

SPURIOUS CORRELATIONS DUE TO INTERNAL PSYCHOLOGICAL BIASES WHEN MEASURING WITH SUBJECTIVE QUESTIONNAIRES

Georgi Petkov

Institute of Population and Human Studies

Bulgarian Academy of Sciences

Sofia 1113, Acad. Georgi Bonchev St., Bl. 6, Fl. 5/6, Bulgaria

georgi.petkov@iphs.eu

Abstract. *Estimations of correlation could be overestimated due to sample errors, measurement noise, and biases, but also because of the nature of the correlational coefficient. In particular, it is a dispute whether the correlation reflects a single common factor (or few), or that it may emerge from a huge number of common factors, including contextual ones, so many, that it is impossible to control them all and therefore, it is needless to explore them at all. It is a widespread method in psychology to use various questionnaires, assuming that they measure some not directly observable psychological constructs. In this context, some temporal factors like mood, the current condition, and some more stable ones like cultural or personal specifics may produce spurious correlations. Two known psychological biases – ends effect and favorite number effect, when answering to questionnaires were simulated and it is demonstrated how they could produce moderate spurious correlation.*

Keywords: spurious correlation; Likert scale; questionnaires.

The article can be cited as follows:

Petkov, G. (2022). Spurious Correlations Due to Internal Psychological Biases When Measuring with Subjective Questionnaires. *Psychological Research (in the Balkans)*, Volume 25, Number 1, 2022, 77-87. ISSN 0205-0617 (Print); ISSN 2367-9174 (Online). DOI: <https://doi.org/10.7546/PsyRB.2022.25.01.06>

© G. Petkov, 2022

Received: February 17, 2022

Accepted: June 13, 2022

The author has read and approved the final manuscript.

INTRODUCTION – CORRELATIONS COULD BE OVERESTIMATED DUE TO SAMPLE ERRORS, MEASUREMENT NOISE, AND BIASES

In the context of replication crisis in psychology, Stanley & Spence (2014) demonstrate that the measurement error alone might produce large discrepancies between the real value of a certain size effect and its empirical estimations and that this error could be in both directions. Loken & Gelman (2017) further on rebut the widespread assumption that both measurement error and sample error always underestimate size effects. On the contrary, small sample sizes and in general, low-powered studies might often cause overestimating it. Hong & Cheng (2019) specify how the lack of independence of two responses, combined with increasing the measurement noise (in the author's example – by speeding the answers) may produce random fluctuation of the estimation of the effect in both directions. In addition, the assumption that both measurement and sample errors are independent of the true scores is often questionable.

Any biases during the measurement procedure could obviously produce additional displacement of the estimation of any parameter from its real value in both directions. Some confounding variables could be source of spurious correlations.

However, it is also possible that the discrepancy between the true scores and their estimation is sometimes due to the nature of the correlational coefficient. In particular, it is a dispute whether the correlation reflects a single common factor (or few only), or that it may emerge from a huge number of common factors, including contextual ones, so many, that it is impossible to control them all and therefore, it is needless to explore them at all.

WHAT DOES CORRELATIONAL COEFFICIENT MEAN?

Since the concept of correlation emerged at the second half of 19th century (see for example Stigler, 1989 for a historical review), there was a considerable debate about its es-

sence. There are several different views about what does correlation mean.

Correlation reflects common factors

It is almost a dogma in psychology that any correlation has to have a reason expressed with few common factors (or a single one) underlying the correlated variables. Any correlation between A and B means either that A causes B, or B causes A, or it is a third, perhaps unseen factor C (latent variable), which influences both A and B. Following this view, there is a widespread method in psychology to use various questionnaires, assuming that they measure some psychological constructs with a high precision, and then to interpret eventual correlations among these questionnaires as being caused by additional not directly observable factors.

