Eva van Reijmersdal & Amber van der Wal,
Faculty of Social and Behavioural Sciences
Course: Research & Data Ethics

UNIVERSITEIT VAN AMSTERDAM

# Individual Essay:

## A critical reflection on the social sciences

**Helge Moes**
11348801
11348801@uva.nl

Date: 02.06.2023
Word count: 2310

# Table of Content

# Introduction

At the beginning of the COVID-19 pandemic, governments were held accountable for preventing the spread of the virus. In order to do so, multiple measures were conducted for the public's well being. Yet these measures were met with a degree of doubt and mistrust by the public (Vliegenthart et al., 2021). In order to understand and investigate this phenomenon, Bakker, van der Wal and Vliegenthart (2020) have investigated public opinions and sentiments regarding the coronavirus in the Netherlands. This was done by asking participants a series of questions about their views and feelings towards the coronavirus. The aim of this research was to gain a better understanding of the behavior and attitudes of Dutch individuals towards the coronavirus.

Nevertheless, these behaviors are challenging to research and may hold biases when examined (Franco et al., 2014). This is a common deficit of research in social sciences. In order to evaluate research in social sciences, studies have to be transparent, robust, valid, representative and be able to communicate uncertainty without compromising research ethics (van der Bles et al., 2020). Thus, for this essay an analysis shall be conducted on the public's sentiment on COVID-19 to explore the freedom researchers have in conducting research, what the ramifications are of this and the need for transparency in research.

In recent years, social sciences has been experiencing a publication bias (Franco et al., 2014). In the academic field, articles that produce null results face greater challenges on being published than those that demonstrate statistically significant findings (Franco et al., 2014). Moreover, this phenomenon implies that it is not important whether findings are biased or incorrect, rather whether something is statistically significant or not. This drawback within academic research in the social sciences is more often than necessary accompanied with questionable research practices (QRP) that may affect results of studies (Bakker et al., 2021). These QRPs increase the chance for discovering a finding to support a false hypothesis and makes it impossible to replicate. Therefore, it is fundamental for scientists to produce robust and credible outputs while also being transparent on these measures (van der Bles et al., 2020). If QRPs are indeed widespread in communication research, it would have significant implications for the reputation of the social sciences field within both external perceptions and among the scholars within the field (Bakker et al., 2021).

### Problem definition

An examination conducted by researchers (Özdil et al., 2021) explored the connection between personal compliance to COVID-19 guidelines and feelings of apprehension and fatalism due to fear. The findings indicated that fear can strongly influence one's stance towards the regulations. Similarly, education levels also played a significant role, with higher education correlating to greater compliance, although the impact was comparatively smaller than that of fear (Özdil et al., 2021). Hence, within the context of examining research ethics, the hypothesis under investigation in this study posits that fear holds a greater influence on compliance predisposition than education.

The primary objective of this study is to highlight the importance of enhanced transparency in social science research by illustrating how diverse analytical approaches employed within the same study can yield different findings, potentially distorting our comprehension of a social phenomena. To accomplish this, the essay examines the case of exploring the public's reactions and attitudes towards COVID-19 regulations in the Netherlands presented in the OSF data from the study that was conducted by Bakker et al. (2020). A two-tailed correlation and a linear regression shall be executed in order to highlight the different types of statistical models and their results.

## Empirical Results

### Variables

The research shall contain an exploratory nature that attempts to provide evidence to portray how sensitive data to different analytical methods. Consequently, a two-tailed correlation and a linear regression model were run with multiple randomly selected variables in order to find significant variables with the dependent variable compliance. Based on this examination, ‘opleidingsniveau’ was determined to be significant. Thus, these two strategic methods shall be compared in order to assess the compliance of people based on their fear.

The variable ‘fear’ was treated as a continuous variable on a scale from 1 to 7 and was created by merging all the fear variables together. The variable ‘opleidingsniveau’ was presented on three different levels: laag (1), midden (2), hoog (3). In order to conduct a linear regression, dummy variables were created for ‘opleidingsniveau’, which consisted of ‘opleidingniveau2’ (midden, middle educated) and ‘opleidingniveau3’ (hoog, high educated). In order to prevent collinearity, a dummy variable was not made for laag (low educated), since it is already implemented in the constant.

### Model 1: Two-Tailed Correlation

The Pearson's correlation coefficient is typically used to determine if there is a linear association between two variables and to what degree they are correlated. In this case, ‘opleidingsniveau’ and ‘fear’ of COVID-19 are examined to see whether they correlate. If the findings are significant, then one can conclude that the level of education has an effect on whether or not an individual complies based on fear of the virus.

