be obtained were included. This resulted in models that contained up to 1,046 observations from 176 countries covering a span of 65 years.

Religiosity. Country-level religiosity was operationalized as the percentage of the population that practiced religion (Association of Religion Data Archives, or ARDA; Maoz & Henderson, 2013). Every half-decade from 1945 to 2010, ARDA provided estimates of the average (over the previous 5-year period) percentage of the population that was affiliated with any religious party. To our knowledge, these are the best available country-level longitudinal data for religiosity.

IQ. No large-scale country-level longitudinal data for IQ exist (e.g., data on the Flynn effect include only 31 countries; Pietschnig & Voracek, 2015), so we used three separate (but related) average IQ estimates by country drawn from the National IQ (NIQ) data set (Becker, 2019); these estimates are labeled NIQ_LV12GeoIQ (henceforth referred to as LV12GeoIQ), NIQ_QNWSAS (henceforth referred to as NIQ), and NIQ_R (henceforth referred to as RIQ) in the data set. LV12GeoIQ is a set of psychometric test data from Lynn and Vanhanen (2012), with missing nations supplemented by geographic means of neighboring countries. NIQ_QNWSAS is a set of combined data from school assessment studies—mainly Progress in International Reading Literary Study (PIRLS), Programme for International Student Assessment (PISA), and Trends

in International Mathematics and Science Study (TIMSS)—and psychometric test data from Lynn and Vanhanen (2012), weighted and adjusted for sample size, data quality, and population composition without geographical replacement (i.e., relevant samples were obtained from each included country). Becker (2019) described the NIQ data set as less in quantity but higher in quality than the LV12GeoIQ data set. RIQ data (Becker & Rindermann, 2016) were calculated from Lynn and Vanhanen's (2012) data and school assessment studies (mainly PIRLS, PISA, and TIMSS) corrected for schooling such that populations with lower school attendance rates were adjusted slightly upward (these data also include geographical replacement).

Note that all three data sets are based, at least in part, on Lynn's data (Lynn & Vanhanen, 2012). To our knowledge, these are the most complete and well-validated country-level IQ data available (Lynn & Meisenberg, 2010), but the quality of the data varies by country. We included the NIQ data set precisely because it attempts to correct and adjust for differences in data quality. Lynn and Meisenberg (2010) thoroughly discuss the validity of Lynn's data, but a few points are worth mentioning: (a) These country-level data are strongly correlated with educational attainment, gross domestic product (GDP) per capita, and various health outcomes ($r_s > .60$); (b) within-country IQ studies are highly correlated ($r = .92$); and (c) the date the IQ studies were conducted (some decades ago, some more recent) does not influence the relationship between IQ and (more recent) educational attainment, which suggests that the year the IQ data were collected does not substantially reduce their predictive validity. This all supports our use of these time-invariant (time-stable) IQ data as estimates of country-level IQ. Note also that noise in the data, if anything, should obscure our hypothesized pattern of results.

Homicide and GDP. Annual homicide rates by country over time (our dependent variable) were drawn from the Clio Infra data sets and are available beginning in the 1800s (Bierman & van Zanden, 2014). Because of the limited availability of other relevant time-varying covariates, the only time-varying covariate included in Study 1 was GDP (The World Bank, 2017a).¹ GDP data were available beginning in 1960. Because GDP had a very large positive skew and the range was much larger than the range for the other analysis variables (range of original GDP variable = 35.4–116,612.9), GDP was square-root transformed prior to analysis. Additionally, because religiosity was collected in half-decade intervals, homicide rates and GDP were averaged in 5-year intervals to align with religiosity. Table S2 in the Supplemental Material available online reports correlations between all variables within each 5-year time period.

Time. All models also controlled for measurement year with a series of binary variables (less one) to account for historical changes in homicide rates. This technique is advantageous because it allows the model to account for natural changes over time without imposing a structure (e.g., linear, quadratic) on the functional form of change. Because of data constraints and the need for overlapping assessments for the variables, the time frame for Study 1 was 1945 to 2010 for models without GDP and from 1960 to 2010 for models that controlled for GDP.

Analytic plan. We used fixed-effects, within-country linear regressions (Allison, 2009) to examine (a) whether changes in religiosity were associated with simultaneous changes in homicide rates and (b) whether the strength of this association varied on the basis of country-level IQ. These models are appropriate for panel data because the time points are nested within country and the estimates are adjusted for this dependence. The primary advantage of this strategy is that the models automatically control for time-stable variables that might differ between countries (geographic location, stable population, and environmental characteristics). In essence, each country is used as its own control variable (Allison, 2009). This strategy therefore limits possible third-factor explanatory variables to unobserved within-country factors that changed during the study period.

Because of the focus on within-country variability, it was not possible to obtain main effects for unchanging variables (i.e., IQ; although this was tested in Study 2). For example, time-stable variables might explain average differences in homicide rates between countries, but they do not explain why a particular country's homicide rates might fluctuate over time. Using these models, we were able to examine whether country-level homicide rates systematically increased as a country decreased in religiosity.

Although it is not possible to obtain main effects for time-stable variables, it is possible to examine interactions between time-stable (in our case, IQ) and time-varying (i.e., religiosity) variables. Thus, we were able to use these models to examine the critical question of whether the association between religiosity and homicide rates varied for countries with different average IQ levels. All fixed-effects models were estimated with robust standard errors. The general structure of the fixed-effects models used in the present study is as follows (based on Allison, 2009):

$$\text{homicide}_{it} = \mu_t + \beta_1 \text{religiosty}_{it} + \beta_2 \text{IQ} \times \text{Religiosity}_{it} + \beta_3 \text{GDP}_{it} + \sum \beta_z \text{measurementyear}_{it} + \alpha_i + \epsilon_{it},$$

Table 1. Results of Fixed-Effects Linear Regressions Probing Within-Country Associations Between Changes in Religiosity and Simultaneous Changes in Homicide Rates by Average Country IQ (Study 1)

Model and predictor	<i>b</i>	95% CI	Robust <i>SE</i>	<i>p</i>
Model 1 (<i>N</i> = 176, obs. = 1,046)				
Religion	0.04	[-0.05, 0.14]	0.05	.350
Model 2a (<i>N</i> = 136, obs. = 922)				
Religion	-2.82	[-5.21, -0.43]	1.21	.021
Religion × NIQ	0.03	[0.01, 0.06]	0.01	.018
Model 2b (<i>N</i> = 173, obs. = 1,038)				
Religion	-3.43	[-5.61, -1.24]	1.11	.002
Religion × LV12GeoIQ	0.04	[0.01, 0.06]	0.01	.002
Model 2c (<i>N</i> = 173, obs. = 1,038)				
Religion	-2.98	[-4.96, -0.99]	1.01	.004
Religion × RIQ	0.03	[0.01, 0.05]	0.01	.003
Model 3 (<i>N</i> = 164, obs. = 864)				
Religion	-0.07	[-0.20, 0.05]	0.06	.255
GDP	-0.04	[-0.08, -0.01]	0.02	.010
Model 4a (<i>N</i> = 130, obs. = 762)				
Religion	-2.06	[-4.45, 0.32]	1.20	.089
Religion × NIQ	0.02	[0.00, 0.05]	0.01	.090
GDP	-0.03	[-0.06, 0.00]	0.01	.073
Model 4b (<i>N</i> = 163, obs. = 861)				
Religion	-2.88	[-4.59, -1.18]	0.87	.001
Religion × LV12GeoIQ	0.03	[0.01, 0.05]	0.01	.001
GDP	-0.03	[-0.06, 0.00]	0.02	.091
Model 4c (<i>N</i> = 163, obs. = 861)				
Religion	-2.46	[-4.02, -0.91]	0.79	.002
Religion × RIQ	0.03	[0.01, 0.04]	0.01	.002
GDP	-0.03	[-0.06, 0.00]	0.02	.090