This view about correlation as a result of a common cause is advocated by Galton (1889), then by Spearman (1961/1904). Psychology research still follows it assuming that the common cause consists of a single factor or at least only a few ones. Further on, Reichenbach (1956) explicitly formulated the Common Cause Principle of correlated events (i.e. $P(A \cap B) > P(A) \cdot P(B)$) and formalized it mathematically, thus trying to expand the probability theory, incorporating the Common Cause Principle axiomatically.

Independence as a point on a continuum

However, there are other opinions and counterexamples. Starting from the Reichenbach's description of correlated events and using the basic rules of probabilities, Hitchcock & Miklós (2020) propose radically different view of independence and hence, of correlation (see Figure 1).

There is a positive correlation between two events A and B if and only if the inequality $P(AB) > P(A)P(B)$ is satisfied. What is important here is that a lack of correlation, i.e. independence of A and B does not mean that there are not common factors underlying them. In contrary, such a case will result in a strong negative correlation (see Figure 1, left). Independence is just a point on a continuum. When we move A and B horizontally

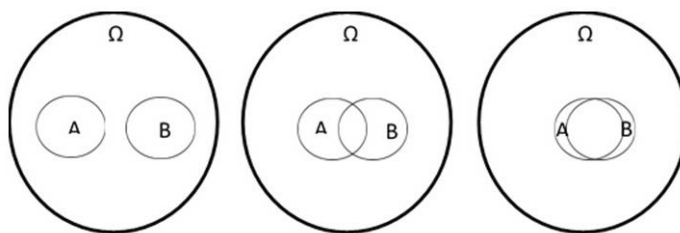


Figure 1. On the middle panel – independent events A and B ($P(AB) = P(A)P(B)$). If we shift the circles representing A and B horizontally, we will see either negative correlation (left panel) or positive one (right panel). Nevertheless that the independence has important properties, it seems like zero correlation (middle panel) is just a point onto a continuum.

(Figure 1), their common area $P(AB)$ changes from zero to $P(A)$ ¹. Independence could be observed during this move only at the point at which $p(AB)$ becomes exactly equal to $P(A)P(B)$. But there is nothing special according to the underlying common factors here (see for additional analysis on another examples also Lewis, 1973).

One can think on the scheme on Figure 1 as a game of darts but also as photons hitting a scintillation screen after passing through a narrow slit. Finally, Figure 1 could be thought as a Venn diagram and hence the argument is valid for various types of events. Consider as well the following example: If we have a standard deck of 52 cards and pick a single card randomly, then the events “picking a heart” and “picking a queen” are independent. However, adding a single joker to the deck (which is neither heart, nor queen) destroys the independence and the two events become correlated. It is hard to imagine any common underlying factor(s) that can explain this correlation.

Going back from the events and probability theory to the variables, consider time series correlations. Sober & Elliott (1984) for example point that sea level in Venice and bread prices in London have both been rising during last few centuries, thus they are correlated. It is very difficult to assume, however, any common factor that causes both risings. Researchers from the field of economy for ex-

ample know well this phenomenon – everything rises and this causes correlations.

Pearson's view of correlation

Being involved deeply in philosophy of science, Pearson looks for answers of how to detect causal relations. According to him, causality is nothing than correlation with temporal precedence (Pearson, 1892). Correlation in turn could be viewed as a weaker form of causation: “It is this conception of correlation between two occurrences embracing all relationships from absolute independence to complete dependence, which is the wider category by which we have to replace the old idea of causation” (Pearson, 1892, p.157). Thus, Pearson actually expand the old concept of causation. According to him, causality is not yes or no phenomena but can have degrees and correlation is the exact measure of that degree. Pearson's view is in accordance with most of the research lines of contemporary economy. When the economists look for relationships between certain prices for example and various other factors like weather, political situation, etc. they rarely look for any common causes. In such cases the causal models are more or less clear and non-controversial, thus correlation directly could be interpreted as a degree of causation.

