Supplemental Material

for

Mede, N. G., Schäfer, M. S., Ziegler, R., & Weißkopf, M. (2020). The "replication crisis" in the public eye: Germans' awareness and perceptions of the (ir)reproducibility of scientific research. *Public Understanding of Science*. Available at: https://doi.org/10.1177/0963662520954370.

Corresponding Author

Niels G. Mede, University of Zurich, Department of Communication and Media Research (IKMZ), Andreasstrasse 15, 8050 Zürich, Switzerland / +41 (0) 44 635 20 64, n.mede@ikmz.uzh.ch

Introduction

In this document we provide tables and figures supplementing the analyses presented in our research note "The 'replication crisis' in the public eye: Germans' awareness and perceptions of the (ir)reproducibility of scientific research", published in *Public Understanding of Science* (available at: https://doi.org/10.1177/0963662520954370). Section 1 contains an overview of the variables we used in the analyses. Section 2 contains results of additional analyses we ran to test assumptions of the regression models shown in Tables 1 and 2 in the original article.

List of Tables and Figures

Section 1

Section 2

Section 1

Supplemental Table A1

Overview of variables used in the analyses

Variable	Question wording	Levels in analyses
Awareness of the "replication crisis"	Part of the scientific research process is also to replicate the work of other scientists. Therefore, studies are undertaken in the same way again to see if they deliver equivalent results.	
Awareness of "replication crisis" in biology[a]	It has recently been reported that within biology a certain number of replication studies did not deliver the same results. Have you heard or read about it in the media lately?	1 = yes, 0 = no, NA = don't know/not answered
Awareness of "replication crisis" in economics[a]	It has recently been reported that within economics a certain number of replication studies did not deliver the same results. Have you heard or read about it in the media lately?	1 = yes, 0 = no, NA = don't know/not answered
Awareness of "replication crisis" in medicine[a]	It has recently been reported that within medicine a certain number of replication studies did not deliver the same results. Have you heard or read about it in the media lately?	1 = yes, 0 = no, NA = don't know/not answered
Awareness of "replication crisis" in psychology[a]	It has recently been reported that within psychology a certain number of replication studies did not deliver the same results. Have you heard or read about it in the media lately?	1 = yes, 0 = no, NA = don't know/not answered
Awareness of "replication crisis" in any of the four disciplines[b]	–	1 = yes, 0 = no, NA = don't know/not answered
Perceptions of the "replication crisis"	Whether you have read or heard about it or not, how do you assess the situation? Considering that scientific results could not be confirmed in replication studies, please tell me to what extent you agree with the following statements.	
Reason to distrust	It shows that you cannot trust science and research in this field.	1 = completely disagree, 2 = somewhat disagree, 3 = undecided, 4 = somewhat agree, 5 = completely agree, NA = don't know/not answered
Shows quality control	Replication shows that quality assurance takes place in science and research.	1 = completely disagree, 2 = somewhat disagree, 3 = undecided, 4 = somewhat agree, 5 = completely agree, NA = don't know/not answered
Indicates self-correction	Errors and their corrections are part of science and research.	1 = completely disagree, 2 = somewhat disagree, 3 = undecided, 4 = somewhat agree, 5 = completely agree, NA = don't know/not answered