Note: Homicide rates and gross domestic product (GDP) were averaged in 5-year intervals to align with religiosity. In all models, we also controlled for historical changes by including a series of dummy-coded time variables representing each of the measurement years (less one). LV12GeoIQ, NIQ, and RIQ are the three country-level average IQ estimates drawn from the National IQ data set (see Becker, 2019). *N* = number of unique countries included in the analysis; obs. = observation count; CI = confidence interval.

where homicide_{it} refers to the homicide rate for country i at time t , μ_t is the intercept for time t , religiosity_{it} refers to the religiosity score for country i at time t , $\text{IQ} \times \text{Religiosity}_{it}$ refers to the interaction between IQ and religiosity for country i at time t , $\beta_3 \text{GDP}_{it}$ refers to the GDP for country i at time t , $\sum \beta_z \text{measurementyear}_{it}$ is the sum of the effect of all dummy-coded time variables for country i at time t , α_i is the combined effect of unobserved time-invariant variables for country i , and ε_{it} is an error term for country i at time t .

No alternate models were tested that are not reported in this article (with the exception of pre-peer-review models that included a lower quality measure of country-level IQ but that demonstrated very similar patterns of results as those reported here). All analyses for Study 1 were conducted in Stata 14 (StataCorp, 2015), and all data and analysis code for this study are publicly available on the Open Science Framework at <https://osf.io/ecdrf/>.

Results

First, we examined the extent to which change over time in religiosity was, on average, associated with change over time in homicide rates before and after adding GDP as a covariate (see Table 1, Models 1 and 3). All models controlled for measurement year as described above, but these variables are excluded from Table 1 for space reasons. Results showed that on average, religiosity was not significantly associated with homicide rates over time, regardless of whether analyses controlled for GDP.

Next, we added interactions between religiosity and each of the three time-invariant IQ variables in models with and without GDP (each product term was examined in its own model; see Table 1, Models 2a, 2b, 2c, 4a, 4b, and 4c). The interaction was significant ($p < .005$ in four of the six models, $p < .05$ in five of the six models, and $p < .091$ in all six models).² This interaction

Table 2. Results of Post Hoc Fixed-Effects Linear Regressions Probing the Religiosity × IQ Interactions in Models Predicting Homicide Rates (Study 1)

Model type and parameter description	<i>b</i>	95% CI	<i>p</i>
Models with Religiosity × NIQ	0.02	[0.00, 0.05]	.090
Religiosity estimate for ~ +1 <i>SD</i> IQ country	0.10	[-0.05, 0.26]	.179
Religiosity estimate for ~ average IQ country	-0.11	[-0.25, 0.03]	.114
Religiosity estimate for ~ -1 <i>SD</i> IQ country	-0.33	[-0.70, 0.05]	.086
Models with Religiosity × LV12GeoIQ	0.03	[0.01, 0.05]	.001
Religiosity estimate for ~ +1 <i>SD</i> IQ country	0.15	[0.06, 0.23]	.001
Religiosity estimate for ~ average IQ country	-0.15	[-0.30, -0.01]	.034
Religiosity estimate for ~ -1 <i>SD</i> IQ country	-0.46	[-0.76, -0.15]	.004
Models with Religiosity × RIQ	0.03	[0.01, 0.04]	.002
Religiosity estimate for ~ +1 <i>SD</i> IQ country	0.12	[0.04, 0.20]	.004
Religiosity estimate for ~ average IQ country	-0.14	[-0.28, 0.00]	.046
Religiosity estimate for ~ -1 <i>SD</i> IQ country	-0.40	[-0.68, -0.11]	.007

Note: Homicide rates and gross domestic product (GDP) were averaged in 5-year intervals to align with religiosity. In all models, we controlled for temporal changes by including a series of dummy-coded time variables representing each of the measurement years (less one) and GDP. Estimated religiosity coefficients for different average IQ levels were obtained by recentering IQ variable. In each model, “-1 *SD* IQ” was approximately 1 *SD* below the sample mean (~80), “average IQ” was around the sample mean (~90), and “+1 *SD* IQ” was approximately 1 *SD* above the sample mean (~100). Because the precise values for the means and standard deviations varied for the three measures of IQ, we probed the interaction with even values that were roughly representative of the means and standard deviations. The precise means and standard deviations for the analytic sample were as follows: NIQ: $M = 86.51$, $SD = 13.56$; LV12GeoIQ: $M = 87.54$, $SD = 11.01$; RIQ: $M = 86.82$, $SD = 11.90$. LV12GeoIQ, NIQ, and RIQ are the three country-level average IQ estimates drawn from the National IQ data set (see Becker, 2019). CI = confidence interval.

suggests that the nature of the association between religiosity and homicide rates over time varied on the basis of the country’s average IQ.

Probing of the significant interactions suggested that increases in religiosity were associated with simultaneous decreases in homicide rates for countries with lower average IQs only. For example, for countries with average IQs approximately 1 standard deviation below the overall mean, declines in religiosity were associated with increases in homicide rates (*bs* from -0.46 to -0.33; see Table 2). However, the positive values for the interaction terms indicated that the slope representing the association between religiosity and homicide rates systematically became more positive as average IQs were higher. For example, in countries with average IQs approximately 1 standard deviation above the mean, the association between religiosity and homicide was near zero or positive (*bs* from 0.10 to 0.15; see Table 2).³

Study 2

Study 1 demonstrated that declines in religiosity from 1945 to 2010 predicted concurrent increases in homicide rates among countries with relatively low average IQs only. In Study 2, we sought to confirm these results with available modern data, which allowed for the inclusion of additional control variables and tests with

multiple operationalizations of religiosity to confirm that the results are not limited to ARDA estimates and to eliminate concerns that the present results were influenced by the Flynn effect (because all data are stable across time).

Method

In Study 2, we examined the interaction between country-level IQ and religiosity on homicide rates. All countries for which the relevant data could be obtained were included. Given that there are no objective best measures of religiosity and IQ or an objective best list of relevant control variables, we conducted a multiverse analysis using three operationalizations of religiosity, three operationalizations of IQ, all possible combinations of four control variables, and additional interactions between those control variables and each operationalization of religiosity. Multiverse analysis reports all (or at least many) of the conceivable statistical models to eliminate researcher degrees of freedom (Steege, Tuerlinckx, Gelman, & Vanpaemel, 2016). Multiverse analysis is preferred to preregistrations of specific analysis plans because preregistrations allow researchers to specify the one statistical model that they think is most likely to produce the hypothesized result. In a multiverse analysis, researchers analyze every

single model they could have chosen and report the results for all models, which eliminates entirely (or nearly entirely) researchers' ability to exert control over the results with variable and model selection. If most or all models in a multiverse demonstrate a meaningful effect size for the hypothesized effect, this is much stronger evidence that the effect is real than demonstrating the effect once in one preregistered model.

