***Comment on “Civic honesty around the globe” – no detectable differences
between countries***

Rongfeng Cui^{1*}, Yanhong Hou², Pingze Zhang¹, Wenjie Yao^{3,4,5}, Xinxin Du^{6,7} and Yongwei Guo²

¹Max Planck Institute for Biology of Aging, Cologne, Germany 50931

²Department of Ophthalmology, University of Cologne, Kerpener Str. 62, Cologne Germany 50937

³Institute for Genetics, Centre for Molecular Medicine (CMMC), University of Cologne, Cologne, Germany 50931

⁴Cologne Excellence Cluster on Cellular Stress Responses in Aging-Associated Diseases (CECAD), University of Cologne, Cologne, Germany 50931

⁵Institute for Neurophysiology, Faculty of Medicine and University Hospital, University of Cologne, Cologne Germany 50931

⁶Department of Gynecology, University of Cologne, Cologne, Germany 50931

⁷Department of Obstetrics and Gynecology, PLA General Hospital, Beijing, China 100853

*Correspondence to: rcui@age.mpg.de

Abstract After correcting for experimental control variables, we found that all country-level correlations originally reported in Cohn et al. vanished. No conclusions can be drawn on cross-country differences in the rates of reporting lost wallets.

Field behavioral studies in humans have been complicated by difficulties in controlling for unobserved variables (1). Although the main aim was to quantify the effect of money incentives on wallet return rate, Cohn et al. (2) also reported a drastic variation in email responding rate among countries, and interpreted them as “objective measurements” of civic honesty. They then claimed that various geographic, cultural, and institutional variables of different countries (2) can explain 65% of the variation in the adjusted email responding rate (their Fig S6).

Principal component (PC) analysis of all independent variables revealed that many situational control variables and candidate explanatory country-level variables both strongly loaded onto PC1 (Fig. S1), suggesting that these two types of variables are confounded. PC1 alone explains 70.65% of the variation in email response rate. Statistical significance cannot be directly interpreted if multicollinear variables are simultaneously included in a linear model (3), an approach adopted by Cohn et al. (2) when attempting to remove the influence of control variables. A proper method to remove the effects of the confounding variables is by taking the residuals from a linear regression (3), as Cohn et al. also did with email usage controls (2). After removing the effects of 19 control variables from the country-level variables by linear regression, none of their residuals significantly correlate with response rate after false discovery rate correction (Fig. S2). We present a corrected Fig. 1 with email response residuals, with the emphasis that the country rankings were unexplained by statistically meaningful factors.

Other caveats that we have not accounted for here includes pseudoreplication of countries of shared cultural history, akin to the non-independency problem in phylogenetic comparisons (4), which results in over-confidence of significant tests, if any.

Our analyses confirm that money incentives increase email response rate, but cross-country differences were mainly driven by spurious correlations unrelated to civic honesty.

References

1. J. Henrich, S. J. Heine, A. J. N. Norenzayan, Most people are not WEIRD. *Nature* **466**, 29 (2010).
2. A. Cohn, M. A. Maréchal, D. Tannenbaum, C. L. Zünd, Civic honesty around the globe. *Science*, eaau8712 (2019).
3. M. H. Graham, Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809-2815 (2003).
4. J. Felsenstein, Phylogenies and the comparative method. **125**, (1985).

Figures

Fig. 1. Residuals of email response rates of 40 countries, correcting for 19 variables related to experimental conditions (Fig. S1). Yellow dot – NoMoney treatment. Red dot – Money treatment. Variations between countries were not explained by any examined country-level variables (Fig. S2).

Supplementary Information

Principal component analysis

Due to the very high number of independent variables (including control variables and candidate country-level variables) in the dataset, it is necessary to remove their self-correlations prior to subsequent analysis, to avoid problems caused by multicollinearity during linear regressions (1). Due to extensive missing data, the variable “r_PL_propertylaw” was excluded from the matrix. The following variables contained sporadic missing data: “security_cam”, “security_guard”, “c_WVS_familyties”, “c_GPLS_judicial_independence”, “c_GPLS_constitutional_review”, “c_GPLS_plurality” and “c_GPLS_prop_representation”. In order to maximize usable data, we imputed the missing values with the missingMDA R package, using all other independent variables (but not response rate) as predictors.

The optimal number of PCs were then estimated on the matrix with missing data using the *estim_ncpPCA* R function using the kfold approach, with a maximal number of PCs set to 30. This procedure identified 27 PCs to be sufficient to describe all independent variables. The PCA was performed using the data matrix with imputed cells, with the component number set to 27. We then used a model selection approach to identify which PCs significantly correlated with email response rates.

For model selection, we use a stepwise variable addition/elimination approach, implemented in the R *step()* function, using the Akaike information criterion (2) (AIC) during each iteration to determine the most parsimonious model that yields the lowest AIC. The addition/elimination steps were run in both the direction of increasing and decreasing model complexity (*direction* = “both”), starting from the simplest model including only the intercept. Only main effects were examined. We note that the model choice was robust to the choice of the starting model and the direction of iteration.

This procedure yielded an optimal model containing 14 out of the 27 PCs, explaining 90.73% of the variation in email response rate (10 of 14 are significant at $p < 0.05$). Among these, PC1 explains 70.65%, PC7 explains 7.87% while PC12 explains 5.64%. The rest of the PCs only explain a small proportion of the variation. We visualize the variable loadings onto each PC, and found that money condition uniquely loaded onto PC12, suggesting that this variable is unlikely to be confounded by other effects (Fig. S1). However, many variables, including control and country-level variables, heavily loaded onto PC1, suggesting that their correlations were confounded by each other (Fig. S1).

We next removed the influence of situational and experimental variables (control variables, bolded in Fig. S1) from the candidate country-level variables. First, we performed PCA on all control variables and correlated all PCs with email response rate. Significant PCs ($p < 0.05$) were used to perform linear regressions with each of the country-level variables. By taking the residuals from these regressions, we effectively remove any spurious correlations caused by control variables caused by idiosyncratic experimental conditions. None of the residuals correlated with email response rate after FDR correction (Fig. S2) any longer.

We produced a new Fig. 1 showing the residual email response rates corrected for all control variables, using the same methods outlined above. A prominent feature of this figure is a random distribution of developed and developing nations and a narrower spread of data.

Supplementary Figures

Fig S1. Visualization of principal component loadings of all independent variables, including experimental control (in bold) and candidate country-level explanatory variables (in normal font). The signs of the loadings were calculated such that it is in the direction of predicting higher email response rates. The relative loadings were multiplied by the percentage of variation explained by the PC shown on row 1. PC1 explains the majority of the variation, and receives loading from both control variables and candidate country-level variables. This suggests that control and country-level variables are correlated among themselves (multicollinear).

Fig. S2. After removing the correlation with all potential confounding experimental variables (including "money_condition", "computer", "no_english", "security_cam", "cultural", "busy", "understood_situation", "rejection_frac", "hotel", "other_bystanders", "postal", "public", "age", "maleexperimenter", "bank", "age_experimenter", "male", "security_guard" and "coworkers"), none of the other variables were significantly correlated with email response ($FDR > 0.2$). Bar plot shows the correlation coefficients, error bars are standard errors of the means.

References

1. M. H. Graham, Confronting multicollinearity in ecological multiple regression. *Ecology* **84**, 2809-2815 (2003).
2. T. Yamashita, K. Yamashita, R. J. C. i. S. T. Kamimura, Methods, A stepwise AIC method for variable selection in linear regression. *Communications in Statistics—Theory and Methods* **36**, 2395-2403 (2007).