Pearson, however, does not reject the role of eventual common factors for producing a correlation between two variables. On

¹ Assume for simplicity that $P(A)=P(B)$ and $P(\Omega)=1$

the contrary, he studies their role intensively. However, he opposes to the advocates of the common factor explanation of correlation regarding the number of the eventual common factors. Namely, he assumes that the number of common factors usually is not one, two, or few, but could be very large and even infinite (see for deeper details Aldrich, 1995). The imperative implication of this position is that it is not interesting to study common factors. Pearson's position does not allow us to deduce any common factor, nor few such, from the eventual correlations among variables. Even more, a significant part of the observed correlations could be spurious, i.e. produced by uncontrolled common factors. For example: "We are thus forced to the conclusion that a mixture of heterogeneous groups, each of which exhibits in itself no organic correlation, will exhibit a greater or less amount of correlation. This correlation may properly be called spurious." (Pearson, Lee, and Bramley-Moore, 1899, cited by Aldrich, 1995, p. 369). Different reasons for eventual spurious correlation will be shortly discussed in the next section.

SPURIOUS CORRELATIONS

Aldrich (1995) analyzes the dispute between Pearson, Yule, and others more than a century ago. In addition, he summarizes various reasons for observing spurious correlation.

Probably the dispute between Pearson and Yule had been initiated from a Pearson's critique of the analysis of one of his biologist colleagues, Weldon (Aldrich, 1995). Weldon (Weldon, 1892, cited from Aldrich, 1995) normalized the size of different animal's parts in terms of body lengths. However, Pearson notices that if we pick randomly three types of bones from different animals and calculate two ratios – between the first and third type and between the second and the third type, we would obtain a spurious correlation. Generally speaking, if we take three random variables A, B, and C, there always be a correlation between the ratios A/C and B/C, even if A, B, and C themselves are independent. This is because if a concrete value of C happens to be small, both concrete values of A/C and B/C

would be large and vice versa. This simple observation, however, poses difficult questions. For example, is it correct to correlate entities like GDP per capita and birth rate per capita? Actually, many of the variables we use are normalized in one or another way. Yule's answer was delicate – he argued the "organic nature" of some variables requires ratios, but of some other doesn't. It seems that Pearson dislikes such subjective interpretations.

Some confounding variables could be additional source of spurious correlations. For example, if a psychological research includes calculation of correlation between the performance of two behavioral tasks and the participants do the tasks soon one after the other, the mood could be such a confounding variable. Even if it is no organic correlation between the two performances, some people would be in a good mood during the examination and would achieve good results of both tasks; some participants, vice versa, will be underestimated in both tasks. Obviously, this could produce a spurious correlation.

Let's bring back to two of the different views about the nature of correlation. The modern psychology follows Spearman's and Galton's understanding of the correlation as an effect of a common cause. Pearson, however, realized that it is not necessary to assume a single common cause, even few ones. The common causes could be (and in most of the cases are) huge number, maybe even infinite number. When we observe a correlation between two psychological measures for example, it could be due to a latent factor, but could be also due to a common context of the measurements.

In the next section, two possible sources of a *common context* when measuring psychological constructs with questionnaires on Likert scales will be examined.

PSYCHOLOGICAL EFFECTS WHEN TESTING WITH QUESTIONNAIRES

Ends effect

Batchelor & Miao (2016) propose a meta-analysis of the so-called Extreme Response Style. Some people tend to respond more with

the extreme grades to questionnaires. Some other people, however, vice versa, avoid them and prefer the options from the middle. There are significant culture differences (Wang et. al, 2008). There are also demographic, racial, and maybe gender factors that influence the response style (Bachman et. al, 2010). There are pieces of evidence that the response style is relatively stable (Lau, 2008) but some researchers argue that it depends on contextual factors like mood as well (Paulhus, 1991). When people have to grade some objective characteristics (like lengths of lines, etc.) on a Likert scale, there is a rigid tendency people to prefer the answers in the middle of the scale and to avoid the extremes (Petrov & Anderson, 2000, Petkov, 2013).