Predictors

Age	How old are you?	[Measured in years]
Gender	[Determined by interviewer]	1 = female, 0 = male
Net household income per month	If you add it all up, what is the net monthly income that you all have in the household after paying taxes and social security premiums? I will now read out some groups and you please tell me which of them applies.	1 = less than 1,000 Euros, 2 = 1,000 to less than 2,000 Euros, 3 = 2,000 to less than 3,000 Euros, 4 = 3,000 to less than 4,000 Euros, 5 = 4,000 to less than 5,000 Euros, 6 = more than 5,000 Euros
Education[b]	Do you have a degree from a university or a university of applied sciences?	1 = university degree, 0 = no university degree
Region of residence place[b]	–	1 = East Germany, 0 = West Germany
Party preference[b]	Which political party would you vote for if national elections would be held next Sunday?	1 = preference for AfD, 0 = preference for other party/no preference/don't know, NA = not answered
Religiosity	To what extent would you describe yourself as religious?	1 = not at all religious, 2 = somewhat not religious, 3 = so-so, 4 = somewhat religious, 5 = very religious, NA = don't know/not answered
Interest in science	Please tell me how strong your interest is in science and research.	1 = very low, 2 = somewhat low, 3 = so-so, 4 = somewhat strong, 5 = very strong, NA = don't know/not answered
Frequency contact with science[b]	[Mean index of six items asking how frequently one comes in contact with science via (1) conversations, (2) events, (3) print media, (4) TV, (5) radio, (6) the Internet]	[Mean index ranges from 1 to 5, higher values indicate higher frequency (Cronbach's Alpha = .72). Levels of items: 1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = very often, NA = don't know/not answered]
Proximity to science[b]	[Score composed from two variables asking for respondents' personal involvement with science: (1) currently or previously working in science, (2) personal acquaintance with scientists]	[Score ranges from 1 to 3, higher values indicate greater proximity: 1 = no acquaintance with scientists, 2 = personal acquaintance with scientists, 3 = currently/previously working in science]
Split-ballot group assignment: biology[b]	–	1 = true, 0 = false
Split-ballot group assignment: economics[b]	–	1 = true, 0 = false
Split-ballot group assignment: medicine[b]	–	1 = true, 0 = false
Split-ballot group assignment: psychology[b]	–	1 = true, 0 = false

Note. Original items were in German (Wissenschaft im Dialog, 2019). English translations were carried out by a professional translator from the Wissenschaftsbarometer team (Wissenschaft im Dialog, 2018).

[a] Each of the four variables marked with this superscript was measured for a different quarter of randomly selected respondents (split-ballot procedure).

[b] Variable was constructed for the analyses.

Section 2

Assumption Tests for Binary Logistic Regression Analysis on Weighted Data (Table 1)

We checked assumptions of the binary logistic regression analysis we used to investigate predictors of awareness of replication failures (see Table 1 in the original article). In particular, we tested for multicollinearity and non-linearity of the logit of continuous predictors (Field, Miles, & Field, 2012, pp. 342–345).

Supplemental Table A2

Assumption checks for binary logistic regression analysis on weighted data (see Table 1)

	Test for multicollinearity	Test for non-linearity of the logit	
Predictor	*VIF*	*t*	*p*
(Intercept)	—	-1.57	0.12
Age	1.43	0.14	0.89
Gender (1 = female)	1.20	-1.45	0.15
Household income	1.71	1.61	0.11
Education (1 = university degree)	1.48	2.06	0.04
Region (1 = East Germany)	1.31	0.62	0.54
Party preference (1 = AfD)	1.26	0.40	0.69
Religiosity	1.20	-0.43	0.67
Interest in science	1.73	1.61	0.11
Frequency contact with science	2.01	-0.59	0.56
Proximity to science	1.64	0.27	0.79
Split-ballot group assignment (ref.: replication failures in biology)			
Replication failures in economics	1.74	-1.42	0.16
Replication failures in medicine	1.68	-1.11	0.27
Replication failures in psychology	2.12	-2.47	0.01
Age × log(Age)	—	0.02	0.99
Household income × log(Household income)	—	-1.71	0.09
Religiosity × log(Religiosity)	—	0.34	0.74
Interest in science × log(Interest in science)	—	-1.52	0.13
Frequency contact with science × log(Frequency contact with science)	—	0.84	0.40
Proximity to science × log(Proximity to science)	—	-0.21	0.83

To test for multicollinearity, we computed variance inflation factors (Marquardt, 1970), employing the function `summ` from the package `jtools` v2.1.0 (Long, 2020). Variance inflation factors of predictors in the regression model were even below conservative cutoff values such as 5.00 (Menard, 1995, p. 66; see Supplemental Table A2). Therefore, we did not assume severe multicollinearity and kept all predictors in the model.