Religiosity. Religiosity was operationalized as the percentage of the population affiliated with any religion (Pew Research Center, 2012), the percentage of the population that practices religion (ARDA; Maoz & Henderson, 2013), and the percentage of the population that reports that religion is an important part of its daily life (Crabtree, 2010).

IQ. The same three average IQ estimates by country from Study 1 were again used in Study 2: LV12GeoIQ, NIQ, and RIQ.⁴

Homicide. Per capita homicide rates were drawn from the United Nations Office on Drugs and Crime (2013; the most recent available year of data was used, with the majority from 2012). To our knowledge, this source provides the best available estimates for homicide rates, and so no other operationalizations of homicide rates were included in the multiverse analysis. See the Supplemental Material for secondary analyses using a different operationalization of violence (tourism safety scores), which demonstrated very similar patterns as those observed for homicide rates.

Controls. In all possible combinations, we controlled for various other factors generally regarded to be related to homicide rates: GDP and the Gini index of income inequality (2015 CIA *World Factbook*; latest available estimates were used if 2015 estimates were not available), population density (The World Bank, 2015), and educational attainment (secondary education completion rate; The World Bank, 2019).⁵ At the request of a reviewer, we analyzed additional models (in models with all controls) also controlling for the interactions between each operationalization of religiosity and GDP, each operationalization of religiosity and income inequality, and each operationalization of religiosity and educational attainment (independently, so only one additional interaction was included at a time). Table S1 in the Supplemental Material contains source information for all variables included in both Studies 1 and 2.

Multiverse analysis. This combination of variables and planned analyses produced 171 possible statistical models with up to 195 countries. All variables were z transformed

prior to analysis, except for GDP, which was square-root transformed as in Study 1.^{6,7} Data were analyzed first in SPSS and then cross-checked in R (Version 3.4.3; R Core Team, 2017). All data and code for this study are publicly available on the Open Science Framework at <https://osf.io/ecdrf/>.

Results

Correlations. As can be seen in the correlation matrix (Table 3), higher homicide rates were associated with lower IQ, GDP, and educational attainment. Higher homicide rates were unrelated to population density and either unrelated (ARDA and Pew) or positively associated (Gallup) with religiosity. Higher IQ was associated with higher GDP, population density, and educational attainment and with lower religiosity and income inequality. Higher religiosity was negatively associated with GDP and educational attainment, positively associated with income inequality, and unrelated to population density.

Multiverse analysis. In separate analyses, homicide rates were regressed on each of the three operationalizations of religiosity, each of the three operationalizations of IQ, and each of the relevant interactions (for nine possible interaction terms), independently and with every possible combination of the four control variables, excluding listwise deletion. This produced a total of 144 possible models. For each of the nine full models (with all four controls), we tested three additional models controlling for the interactions between the relevant operationalization of religiosity and (a) GDP, (b) income inequality, and (c) educational attainment, which produced 27 additional models. Thus, we tested 171 models in total for the multiverse analysis.

We used semipartial r s (the proportion of the variance in homicide rates uniquely explained by the interaction) as estimates of the interaction effect size ($ps < .001$ were coded as .00099). Across all possible models (see Fig. 1), the effect sizes for the interaction between religiosity and IQ ranged from small-medium, semipartial $r = .14$, to medium-large, semipartial $r = .46$ (Cohen, 1992), with a medium average effect size (mean semipartial $r = .30$, $SD = .08$). The interaction was statistically significant ($p < .001$ in 64.9% of models, $p < .010$ in 88.9% of models, $p < .050$ in 97.7% of models, and $p < .078$ in 100% of models). Thus, the multiverse analysis provided very strong support for the hypothesized interaction.⁸

Sample models. To decide which models to expand on for purposes of graphing the interaction, we checked the average semipartial r s for each of the nine interaction terms and selected the smallest (LV12GeoIQ \times Gallup

Table 3. Correlations Between Homicide Rates, All IQ Variables, All Religiosity Variables, and All Control Variables Included in Study 2

Variable and statistic	1	2	3	4	5	6	7	8	9	10
1. Homicide rate	—									
2. NIQ		—								
<i>r</i>	<i>-.421</i>									
<i>p</i>	< .001									
<i>n</i>	146									
3. LV12GeoIQ			—							
<i>r</i>	<i>-.378</i>	.856								
<i>p</i>	< .001	< .001								
<i>n</i>	195	147								
4. RIQ				—						
<i>r</i>	<i>-.375</i>	.870	.978							
<i>p</i>	< .001	< .001	< .001							
<i>n</i>	195	147	199							
5. ARDA religiosity					—					
<i>r</i>	.082	<i>-.477</i>	<i>-.536</i>	<i>-.528</i>						
<i>p</i>	.259	< .001	< .001	< .001						
<i>n</i>	191	140	185	185						
6. Pew religiosity						—				
<i>r</i>	.101	<i>-.500</i>	<i>-.534</i>	<i>-.528</i>	.870					
<i>p</i>	.135	< .001	< .001	< .001	< .001					
<i>n</i>	219	147	197	197	191					
7. Gallup religiosity							—			
<i>r</i>	.244	<i>-.698</i>	<i>-.727</i>	<i>-.750</i>	.715	.730				
<i>p</i>	.003	< .001	< .001	< .001	< .001	< .001				
<i>n</i>	146	125	144	144	141	146				
8. Gross domestic product								—		
<i>r</i>	<i>-.168</i>	.700	.700	.712	<i>-.333</i>	<i>-.310</i>	<i>-.598</i>			
<i>p</i>	.014	< .001	< .001	< .001	< .001	< .001	< .001			
<i>n</i>	212	146	196	196	190	222	144			
9. Gini index									—	
<i>r</i>	.509	<i>-.468</i>	<i>-.507</i>	<i>-.536</i>	.340	.184	.505	<i>-.368</i>		
<i>p</i>	< .001	< .001	< .001	< .001	< .001	.031	< .001	< .001		
<i>n</i>	138	116	138	138	135	138	123	138		
10. Population density										—
<i>r</i>	<i>-.102</i>	.189	.199	.180	<i>-.041</i>	<i>-.106</i>	<i>-.103</i>	.227	.088	
<i>p</i>	.144	.023	.006	.013	.579	.125	.220	< .001	.303	
<i>n</i>	205	143	192	192	189	211	144	212	138	
11. Educational attainment										
<i>r</i>	<i>-.248</i>	.585	.668	.667	<i>-.352</i>	<i>-.303</i>	<i>-.595</i>	.693	<i>-.316</i>	.123
<i>p</i>	.001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	< .001	.114
<i>n</i>	169	126	162	162	160	169	128	170	122	168

Note: Significantly negative correlations are given in italics, and significantly positive correlations are given in boldface. LV12GeoIQ, NIQ, and RIQ are the three country-level average IQ estimates drawn from the National IQ data set (see Becker, 2019). ARDA = Association of Religion Data Archives.