Favorite numbers

Another example of bias when people answer on Likert scales is the fact that often they have favorite numbers and arrange their answers close around them. The presented individual differences according to the preferences of the extremes or the middle points are examples for specific personal preferences. In addition, Moors (2008) find out that the number of the available grades (specifically, whether there are odd or even number of options) influence people's preferences. Chan (1991) reports for dependencies of the order of presentation of the available response options to the response style of the people. Finally, social desirability bias differs in its strength among people and may provoke individual differences of consistently lower or higher grades (see for example Garland, 1991).

Certainly, there are statistical methods for evaluating the strength of exactly these effects and hence, to eliminate them. However, these are just two arbitrarily chosen biases when people answer to questionnaires on a Likert scale. The *context* of filling the surveys may include huge number of similar biases. Even if we can evaluate statistically the strength of each concrete bias, it is challenging to account for all of them. At least, it is difficult to find evidence for separating the influence of contextual factors from those of (again not directly observable) latent variables of interest.

In the next section some computer-simulated experiments are presented, aiming to propose some evidences that the context factors may produce significant spurious correlations between fully independent variables.

SIMULATIONS

The purpose of the simulations were to investigate further the role of the contextual factors of end effects and number preference on the correlation among results of judgment on Likert scale. Although the spurious correlation among ratings A/C and B/C even when the variables A and B are independent themselves could be theoretically implied, the exact values of this correlation when the strength of the contextual factors vary, is difficult to be calculated algebraically. Thus, the simulations would give deeper picture of the impact of the factors explored.

Procedure

First, two random variables were independently generated from a uniform distribution. They were fitted algebraically to 1 to 7 Likert scale thus simulating independent (non-correlated) 'answers' of a sample of people to two different surveys on a 1 to 7 Likert scale.

Second, additional variables were generated representing various contextual factors. Thus, for each simulated participant, we had two numbers representing their respective simulated latent values for the two questionnaires, as well some numbers representing their contextual situation.

Third, additional algebraic manipulations were performed, thus adjusting the original values to 'answers' of the two questionnaires according to the contextual factors. Finally, the Pearson's correlational coefficient between the modified 'answers' was calculated.

In the first simulation both *preference to the middle* and *favorite number* biases were simulated. The second simulation explored the role of the *preference to the middle* tendency alone; the third simulation was about the *favorite number* alone.

Simulation 1. Measuring non-correlated constructs on Likert scales with ends effect and favorite numbers

A set of 100 uniformly distributed random numbers were generated and were adjusted to discrete answers from 1 to 7 Likert scale. Thus, answers of 100 people to one-item questionnaire were simulated assuming uniform distribution. Let's call these numbers answers of the A test.

The procedure was repeated and 100 independent answers of B test were generated.

Not-surprisingly, there was not a significant correlation between the two lists of 100 random numbers. In order to obtain statistical results, the procedure was repeated 100 times.

On the next step, the following adjustments were made: Two additional random variables were generated with 100 numbers in each. Let's call the first variable "Skew_bias" and the second one "Mean_bias". The Skew_bias variable was extracted from an uniform distribution $U(0,1)$. It represents the tendency of people to concentrate their answers around the middle of the scale. If we assume a simulated person with Skew_bias = 1, then their resulting distribution would be uniform; if Skew_bias = 0, all their answers would be equal. The Mean_bias represent eventual favorite number of the simulated person. It was extracted uniformly from the whole numbers between 1 and 7.

Each of the original variables A and B was transformed as a weighted sum of the initial numbers and the Mean_bias, using Skew_bias as a weight:

$$A[i] = A[i] * \text{Skew_bias}[i] + \text{Mean_bias}[i] * (1 - \text{Skew_bias}[i])$$

$$B[i] = B[i] * \text{Skew_bias}[i] + \text{Mean_bias}[i] * (1 - \text{Skew_bias}[i])$$

for each i between 1 and 100.

Thus, the role of Mean_bias could be thought also as altering the mean of the original variable; and the Skew_bias alters it's variance. The essence of the simulations was to look for spurious correlations when altering the mean and the variance of two variables synchronically.