To test for non-linearity of the logit of continuous predictors, we applied the Box-Tidwell approach suggested by Hosmer and Lemeshow (1989). It involves including additional predictors in the model, which are interaction terms of each continuous predictor and its natural logarithm (Box & Tidwell, 1962). If any of these terms reaches significance, the assumption of linearity of the logit of the respective predictor would be violated (Tabachnick & Fidell, 2005). We did not find significant interaction terms, so we kept all predictors in our regression model (see Supplemental Table A2).

Assumption Tests for Multiple Linear Regression Analyses on Weighted Data (Table 2)

We checked assumptions of the multiple linear regression analyses we used to investigate predictors of perceptions of replication failures and replication efforts (see Table 2 in the original article). In particular, we tested for multicollinearity as well as for non-normality and heteroscedasticity of the residuals (Field et al., 2012, pp. 292–297).

To test for multicollinearity, we relied on VIFs once more. VIFs of predictors in the regression models were below conservative cutoff values such as 5.00 (see Supplemental Table A3). Therefore, we did not assume severe multicollinearity and kept all predictors in the models.

To test for non-normality of the residuals, we relied on histograms of studentized residuals and on Q-Q plots (Field et al., 2012, pp. 294–297). These histograms and plots suggested that the assumption of normality was met for the regression predicting respondents’ agreement that

non-replicability shows that one cannot trust science (see Supplemental Figures A1 and A2). However, residuals of the regression predicting agreement that replication efforts demonstrate quality assurance in science did not seem to be normally distributed. The histogram and the Q-Q plot indicated unexpectedly many medium positive (studentized) residuals and many large negative (studentized) residuals (see Supplemental Figures A4 and A5). The normality assumption was also violated for the regression predicting agreement that errors and their corrections are part of science. Here we found a clearly skewed distribution (see Supplemental Figures A7 and A8).

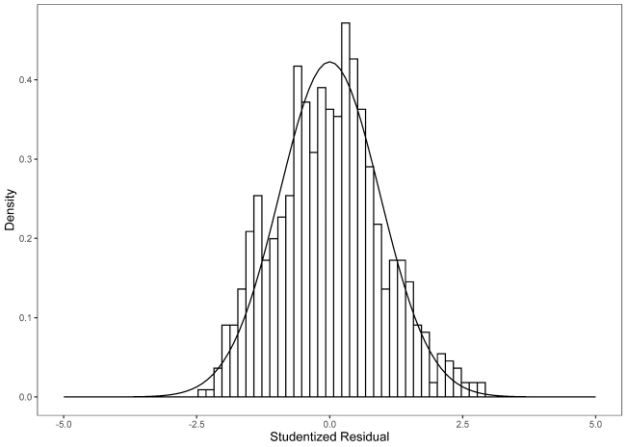
Supplemental Table A3

Multicollinearity tests for multiple linear regression analyses on weighted data (see Table 2)

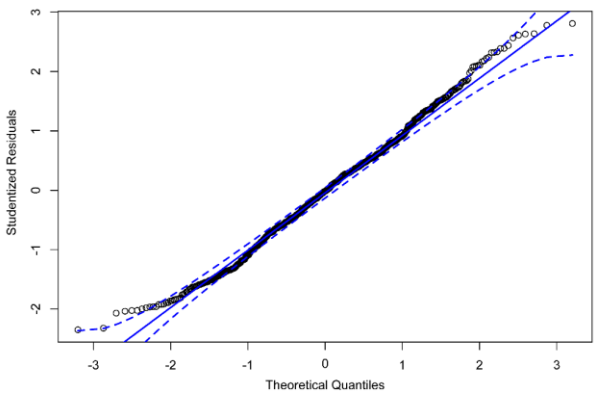
	“[Non-replicability] shows that you cannot trust science and research.”	“Replication shows that quality assurance takes place in science and research.”	“Errors and their corrections are part of science and research.”
Predictor	*VIF*	*VIF*	*VIF*
(Intercept)	—	—	—
Age	1.17	1.42	1.35
Gender (1 = female)	1.24	1.27	1.26
Household income	1.34	1.43	1.49
Education (1 = university degree)	1.58	1.73	1.77
Region (1 = East Germany)	1.34	2.14	1.61
Party preference (1 = AfD)	1.30	1.14	1.42
Religiosity	1.21	1.23	1.34
Interest in science	2.14	2.18	1.55
Frequency contact with science	1.88	3.27	1.77
Proximity to science	1.83	1.82	1.98
Awareness of “replication crisis”	1.34	1.56	1.11
Split-ballot group assignment (ref.: replication failures in biology)			
Replication failures in economics	1.94	1.63	1.66
Replication failures in medicine	1.74	2.02	1.71
Replication failures in psychology	1.79	1.70	1.67