The correlation output indicates the relationship between the fear of COVID-19 and educational attainment that affect the rate of compliance. The Pearson correlation coefficient for the variables ‘fear’ and ‘opleidingsniveau’ is $r = -0.084$. The negative finding indicates a weak negative correlation, implying that as the fear of contracting COVID-19 increases, people with a lower education level tend to comply less to COVID-19 regulations. However, the correlation is quite small.

The significance level associated with the correlation coefficient is $p < 0.001$ with a sample size of 1429, which suggests that this correlation is statistically significant. This means that the observed correlation is unlikely to have occurred by chance. Hence, this correlation output suggests that there is a small but statistically significant negative relationship between the fear of contracting COVID-19 and education level on compliance.

### Model 2: Linear Regression

The linear regression model examining the relationship between the predictor variable ‘opleidingsniveau’ and the dependent variable ‘fear’ yielded a weak positive association ($R = .178$, $R^2 = .032$, adjusted $R^2 = .031$). In addition, the analysis of variance indicated a significant effect of the regression model ($F(1, 1739) = 56.908$, $p < .001$). The regression model accounted for 3.2% of the variance in the dependent variable, as indicated by the regression sum of squares ($SS = 172.490$) compared to the residual sum of squares ($SS = 5270.978$). Finally, the predictor variable ‘opleidingsniveau’ exhibited a significant negative standardized coefficient ($\beta = -.178$, $t(1739) = -7.544$, $p < .001$), indicating that higher levels of ‘opleidingsniveau’ were associated with decreased levels of fear regarding the coronavirus ($B = -.414$, $SE = .055$). Furthermore, the 95% confidence interval for the coefficient ranged from -0.521 to -0.306 and the intercept was 3.318 ($SE = .125$) with a confidence interval ranging from 3.072 to 3.564.

The dummy variable ‘opleidingsniveau2’ also revealed a very weak positive association ($R = .050$, $R^2 = .002$, adjusted $R^2 = .002$). Additionally, the analysis of variance

indicated a marginally significant effect of the regression model ($F(1, 1739) = 4.293$, $p = .038$). This regression model only accounted for 0.2% of the variance in the dependent variable of ‘fear’, as shown by the regression sum of squares (SS = 13.404) compared to the residual sum of squares (SS = 5430.064). The predictor variable of ‘opleiding2’ exhibited a marginally significant positive standardized coefficient ($\beta = .050$, $t(1739) = 2.072$, $p = .038$). This suggests that higher levels of ‘opleidingsniveau2’ were associated with slightly increased levels of fear of the coronavirus (B = .179, SE = .087). The 95% confidence interval for the coefficient ranged from 0.010 to 0.349 and the intercept was 2.354 (SE = .055) with a confidence interval ranging from 2.247 to 2.461.

Lastly, the regression model with ‘opleidingsniveau3’ as dependent variable also portrayed a weak positive association ($R = .164$, $R^2 = .027$, adjusted $R^2 = .026$). Moreover, the analysis of variance indicated a significant effect of the regression model ($F(1, 1739) = 48.325$, $p < .001$). The regression model accounted for 2.7% of the variance in the dependent variable, as shown by the regression sum of squares (SS = 147.178) compared to the residual sum of squares (SS = 5296.290). Additionally, the predictor variable "opleidingsniveau3" displayed a significant negative standardized coefficient ($\beta = -.164$, $t(1739) = -6.952$, $p < .001$). This suggests that higher levels of "opleiding3" were associated with decreased levels of fear for the coronavirus (B = -.599, SE = .086). The 95% confidence interval for the coefficient ranged from -0.768 to -0.430. The intercept was 2.653 (SE = .053) with a confidence interval ranging from 2.549 to 2.757.

The results of these three tests are significant, which therefore supports the hypothesis that fear holds a greater influence on compliance predisposition than education. Based on the output of these three linear regressions, an intriguing finding is that the dummy variable ‘opleidingsniveau2’ portrayed a positive relationship in contrast to the other findings in the linear regression. This implies that those who hold an education level in the range ‘midden’ are more likely to comply with the regulations based on fear, whereas the other education levels hold a diminishing influence on compliance.

## Reflection & Conclusion

In the context of the COVID-19 pandemic, this research emphasizes the freedom researchers have and shows how it plays a vital role in informing public health policies and interventions. Moreover, it is important to be aware of questionable research practices (QRPs) that can impact the reliability of research outcomes (Bakker et al., 2021). Therefore, it is integral to produce results of research in a credible and transparent manner. QRPs, such

as selectively reporting significant results or manipulation of data, can undermine the integrity of scientific research. Despite the fact that all findings are significant, it is difficult to state whether ‘fear’ exhibited a stronger association than ‘opleidingsniveau.’