Religiosity), the largest (NIQ × ARDA Religiosity), and the one closest to the overall mean (RIQ × Pew Religiosity). We expanded on these three interactions without any controls and then with all four controls (for six models total). Note that none of the additional included interactions—between each of the three operationalizations of religiosity with (a) GDP, (b) income inequality, and (c)

educational attainment within each of the 27 additional models—were even consistently in the same direction across models, and only 1 of the 27 tested interactions was statistically significant (between income inequality and ARDA religiosity in the models with NIQ, semipartial $r = -.160$, $p = .027$). Thus, we did not test these additional interactions and will not discuss them further. As can be

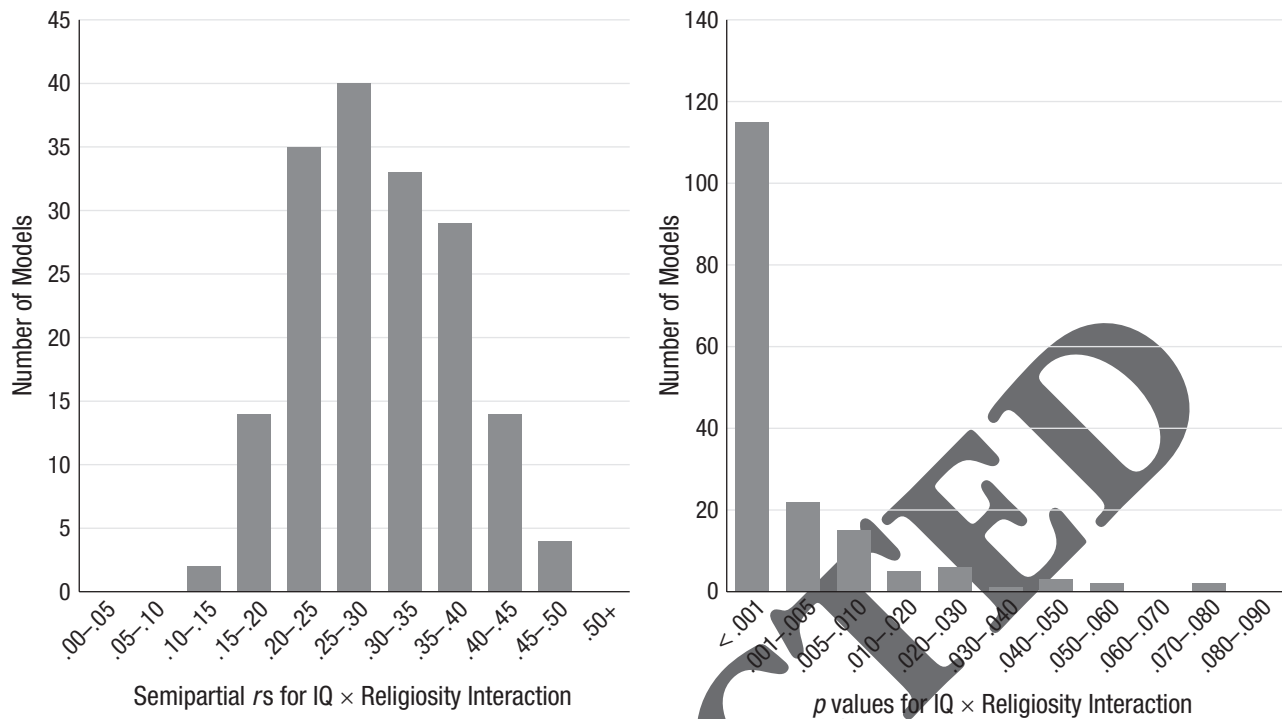


Fig. 1. Histograms from the multiverse analysis showing the distribution of semipartial r_s and p values for the interactions between IQ and religiosity across 171 models (Study 2).

seen in Table 4, higher IQ was a significant predictor of lower homicide rates in four of the six models, religiosity was a significant predictor of lower homicide rates in five of the six models, and their interaction was significant in all six models.

As can be seen in Figure 2, simple slopes 1 standard deviation above and below the overall mean of IQ indicated that in countries with relatively high average IQ (IQ \approx 95–98), there were small to nonexistent relationships between higher religiosity and lower homicide rates, LV12GeoIQ \times Gallup: $b = -0.10$, $p = .452$; NIQ \times ARDA: $b = -0.16$, $p = .092$; RIQ \times Pew: $b = -0.22$, $p = .016$, but in countries with lower average IQ (IQ \approx 70–73), higher religiosity was associated strongly with lower homicide rates, LV12GeoIQ \times Gallup: $b = -0.89$, $p = .004$; NIQ \times ARDA: $b = -2.06$, $p < .001$; RIQ \times Pew: $b = -1.44$, $p < .001$.

Galton's problem and spatial autocorrelation. Galton's problem is an issue with cross-cultural data (and perhaps statistical inference more generally) regarding drawing statistical inferences from nonindependent data. Countries are treated as independent observations, yet neighboring societies (e.g., the United States and Canada) or otherwise historically related societies (e.g., the United States and the United Kingdom) share numerous traits and in some cases are near duplicates of each other, which can bias results in unpredictable ways. For example, if we are

oversampling one particular type of culture (because one culture spreads across numerous countries), that particular culture can have a heavy influence on the overall results. Lines between countries are at least somewhat arbitrary in terms of dividing up distinct populations.

After consulting with several Galton's problem experts, we sought to deal with this issue in three ways. First, following Hruschka and Henrich (2013), we reran the nine full models controlling for dummy-coded world regions. Second, we tested the interactions within world regions and within countries that share the same majority religion to assess whether the interaction is particularly strong or weak in particular world regions or among countries of particular majority religions. Third, we hired a statistical expert to rerun our analyses controlling also for spatial autocorrelation between countries. Thorough results of these additional analyses are reported in the Supplemental Material.