To test for heteroskedasticity, we relied on plots of studentized residuals against predicted values (Field et al., 2012, pp. 294–295). When inspecting these plots, we did not find clear signs of increasing or decreasing variance across residuals for any of the three models (see Supplemental Figures A3, A6, and A9). Hence, we assumed that the assumption of homoscedasticity was met for each model.

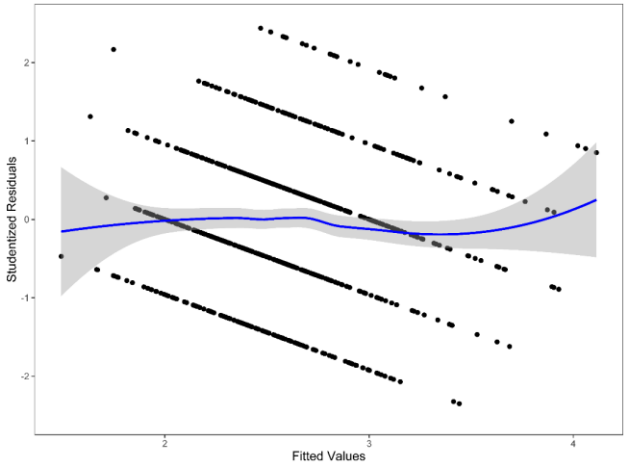
Because we could not assume normality of residuals for the regression analyses predicting agreement that replication efforts demonstrate quality assurance in science and that errors and their corrections are part of science, we aimed to run regression analyses with robust standard error estimation. Yet we are not aware of an R package that can handle such estimation in regression analyses employing survey weights such as those we computed using the package `survey` v4.0 (Lumley, 2020). Therefore, we decided to bootstrap the standard errors and confidence interval bounds of the regression coefficients (Field et al., 2012, pp. 298–301). After all, bootstrapping is related to robust standard error estimation anyways (Pek, Wong, & Wong, 2018, pp. 6–7). For bootstrapping we employed the package `boot` v1.3-25 (Ripley, 2020), using non-studentized pivotal confidence interval estimation (Carpenter & Bithell, 2000).



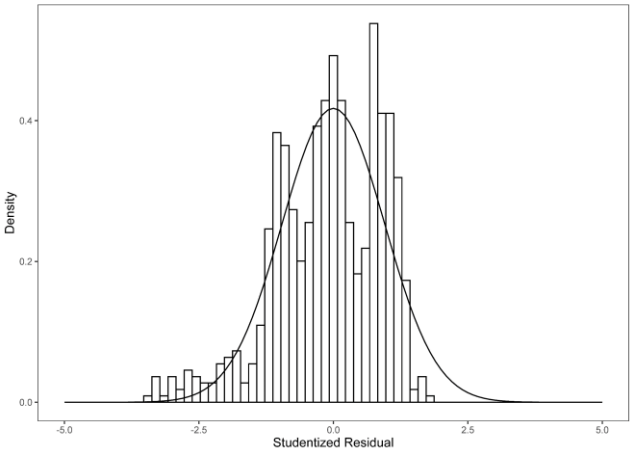
Supplemental Figure A1. *Histogram of studentized residuals of regression model for “[Non-replicability] shows that you cannot trust science and research.”*