Thus, the resulting scores are still generated from random numbers but reflect the respective Mean_Biases and Skew_Biases for each individual simulated person. The important thing is that we assume that people differ each other on the degree of these biases. Thus, each person has their own values for both biases and these values reflect their both answers to A and B.

Not surprisingly, a moderate correlation between A and B was observed. A histogram of 100 runs of the simulation is presented on Figure 2. The correlation comes fully from the third and fourth of the variables, which

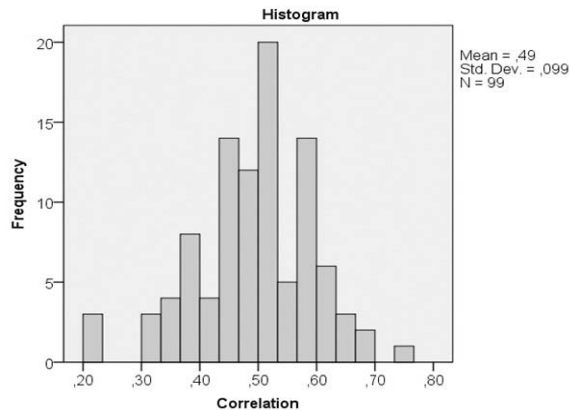


Figure 2. Histogram of the correlation coefficients between two adjusted lists of 100 random numbers (see text for details)

participate in the algebraic calculations. However, they represent natural biases of people’s behavior when answering to questionnaires’ on Likert scales.

Simulation 2. Exploring only the non-uniformity of the answers bias

During the second simulation only the Skew_{bias} was varied, fixing Mean_{bias} for all simulated participants to the middle grade 4. The correlational coefficients vary in accordance of a normal distribution, centered in 0.0, thus no any spurious correlation had been observed.

On the next step, however, the Mean_{bias} was systematically varied among the numbers from 1 to 7. Thus, the same influence of Skew_{bias} was explored but for various fixed values of the Mean_{bias}. The results are represent on Figure 3.

Simulation 3. Exploring only the role of the favorite number

During the third simulation only the Mean_{bias} varied (it was a generated random variable), fixing Skew_{bias} to the middle value of 0.5. The histogram after 100 runs is presented on Figure 4.

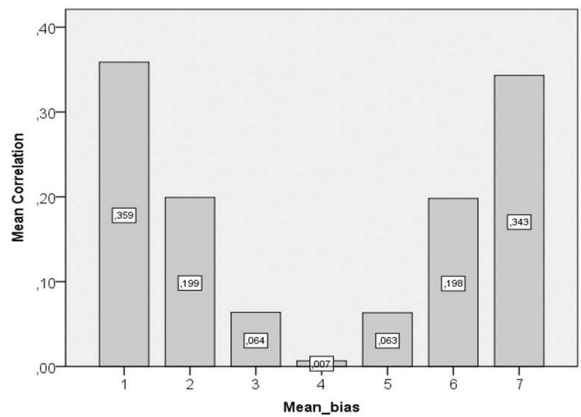


Figure 3. The mean spurious correlational coefficients produced by the non-uniformity bias for various fixed values of the favorite number bias

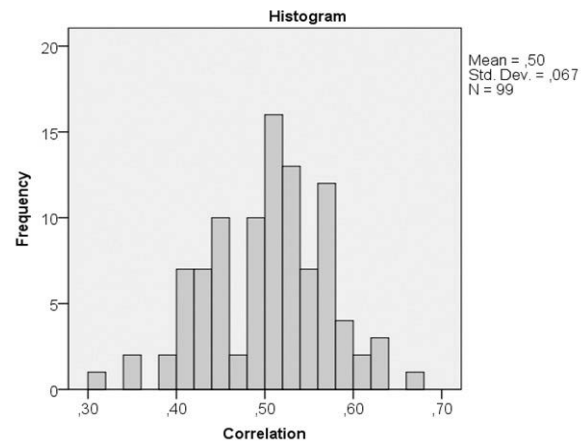


Figure 4. Histogram of the correlation coefficients between two adjusted lists of 100 random numbers varying only the favorite number bias among the simulated participants