Controlling for world region. Controlling for world regions in the nine full models (now with nine control variables each and 97 to 122 countries each) did have a small influence on the size of the interaction effects but generally did not affect our interpretation of the findings. Six of nine models continued to show significant IQ \times Religiosity interactions with small to medium effect sizes (ARDA models: semipartial $r_s = .190$ – $.234$, $p_s < .006$; Pew

Table 4. Results of Fixed-Effects Linear Analyses Regressing Homicide Rates on IQ, Religiosity, the IQ × Religiosity Interaction, and Controls in Study 2

Model and predictor	<i>F</i>	<i>R</i> ²	β	<i>t</i>	<i>p</i>	95% CI	Semipartial <i>r</i>
Sample models for LV12GeoIQ × Gallup Religiosity (small estimate)							
Model 1 (<i>n</i> = 144)	<i>F</i> (3, 140) = 7.89	.13			< .001		
IQ (LV12Geo)			−0.41	<i>t</i> (140) = −3.45	.001	[−0.67, −0.18]	−.270
Religiosity (Gallup)			−0.20	<i>t</i> (140) = −1.39	.166	[−0.51, 0.09]	−.109
IQ × Religiosity			0.24	<i>t</i> (140) = 2.41	.017	[0.05, 0.50]	.188
Model 2 (<i>n</i> = 111)	<i>F</i> (7, 103) = 6.68	.31			< .001		
IQ (LV12Geo)			−0.20	<i>t</i> (103) = −1.07	.287	[−0.61, 0.18]	−.087
Religiosity (Gallup)			−0.48	<i>t</i> (103) = −2.70	.008	[−0.90, −0.14]	−.221
IQ × Religiosity			0.35	<i>t</i> (103) = 2.76	.007	[0.11, 0.67]	.225
GDP			0.10	<i>t</i> (103) = 0.56	.578	[−0.56, 1.00]	.046
Income inequality			0.49	<i>t</i> (103) = 4.74	< .001	[0.31, 0.77]	.387
Population density			−0.08	<i>t</i> (103) = −0.86	.390	[−0.58, 0.23]	−.070
Education			−0.29	<i>t</i> (103) = −1.84	.068	[−0.66, 0.02]	−.151
Sample models for NIQ × ARDA Religiosity (large estimate)							
Model 1 (<i>n</i> = 140)	<i>F</i> (3, 181) = 26.20	.37			< .001		
IQ (NIQ)			−0.67	<i>t</i> (181) = −8.08	< .001	[−0.90, −0.55]	−.552
Religiosity (ARDA)			−0.84	<i>t</i> (181) = −6.20	< .001	[−1.16, −0.60]	−.423
IQ × Religiosity			0.74	<i>t</i> (181) = 6.06	< .001	[0.52, 1.03]	.413
Model 2 (<i>n</i> = 101)	<i>F</i> (7, 111) = 13.64	.51			< .001		
IQ (NIQ)			−0.49	<i>t</i> (111) = −3.66	< .001	[−0.84, −0.25]	−.266
Religiosity (ARDA)			−1.07	<i>t</i> (111) = −6.33	< .001	[−1.49, −0.78]	−.461
IQ × Religiosity			0.89	<i>t</i> (111) = 5.92	< .001	[0.62, 1.24]	.431
GDP			0.11	<i>t</i> (111) = 0.87	.388	[−0.37, 0.94]	.063
Income inequality			0.37	<i>t</i> (111) = 4.21	< .001	[0.22, 0.62]	.306
Population density			−0.01	<i>t</i> (111) = −0.11	.916	[−0.45, 0.41]	−.008
Education			−0.26	<i>t</i> (111) = −2.38	.019	[−0.67, −0.06]	−.174
Sample models for RIQ × Pew Religiosity (middle estimate)							
Model 1 (<i>n</i> = 195)	<i>F</i> (3, 191) = 19.27	.23			< .001		
IQ (RIQ)			−0.53	<i>t</i> (191) = −6.84	< .001	[−0.68, −0.37]	−.433
Religiosity (Pew)			−0.68	<i>t</i> (191) = −4.69	< .001	[−0.91, −0.37]	−.297
IQ × Religiosity			0.58	<i>t</i> (191) = 4.46	< .001	[0.24, 0.62]	.283
Model 2 (<i>n</i> = 122)	<i>F</i> (7, 114) = 11.01	.4			< .001		
IQ (RIQ)			−0.22	<i>t</i> (114) = −1.33	.187	[−0.60, 0.12]	−.096
Religiosity (Pew)			−0.90	<i>t</i> (114) = −4.71	< .001	[−1.25, −0.51]	−.341
IQ × Religiosity			0.82	<i>t</i> (114) = 4.71	< .001	[0.34, 0.84]	.340
GDP			0.02	<i>t</i> (114) = 0.10	.924	[−0.64, 0.70]	.007
Income inequality			0.39	<i>t</i> (114) = 4.27	< .001	[0.22, 0.61]	.309
Population density			−0.04	<i>t</i> (114) = −0.52	.605	[−0.45, 0.26]	−.038
Education			−0.17	<i>t</i> (114) = −1.29	.198	[−0.46, 0.10]	−.094

Note: LV12GeoIQ, NIQ, and RIQ are the three country-level average IQ estimates drawn from the National IQ data set (see Becker, 2019). CI = confidence interval; ARDA = Association of Religion Data Archives.

models: semipartial *r*s = .209–.258, *ps* < .003). The three Gallup models no longer reached statistical significance but maintained generally small effects in the same direction (semipartial *r*s = .08–.10, *ps* < .260).

Testing within world region. We collapsed the seven world regions into four world regions (Europe and Central Asia, Middle East and Africa, South Asia and East Asia Pacific, and North and Latin America and the Caribbean)

in an effort to get large enough sample sizes to test the interactions within regions. However, even after doing so, we found that the samples were very small across models (18–67 countries each), and so we caution against interpreting any of these specific interaction terms in isolation. Within each of the four world regions, we analyzed each of the nine interaction terms in the base models (without controls) and then again controlling for income inequality only (our analyses were already severely

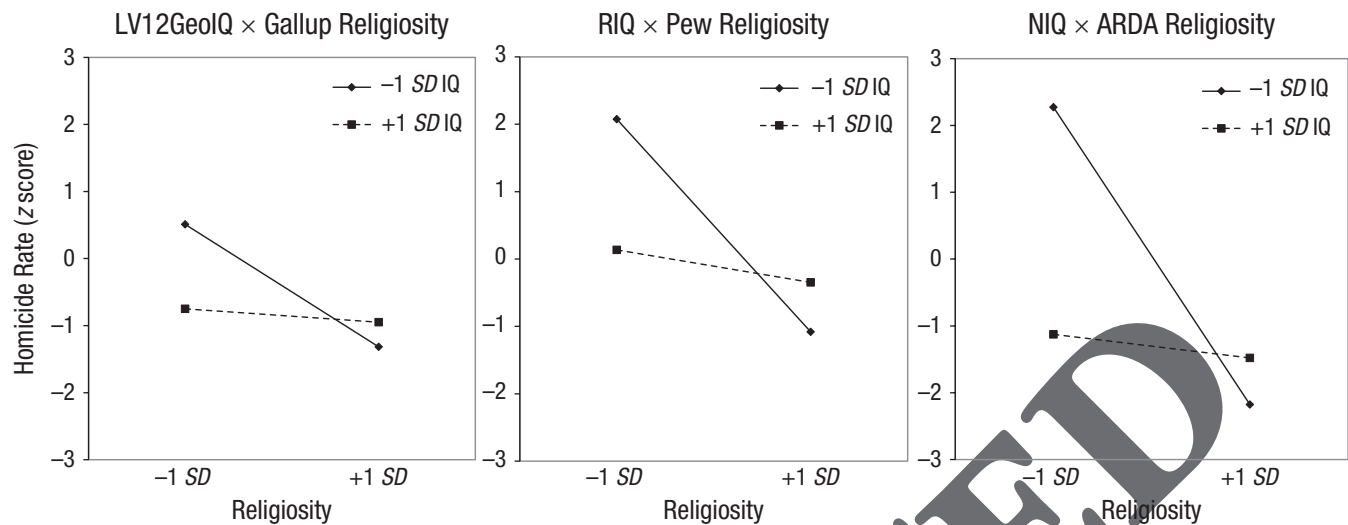


Fig. 2. Interactions between IQ and religiosity on homicide rates with all four controls in Study 2. LV12GeoIQ, NIQ, and RIQ are the three country-level average IQ estimates drawn from the National IQ data set (see Becker, 2019).