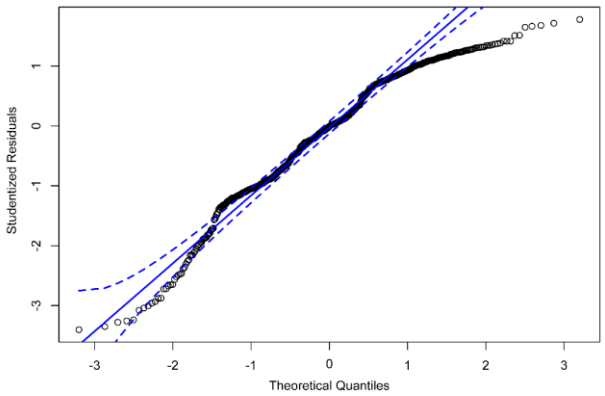
Supplemental Figure A2. *Q-Q plot of studentized residuals of regression model for “[Non-replicability] shows that you cannot trust science and research.”*



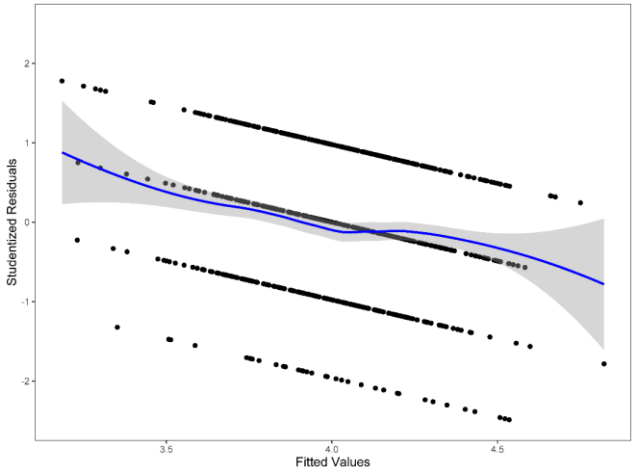
Supplemental Figure A3. *Plot of studentized residuals against predicted values of regression model for “[Non-replicability] shows that you cannot trust science and research.”*



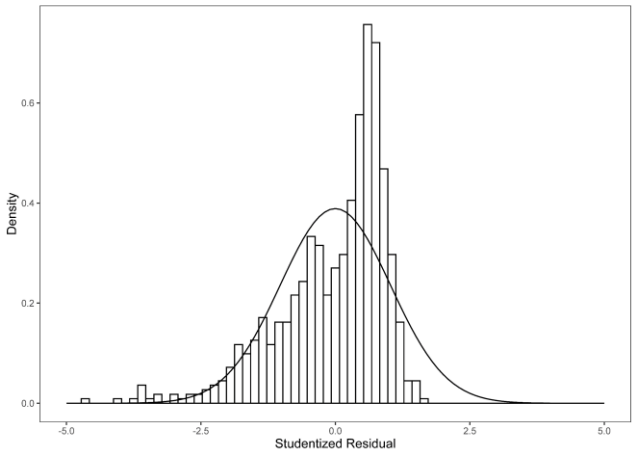
Supplemental Figure A4. *Histogram of studentized residuals of regression model for “Replication shows that quality assurance takes place in science and research.”*