The empirical results presented in this study builded on the literature of Özdil et al. (2021) and Bakker et al. (2020) and shed light on the relationship between the compliance of education level and fear of COVID-19. The findings suggest that there is a weak, but statistically significant and overall negative relationship between compliance based on education and fear, which in first glance seems to support the hypothesis.

It is important to note that correlation measures the degree of association between two variables in a linear fashion, while regression represents this relationship through an equation. Despite the fact that these measures may have similar outputs, the implementation of these methods are vastly dissimilar and portray inherently different perspectives to the results. Furthermore, the use of dummy variables in the linear regression model adds support and credibility to the findings, but does not establish a definitive causal relationship between compliance and fear.

With regard to the correlation, the research implies that the level of education has an effect on compliance based on fear of the virus. However, a correlation does not necessarily imply causation. It is possible that other factors or variables could be influencing both education level and compliance behavior, leading to a spurious correlation. Moreover, the findings are based on a specific sample size of 1429 participants. The generalizability of the results to a larger population or different contexts should be considered. The specific characteristics and demographics of the sample might limit the external validity of the findings.

In addition, regarding model 2, the effect size is not stated in this research output on purpose. This has been done to show how information that is left out can conjure issues when formulating definitive conclusions based on the tests that are run. Due to this fact, it is not clear what the magnitude is of the differences found in the different tests of the linear regression. In this case, due to transparency issues, the findings may have artificially inflated the effect size in order to generate significant results when there are none, this is also referred to as p-hacking (Bakker et al., 2021).

Moreover, it is fundamental to add that a theoretical justification enables the author to implement any variables or methods they want to interpret findings (van der Bles et al., 2020). In this study, it was found that 'fear' had a significantly stronger influence on determining compliance compared to education levels. This suggests that even if a linear

regression is conducted using the same variables, but different functional specifications, it would yield divergent results. However, authors can always resort to theory for support of their findings in order to benefit their own research goals.

Additionally, a researcher can also justify a certain choice to favor the outcome of their hypothesis by reporting findings that support it, which is referred to as HARK-ing (Bakker et al., 2021). In this case, the association of the linear regression is considered to be small, since $R^2$ indicates that 3.2%, 0.2% and 2.7% of the variance in fear could be explained by education level respectively. The findings are significant, so according to journals these findings are deemed to be relevant to report, since it suggests that it supports the hypothesis that is tested in this essay. However, the results of the association should be interpreted with caution, since the test does not seem to fit due to the fact that little can be explained based on the independent variable.

Ultimately, these findings highlight the importance of transparency and rigorous methodology in social science research. It is crucial for researchers to report their findings accurately and provide a clear interpretation of the limitations and implications of their work without attempting to benefit for their own gain. By doing so, the scientific community can engage in critical discussions, replicate studies, and build upon existing knowledge. Therefore, the freedom to present different results in science is a cornerstone of scientific progress. However, it comes with the responsibility to conduct research ethically, transparently, and rigorously. By upholding these principles, researchers can contribute to the advancement of knowledge and the betterment of the social sciences.

# References

Bakker, B.N., van der Wal, A., & Vliegenthart, R. (2020). COVID-19 panel study in the Netherlands. https://doi.org/10.17605/OSF.IO/KWZ7A

Bakker, B. N., Jaidka, K., Dörr, T., Fasching, N., & Lelkes, Y. (2021). Questionable and Open Research Practices: Attitudes and Perceptions among Quantitative Communication Researchers. *Journal of Communication*, *71*(5), 715–738. https://doi.org/10.1093/joc/jqab031

Özdil, K., Bulucu Büyüksoy, G. D., & Çatıker, A. (2021). Fatalism, fear, and compliance with preventive measures in COVID-19 pandemic: A structural equation modeling analysis. Public Health Nursing, 38(5), 770–780. https://doi.org/10.1111/phn.12898

van der Bles, A. M., van der Linden, S., Freeman, A. L. J., & Spiegelhalter, D. J. (2020). The effects of communicating uncertainty on public trust in facts and numbers. Proceedings of the National Academy of Sciences, 117(14), 7672–7683. https://doi.org/10.1073/pnas.1913678117

Vliegenthart, R., Bakker, B. N., & Vreese, C. (2021). Verschuivingen in informatievoorziening tijdens Covid-19. *Beleid en Maatschappij*, *48*(1), 75–83. https://doi.org/10.5553/BenM/138900692021048001005

Franco, A., Malhotra, N., & Simonovits, G. (2014). Social science. Publication bias in the social sciences: Unlocking the file drawer. *Science (New York, N.Y.)*, *345*(6203), 1502–1505. https://doi.org/10.1126/science.1255484