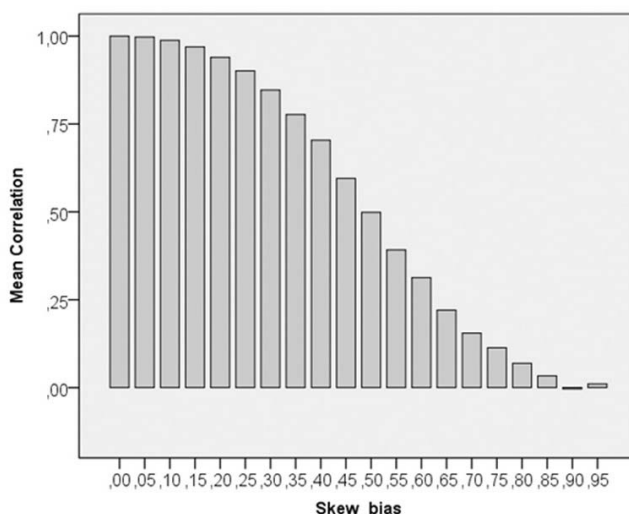


Figure 5. The mean spurious correlational coefficients produced by the favorite number bias for various fixed values of the non-uniformity bias

On the next step, the *Skew_bias* systematically varied among the interval $[0, 1]$ with step 0.05. Thus, the same influence of *Mean_bias* was explored but for various values of the *Skew_bias*. The results are represent on Figure 5.

Summary of the results and discussion

The first simulation demonstrates a moderate spurious correlation (about $r=0.50$) when altering two independent variables algebraically. The results of the simulation are not surprising. The findings could be proven analytically, thus the simulation could be viewed as a demonstration only. It is interesting that Pearson (1897) found out that correlations about 0.4-0.5 might arise when calculating ratios of variables with equal variance. He argued that the null hypothesis should be set to the value of 0.5 of correlational coefficient instead of 0.0. Only correlations significantly above 0.5 should be interpreted. Otherwise, one could be a spurious correlation.

The altering of the independent variables in the first simulation resulted in similar value for the spurious correlation. This result came from algebraic manipulations with random variables, thus the setting could be interpreted

as relatively close to a real-world situation.

Instead, during the next simulations the roles of the two biases were studied systematically. As seen from the second simulation, it seems the role of moving the mean have much greater impact than changing the variance. When we move each pair of the values of two independent variables on one or another direction, obviously a correlation will emerge. When we move the variance only, however, there is no reason for that.

The third simulation, however, demonstrates again the combined role of the two biases on the emergence of spurious correlations. Changing the mean alone has a large impact. Changing the variance alone doesn't. However, altering the variance modulates the effect of the mean changing. Actually all positive numbers from 0.0 to 1.0 are possible values for spurious correlation in certain circumstances.

CONCLUSIONS

The psychological program for looking for a net of latent variables that may explain a lot of the people's behavior had been launched more than a century ago with the work of Spear-

man, Galton, and others. The general idea is that looking on the correlations among various behavioral manifestations, we may deduce the existence of such latent variables. Further on, we may design questionnaires for a precise measuring these invisible latent variables. Even further on, we may produce for a network of psychological constructs and precisely calculate the distances among them.

Even Spearman had noticed some of the challenges of this ambitious program. His analysis of the previous studies resulted in a large list of methodological issues when doing such psychological conclusions.

Nowadays, although the statistical tools are very advanced, still many of the studies could be criticized for methodological issues. There are hundreds of psychological constructs proposed, with various different possible interconnections among them, without a clear consensus among the researchers almost everywhere. In addition, we witnessed the replication crisis in psychology (see for ex. Maxwell et. al., 2015, Loken & Gelman, 2017) and the publication bias one (for ex. Easterbrook et. al, 1991). Most of these problems could be due to some bad scientific practices. However, we should admit that the research program for finding a clear rigid network of latent psychological constructs that may explain a lot of the human behavior is not much advanced from the earlier Spearman's work. This failure could not be explained purely with some bad scientific practices only, nor with methodological issues. Maybe we should reconsider the essence of the program itself and in particular, the assumption that exploring the correlational coefficients among some measures may lead us to understanding behavior in terms of invisible, latent factors.