underpowered to test models with additional controls, but income inequality did stand out as the most important control in the full multiverse). Because of the small sample sizes, the interactions were rarely statistically significant in any of the world regions. We arbitrarily decided that a semipartial r of .07 or higher for the IQ \times Religiosity interaction term would be a consistent effect, that is, consistent with the IQ \times Religiosity interactions found in the multiverse. Of the 18 models tested within each world region, 9 were consistent in Europe and Central Asia, 10 were consistent in South Asia and East Asia Pacific, 12 were consistent in North and Latin America and the Caribbean, and 12 were consistent in the Middle East and Africa. Thus, the effect did not appear to be particularly absent in any world region, which reduced the likelihood that the effects are not (at least somewhat) generalizable globally.

Testing within majority religion. We repeated these analyses within Christian majority countries (71–124 countries) and Muslim majority countries (23–45 countries). Again, we caution against interpreting any specific interaction, especially for Muslim majority countries, because of the small sample sizes. Here, there at least appeared to be a difference. Of the 18 models tested within each religion, 17 were consistent in Christian majority countries, whereas only 3 were consistent in Muslim majority countries—potentially due to reduced variance in religiosity in Muslim majority countries, which often feature uniformly high levels of reported religiosity. Nevertheless, the difference led us to create two additional dummy variables, whether a country was majority Christian or not and whether a country was majority Muslim or not, and to test whether either of these dummy variables moderated the nine IQ \times Religiosity interactions (in the

base models, without controls). None of the 18 three-way interactions were statistically significant, and so we did not interpret this possible difference between Christian majority countries and Muslim majority countries. However, whereas we are quite certain the pattern is real in Christian majority countries, we are less certain about whether it holds in Muslim majority countries. Table S5 in the Supplemental Material reports the semipartial r s and p values for the IQ \times Religiosity interactions within each of these new models (9 models controlling for region, 18 models within each of the four world regions and within each of the two religions, for 117 additional models total).

Accounting for spatial autocorrelation. Last, we hired a statistical consultant to account for spatial autocorrelation between countries (correlation due to spatial proximity). He ran Bayesian multilevel regressions including a Gaussian process (McElreath, 2016) to account for spatial autocorrelation between countries in 18 models: the 9 main models and the 9 models with the four main controls. The interaction was statistically significant, as indicated by pMCMC values, which refer to the proportion of posterior samples falling below zero (pMCMC < .001 in 1 model, pMCMC < .010 in 6 models, pMCMC < .050 in 8 models, pMCMC < .100 in 13 models, and pMCMC < .228 in all 18 models). He concluded that accounting for spatial autocorrelation weakened but did not abolish the effect. The full report, R code, and output for these analyses are available in the Supplemental Material. Researchers who use his R code should cite him (rather than the present article) as described in the Supplemental Material.

Galton's problem conclusions. Although none of our efforts for dealing with Galton's problem may rule out

concerns related to nonindependence of country-level comparisons completely, they did provide evidence that the effect likely cannot be attributed to one particular world region (although, as noted above, they may be more true of Christianity than Islam). Despite this, that controlling for region weakens the effect suggests the possibility that the strength of the interaction varies at least somewhat by subregion. Future research might conduct multiple within-country or within-region analyses to identify countries or regions that do not display the interaction reported here.

Data auditor. As a final step to testing the robustness of the reported interaction, we hired an external adversarial data analyst to audit and cross-check our results. She cross-checked two additional 171 model multiverses, one with a different standardization approach and one with median-split dummy-coded indicators of each operationalization of religiosity and IQ. The results largely confirmed those reported here. The full auditor report is available in the Supplemental Material. Researchers who use her multiverse R code should cite her (rather than the present article) as described in the Supplemental Material.

General Discussion

Whether religion serves a social function in suppressing antisocial behavior has been discussed for well over 2,000 years, and psychological research has recently begun in earnest to investigate this idea empirically (e.g., Purzycki et al., 2016; Shariff, 2015; see Norenzayan et al., 2016, for a review). Here, we introduce a possible moderator for the contested relationship between religiosity and moral behavior—intelligence. Our results indicated that higher religiosity was largely unrelated to homicide rates in societies with relatively high average intelligence, whereas religiosity was a significant predictor of reduced homicide rates in societies with relatively low average intelligence. Study 1 supported this in an examination of changes over the past 65 years. Study 2 confirmed this pattern in a comparison of the majority of countries in the world at the same time in cross-sectional analyses with various controls. Thus, the results supported our hypothesis that religiosity would have greater violence-detering utility among populations with relatively lower mean cognitive ability than among more cognitively advantaged populations.

Although we scrutinized the reported interaction in several hundred ways and found quite consistent and robust support, our results should be interpreted with caution. All three of our main variables of interest (religion, intelligence, and morality) are multifaceted and challenging to measure and even more challenging to compare across cultures. First, for example, the present results might apply more to some religions than others,

and we imagine the effect could vary in countries experiencing religious conflict. So whereas the interaction may be true in the aggregate, it almost certainly is not true in every type of cultural system. Although we found supportive evidence for the interaction in each of the four world regions we tested, controlling for world region weakened the interaction effect, which suggests that the interaction might vary in strength and significance in different regions. Moreover, whereas we found evidence for the effect in the present and over the past 65 years, the nature of the effect could change in the future as secularization likely continues to increase. Future research should investigate possible variation and potential reasons for it. Second, although we reported the interaction between intelligence and religiosity on homicide rates (mainly because homicide rates are the most reliable cross-national measure of violence), our theorizing focused more on violence or antisocial behavior generally. As reported in the Supplemental Material, we tested the effect with an alternate measure of violence (based largely on citizens' reports of perceived violence in their own country) and found a similar pattern, but future work should explore whether the interaction emerges for other types of violent crime and antisocial behavior (should reliable sources of cross-national violence be identified).

Last, although country-level IQ appears to be an important predictive variable, it is controversial because IQ varies substantially within countries, and such differences may be caused (at least in part) by differences in schooling and other cultural differences (e.g., nutrition). We controlled for at least one sort of education (secondary education completion rate), and the RIQ analyses adjusted for schooling, but we would not be surprised if a thorough index of all educational differences (in both quality and quantity) explained at least a large portion of the present effect (Rindermann & Ceci, 2009). However, we are not sure whether this would be a confound (the effect is driven by education, not intelligence) or a mechanism (higher intelligence leads to better educational systems and participation in those systems). Moreover, given the links between higher self-control with higher intelligence, higher religiosity, and lower antisocial behavior, we suspect that self-control may be an important mediating variable or perhaps even the crucial variable that explains the present results. IQ might also be a proxy for a combination of other unmeasured variables that might better explain the pattern observed in the present analyses. We hope future work will investigate these possibilities.