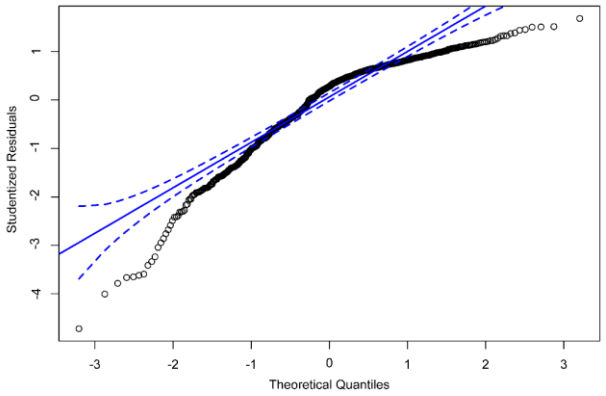
Supplemental Figure A5. *Q-Q plot of studentized residuals of regression model for “Replication shows that quality assurance takes place in science and research.”*



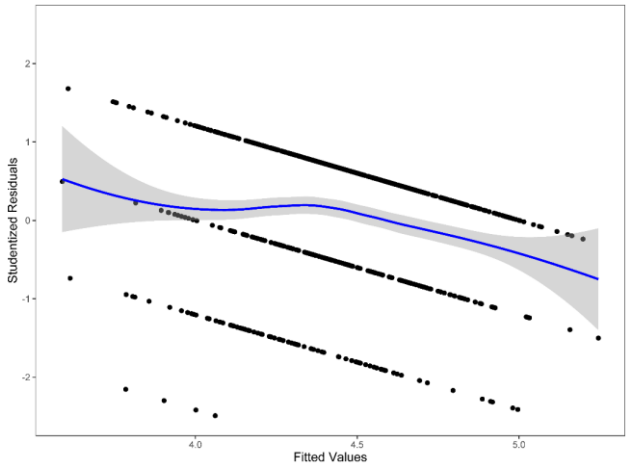
Supplemental Figure A6. *Plot of studentized residuals against predicted values of regression model for “Replication shows that quality assurance takes place in science and research.”*



Supplemental Figure A7. *Histogram of studentized residuals of regression model for “Errors and their corrections are part of science and research.”*



Supplemental Figure A8. *Q-Q plot of studentized residuals of regression model for “Errors and their corrections are part of science and research.”*



Supplemental Figure A9. *Plot of studentized residuals against predicted values of regression model for “Errors and their corrections are part of science and research.”*

References

Box, G. E. P., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, *4*(4), 531. https://doi.org/10.2307/1266288

Carpenter, J., & Bithell, J. (2000). Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*, *19*(9), 1141–1164. https://doi.org/10.1002/(SICI)1097-0258(20000515)19:9<1141::AID-SIM479>3.0.CO;2-F

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, CA: Sage.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York, NY: Wiley.

Long, J. A. (2020). jtools: Analysis and presentation of social scientific data. R package version 2.1.0. Retrieved from https://cran.r-project.org/web/packages/jtools

Lumley, T. (2020). survey: Analysis of complex survey samples. R package version 4.0. Retrieved from https://cran.r-project.org/web/packages/survey

Marquardt, D. W. (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, *12*(3), 591–612. https://doi.org/10.1080/00401706.1970.10488699

Menard, S. (1995). *Applied logistic regression analysis: Sage university series on quantitative applications in the social sciences*. Thousand Oaks, CA: Sage.

Pek, J., Wong, O., & Wong, A. C. M. (2018). How to address non-normality: A taxonomy of approaches, reviewed, and illustrated. *Frontiers in Psychology*, *9.* https://doi.org/10.3389/fpsyg.2018.02104

Ripley, B. (2020). boot: Bootstrap functions. R package version 1.3-25. Retrieved from https://cran.r-project.org/web/packages/boot

Tabachnick, B. G., & Fidell, L. S. (2005). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn & Bacon.

Wissenschaft im Dialog (2018). Questionnaire – Science Barometer 2018. Retrieved from www.wissenschaft-im-dialog.de/fileadmin/user_upload/Projekte/Wissenschaftsbarometer/Dokumente_18/Downloads_allgemein/Sciencebarometer2018_Questionnaire.pdf

Wissenschaft im Dialog (2019). ZA6965: Science Barometer 2018: Data file version 1.0.0. GESIS Data Archive, Cologne. Retrieved from https://doi.org/10.4232/1.13241