The early Pearson's work may guide us why this assumption could be wrong. He opposed to Yule, Spearman and others on the idea of an existence of a single factor (or few only) that could explain the correlations. Instead, he assumed that correlations may emerge from a huge number of common factors, so many, that it is impossible to control them all and therefore, it is needless to explore them at all.

In particular, it is a widespread method among the psychologists to use various questionnaires assuming that they measure with a high precision some psychological constructs and to correlate them to interpret eventual correlations as being caused by additional not directly observable factors. Following the line of thought of Pearson, however, the basic assumption of this method is questionable. The correlational coefficients are due to *all* common factors. Some of them could be the wanted variables, but others may come from the *context*. Part of this context could be temporal, like mood and the current condition. Another part of this context could be more stable and could come from the culture or from some personal specifics.

The presented simulations explored two such possible contextual factors – the tendency of people to have preferred ratings and the bias for skewed, non-uniformly distributed answers. People differ each other on these biases depending on the culture, demographics and some personal specifics. Relatively large spurious correlations (about $r=0.5$) emerged in simulated settings for these biases only, without any hidden latent factors. Although the simulations were very simple (they are implemented on few lines Python code), they could be easily replicated and even analytically could be supposed that the results would be rigid.

Of course, the results from the simulations do not prove that many of the psychological conclusions based on correlations between questionnaires are wrong. The simulations show only that it is possible *some* of them to be not accurately estimated. Unfortunately, it is not easy to distinguish what portion of the correlational coefficient is due to important latent factors and what portion is due to contextual factors.

There are some additional limitations of the presented simulations. For example, additional effort is needed to judge whether the negative correlations, which are part of the psychological program for extracting a network of interconnected not directly observable psychological constructs, could be due to some contextual factors, irrelevant to the

concrete studies. Although there are a lot of evidences for the two simulated biases, additional empirical investigation of the assumptions behind the simulations would be helpful for better understanding of the problem.

Finally, the presented study is focused on the nature of the correlation itself, even without the context of the psychological research. Correlation is probably the concept that is in the hearth of the whole statistics. It is definitely one of the most important concepts and underlies most of the statistical methods. Deeper philosophical understanding of the essence of this concept would result in a significant progress of research methods.