Future research should also test whether the relationship between religiosity and intelligence on violent (or other antisocial) behavior operates on a group level only or whether similar patterns would be observed on an individual-differences level or from experimental

manipulations of religiosity (Na et al., 2010). If the present results operate on a group level only, this might suggest that it is not intelligence per se that regulates violent behavior even in the absence of religion but rather that having a highly intelligent society contributes to highly functional group-level institutions and norms that help regulate behavior. In the Supplemental Material, we report exploratory analyses with two potential mechanisms, rule of law and democracy, but the interaction was robust to these controls as well. Identifying the most viable mechanism or mechanisms should be a crucial priority for future research.

Admittedly, although the observed pattern of results fit our hypothesis, we do not know exactly what it is about intelligence or religion that is associated with lower violent behavior. The mechanisms for intelligence and for religiosity might be similar (e.g., both might increase self-control), or they might be quite different (e.g., each might lead to different effective attitudes, norms, or institutions), but both appear to have some advantages for regulating violent behavior on a group level. The present analyses were not intended to reach final conclusions but rather to shine light on a potentially important and consequential relationship among these variables. We regard our research as a first step and welcome further input from other researchers.

The present work might inspire a bit of cautious reflection on the prescriptive values of Western, educated, industrialized, rich, and democratic (WEIRD) societies (Henrich, Heine, & Norenzayan, 2010). Educated societies might promote secularization without considering potentially disproportionately negative consequences for more cognitively disadvantaged groups. Some potential suppressors of violence (e.g., rule of law, trustworthy secular institutions, widespread concerns for fairness) may be more effectively implemented by populations with relatively high cognitive capacity (Kanyama, 2014), at least at the present moment. The benefits of religion may not be confined to homicide, and so there may be sweeping, multifaceted ways in which religion reduces violent, antisocial behavior, particularly among societies with relatively low average cognitive ability.

We suspect that similar patterns might emerge for numerous cultural narratives. The prescriptive values of highly educated groups (e.g., secularism but also libertarianism, criminal justice reform, and unrestricted sociosexuality, among others) may work for groups that are highly cognitively sophisticated and self-controlled, but they may be injurious to groups with lower self-control and cognitive ability. Highly educated societies with global esteem have more influence over global trends, and so the prescriptive values promulgated by these groups are likely to influence others who may not share their other cognitive characteristics. Perhaps,

then, highly educated and intelligent groups should be humble about promoting the unique and relatively novel values that thrive among them and perhaps should be cautious about mocking certain cultural narratives and norms that are perceived as having little value in their own society.

One-size-fits-all social prescriptions for complicated social problems may lack important nuance. And indeed some cultural institutions (e.g., religion, but also others such as monogamous marriage norms; Henrich, Boyd, & Richerson, 2012) that are denigrated as outmoded among high-IQ populations may still serve valuable functions among other groups around the world.

Transparency

Action Editor: Jamin Halberstadt

Editor: D. Stephen Lindsay

Author Contributions

C. J. Clark conceived the idea for these studies with B. M. Winegard. C. J. Clark and A. F. Shariff compiled the data. J. Beardslee analyzed all data and wrote the Results section for Study 1; C. J. Clark analyzed all data and wrote the Results section for Study 2. A. F. Shariff helped C. J. Clark solve the various methodological and statistical challenges of country-level analyses. C. J. Clark, B. M. Winegard, and A. F. Shariff prepared the manuscript. J. Beardslee and R. F. Baumeister provided critical revisions. All of the authors approved the final manuscript for submission.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Open Practices

All data and analysis code have been made publicly available via the Open Science Framework and can be accessed at <https://osf.io/ecdr/>. The design and analysis plans for a separate study that tested potential mechanisms for the effects we report in this article were preregistered at <https://aspredicted.org/sc9z2.pdf>. Results of this analysis can be found in the supplement at <https://osf.io/pe3jk/> (pp. 23–24). The complete Open Practices Disclosure for this article can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797619897915>. This article has received badges for Open Data, Open Materials, and Pre-registration. More information about the Open Practices badges can be found at <http://www.psychologicalscience.org/publications/badges>.



ORCID iD

Cory J. Clark  <https://orcid.org/0000-0002-3083-9179>

Acknowledgments

We thank David Becker for his helpful correspondence regarding the NIQ data set and the relative merits of different country-level IQ measures. We also thank our “data thug”

(auditor), Cindel White; a statistical consultant for Study 1, Paul Allison; three researchers who provided helpful correspondence regarding Galton's problem and spatial autocorrelation, Quentin Atkinson, Joe Henrich, and Ian Rickard; and our statistical expert on spatial autocorrelation, Scott Claeseens.

Supplemental Material

Additional supporting information can be found at <http://journals.sagepub.com/doi/suppl/10.1177/0956797619897915>

Notes

1. We also explored the inclusion of income inequality as an additional control variable (The World Bank, 2017b), but these data were extremely limited (only available beginning in 1981 and only for a limited number of countries), so we ultimately rejected them for Study 1. Income inequality was included in Study 2 to overcome this limitation.
2. We also cross-checked models using GDP and homicide estimates from the same individual years for which the 5-year average religiosity estimates were reported. In these analyses, all six interactions were statistically significant ($ps < .022$). These results are reported in Table S3 in the Supplemental Material.
3. When people describe an IQ score of 100 as average, this is based on the average IQ in the United Kingdom, which is above average relative to other countries.
4. In a second round of revisions, a reviewer suggested that we cross-check these analyses with data from school assessment studies only (i.e., without Lynn's data), so we reran our main analyses (first without controls, then with all four controls) with these data instead (school assessment studies from Becker, 2019). The interaction effect was very similar ($ns = 71$ – 98 countries, semipartial $r_s = .08$ – $.33$), although with the very limited number of countries, the interaction was not always statistically significant.
5. Data for 2014 were used because they were the most complete; closest available estimates were used if 2014 estimates were not available (and only if within 3 years of 2014).
6. Reviewers requested the square-root transformation instead of z transformation for GDP. This was honored in all models except those including the interactions between religiosity and GDP. For these models, we z -transformed GDP for purposes of computing the interaction term.
7. See the Supplemental Material for an initial (pre-peer-review) multiverse analysis, which included parasite stress and average annual temperature and did not include educational attainment and the additional interactions between control variables and religiosity. At the request of a reviewer, parasite stress and temperature were not included in the present multiverse analysis, and although these variables were positively correlated with higher homicide rates ($r_s \approx .33$) and negatively correlated with IQ ($r_s \approx -.68$), in the full model, they accounted for virtually zero variance in homicide rates (semipartial $r_s < .01$).
8. Homicide rates and ARDA and Gallup religiosity were skewed, so analyses were rerun omitting countries greater than 3 standard deviations above the homicide mean (Honduras, Venezuela, Belize) and countries greater than 3 standard deviations below the religiosity mean (Czech Republic, Estonia, South Korea, Japan). This did not affect the effect size or

statistical significance of the interaction with or without controls. In addition, to ensure that the results were not influenced by a lack of representation of certain combinations of religiosity and IQ (e.g., high religiosity and high IQ or low religiosity and low IQ), we performed median splits on religiosity and IQ and cross-checked the interactions in 2×2 analyses of variance. All nine interactions (three IQ measures by three religiosity measures) were statistically significant, $ps < .003$; medium to large effect sizes were found, η_p^2 s = $.064$ to $.156$. In the low-IQ country group, high-religiosity countries consistently had lower homicide rates than low-religiosity countries, $ps < .001$. In the high-IQ country group, there were no significant differences between high- and low-religiosity countries on homicide rates, $ps > .127$ (nor were the differences in the same direction).

References

- Allison, P. D. (2009). *Fixed effects regression models*. Thousand Oaks, CA: SAGE.
- Baier, C. J., & Wright, B. R. (2001). "If you love me, keep my commandments": A meta-analysis of the effect of religion on crime. *Journal of Research in Crime and Delinquency*, *38*, 3–21.
- Becker, D. (2019). *The NIQ dataset (Version 1.3.2)*. Retrieved from http://viewoniq.org/?page_id=9
- Becker, D., & Rindermann, H. (2016). The relationship between cross-national genetic distances and IQ-differences. *Personality and Individual Differences*, *98*, 300–310.
- Bierman, W., & van Zanden, J. L. (2014). *Homicide rates*. Retrieved from <https://clio-infra.eu/Indicators/HomicideRates.html#>
- Boisvert, D., Stadler, W., Vaske, J., Wright, J. P., & Nelson, M. (2013). The interconnection between intellectual achievement and self-control. *Criminal Justice and Behavior*, *40*, 80–94.
- Boutwell, B. B., Barnes, J. C., Beaver, K. M., Haynes, R. D., Nedelec, J. L., & Gibson, C. L. (2015). A unified crime theory: The evolutionary taxonomy. *Aggression and Violent Behavior*, *25*, 343–353.
- Central Intelligence Agency. (2015). *The world factbook*. Retrieved from <https://www.cia.gov/library/publications/the-world-factbook/rankorder/2172rank.html>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.
- Crabtree, S. (2010). *Religiosity highest in world's poorest nations*. Retrieved from <https://news.gallup.com/poll/142727/Religiosity-Highest-World-Poorest-Nations.aspx#1>
- Fukuyama, F. (1995). *Trust: Social virtues and the creation of prosperity*. London, England: Hamish Hamilton.
- Henrich, J., Boyd, R., & Richerson, P. J. (2012). The puzzle of monogamous marriage. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 657–669.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*, 61–83.
- Hruschka, D. J., & Henrich, J. (2013). Institutions, parasites and the persistence of in-group preferences. *PLOS ONE*, *8*(5), Article e63642. doi:10.1371/journal.pone.0063642

- Hunt, E. (2011). *Human intelligence*. New York, NY: Cambridge University Press.
- Inglehart, R., & Norris, P. (2003). *Sacred and secular*. New York, NY: Cambridge University Press.
- Kanyama, I. K. (2014). Quality of institutions: Does intelligence matter? *Intelligence*, *42*, 44–52.
- Lynn, R., & Meisenberg, G. (2010). National IQs calculated and validated for 108 nations. *Intelligence*, *38*, 353–360.
- Lynn, R., & Vanhanen, T. (2002). *IQ and the wealth of nations*. Westport, CT: Praeger.
- Lynn, R., & Vanhanen, T. (2006). *IQ and global inequality*. Augusta, GA: Washington Summit Publishers.
- Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences*. London, England: Ulster Institute for Social Research.
- Maoz, Z., & Henderson, E. A. (2013). The World Religion Dataset, 1945–2010: Logic, estimates, and trends. *International Interactions*, *39*, 265–291.
- McCullough, M. E., & Carter, E. C. (2013). Religion, self-control, and self-regulation: How and why are they related? In K. I. Pargament (Ed.), *APA handbook of psychology, religion, and spirituality: Vol. 1. Context, theory, and research* (pp. 123–138). Washington, DC: American Psychological Association.
- McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., . . . Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences, USA*, *108*, 2693–2698.
- Na, J., Grossmann, I., Varnum, M. E., Kitayama, S., Gonzalez, R., & Nisbett, R. E. (2010). Cultural differences are not always reducible to individual differences. *Proceedings of the National Academy of Sciences, USA*, *107*, 6192–6197.
- Neopolitan, J. L. (1996). Cross-national crime data: Some unaddressed problems. *Journal of Criminal Justice*, *19*, 95–112.
- Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences*, *39*, Article e1. doi:10.1017/S0140525X14001356
- Pew Research Center. (2012). *Religious composition by country, in percentages*. Retrieved from <http://www.pewforum.org/2012/12/18/table-religious-composition-by-country-in-percentages/>
- Piazza, J., & Sousa, P. (2014). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science*, *5*, 334–342.
- Pietschnig, J., & Voracek, M. (2015). One century of global IQ gains: A formal meta-analysis of the Flynn effect (1909–2013). *Perspectives on Psychological Science*, *10*, 282–306.
- Purzycki, B. G., Apicella, C., Atkinson, Q. D., Cohen, E., McNamara, R. A., Willard, A. K., . . . Henrich, J. (2016). Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature*, *530*, 327–330.
- R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rindermann, H., & Ceci, S. J. (2009). Educational policy and country outcomes in international cognitive competence studies. *Perspectives on Psychological Science*, *4*, 551–568.
- Rindermann, H., & Thompson, J. (2011). Cognitive capitalism: The effect of cognitive ability on wealth, as mediated through scientific achievement and economic freedom. *Psychological Science*, *22*, 754–763.
- Shariff, A. F. (2015). Does religion increase moral behavior? *Current Opinion in Psychology*, *6*, 108–113.
- StataCorp. (2015). *Stata Statistical Software: Version 14.0* [Computer software]. College Station, TX: Author.
- Steege, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*, 702–712.
- United Nations Office on Drugs and Crime. (2013). *Homicide counts and rates, time series 2000–2012*. Retrieved from <http://www.unodc.org/gsh/en/data.html>
- The World Bank. (2015). *Population density (people per sq. km of land area)*. Retrieved from <https://data.worldbank.org/indicator/EN.POP.DNST>
- The World Bank. (2017a). *World development indicators: GDP (per capita)*. Retrieved from <http://databank.worldbank.org/data/reports.aspx?source=2&series=NY.GDP.PCAP.CD&country=#>
- The World Bank. (2017b). *World development indicators: GINI index*. Retrieved from <http://databank.worldbank.org/data/reports.aspx?source=2&series=SI.POV.GINI&country=#>
- The World Bank. (2019). *World development indicators: Lower secondary completion rate*. Retrieved from <https://data.worldbank.org/indicator/SE.SEC.CMPT.LO.ZS>
- Zuckerman, M., Silberman, J., & Hall, J. A. (2013). The relation between intelligence and religiosity: A meta-analysis and some proposed explanations. *Personality and Social Psychology Review*, *17*, 325–354.