REFERENCES

1. Aldrich, J. (1995). Correlations genuine and spurious in Pearson and Yule. *Statistical science*, 10(4), 364-376. <https://doi.org/10.1214/ss/1177009870>
2. Bachman, J. G., O'Malley, P. M., & Freedman-Doan, P. (2010). *Response styles revisited: Racial/ethnic and gender differences in extreme responding*. Institute for Social Research. <http://monitoringthefuture.org/pubs/occpapers/occ72.pdf>
3. Batchelor, J. H. & Miao, C. (2016). Extreme response style: A meta-analysis. *Journal of Organizational Psychology*, 16(2), 51-62. <https://articlegateway.com/index.php/JOP/article/view/1790>
4. Chan, J. C. (1991). Response-order effects in Likert-type scales. *Educational and Psychological Measurement*, 51(3), 531-540. <https://doi.org/10.1177/0013164491513002>
5. Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 337(8746), 867-872. [https://doi.org/10.1016/0140-6736\(91\)90201-Y](https://doi.org/10.1016/0140-6736(91)90201-Y)
6. Galton, F. (1889). I. Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London*, 45(273-279), 135-145. <https://doi.org/10.1098/rspl.1888.0082>
7. Garland, R. (1991). The mid-point on a rating scale: Is it desirable. *Marketing bulletin*, 2(1), 66-70. http://marketing-bulletin.massey.ac.nz/V2/MB_V2_N3_Garland.pdf
8. Hitchcock, Ch. & Miklós R. (2020). Reichenbach's Common cause principle. In Edward N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition). <https://plato.stanford.edu/archives/sum2021/entries/physics-Rpcc/>
9. Hong, M. Y. & Cheng, Y. (2019). Clarifying the effect of test speededness. *Applied psychological measurement*, 43(8), 611-623. <https://doi.org/10.1177/0146621618817783>
10. Lau, M. Y. K. (2008). *Extreme response style: An empirical investigation of the effects of scale response format and fatigue*. University of Notre Dame. <https://curate.nd.edu/show/jm214m92q0s>
11. Lewis, D. (1973). Counterfactuals and Comparative Possibility. In: Harper, W.L., Stalnaker, R., Pearce, G. (eds) *IFS* (pp. 57-85). The University of Western Ontario Series in Philosophy of Science, vol 15. Springer, Dordrech. https://doi.org/10.1007/978-94-009-9117-0_3
12. Loken, E. & Gelman, A. (2017). Measurement error and the replication crisis. *Science*, 355(6325), 584-585. <https://doi.org/10.1126/science.aal3618>
13. Maxwell, S. E., Lau, M. Y. & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, 70(6), 487-498. <https://doi.org/10.1037/a0039400>
14. Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *Quality & quantity*, 42(6), 779-794. <https://doi.org/10.1007/s11135-006-9067-x>
15. Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17-59). Academic Press. <https://doi.org/10.1016/B978-0-12-590241-0.050006-X>
16. Pearson, K. (1892). The grammar of science. *Nature*, 46, 247. <https://doi.org/10.1038/046247b0>
17. Pearson, K. (1897). Mathematical contributions to the theory of evolution. On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the royal society of london*, 60(359-367), 489-498. <https://doi.org/10.1098/rspl.1896.0076>
18. Pearson, K., Lee, A. & Bramley-Moore, L. (1899). Reproductive or Genetic Selection. *Science*, 9(217), 283-286. <https://doi.org/10.1126/science.9.217.283>
19. Petkov, G. (2013). *JUDGEMAP: Context-Sensitive Cognitive Model of Judgment on a Scale: Modelling the Process of Judgment on the Basis of the Mechanisms of Analogy-Making*. Lap Lambert Academic Publishing. <https://www.morebooks.de/store/gb/book/judgemap-context-sensitive-cognitive-model-of-judgment-on-a-scale/isbn/978-3-659-46653-3>
20. Petrov, A. A., & Anderson, J. R. (2000). ANCHOR: A memory-based model of category rating. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society* (pp. 369-374). <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.217.1381&rep=rep1&type=pdf>
21. Reichenbach, H. (1956). *The direction of time*. University of California Press. <https://doi.org/10.2307/2216858>
22. Sober, Elliott (1984). Common Cause Explanation. *Philosophy of Science*, 51(2), 212-241. <https://doi.org/10.1086/289178>
23. Spearman, C. (1961/1904). "General Intelligence" Objectively Determined and Measured. In J. J. Jenkins & D. G. Paterson (Eds.), *Studies in individual differences: The search for intelligence*

- (pp. 59–73). Appleton-Century-Crofts. <https://doi.org/10.1037/11491-006>
24. **Stanley, D. J. & Spence, J. R.** (2014). Expectations for replications: Are yours realistic? *Perspectives on Psychological Science*, 9(3), 305-318. <https://doi.org/10.1177/1745691614528518>
 25. **Stigler, S. M.** (1989). Francis Galton's account of the invention of correlation. *Statistical Science*, 4(2), 73-79. <https://www.jstor.org/stable/2245329>
 26. **Wang, R., Hempton, B., Dugan, J. P. & Komives, S. R.** (2008). Cultural differences: Why do Asians avoid extreme responses? *Survey Practice*, 1(3), 1-7. <https://doi.org/10.29115/SP-2008-0011>
 27. **Weldon, W. F. R.** (1892). I. Certain correlated variations in crangon vulgaris. *Proceedings of the Royal Society of London*, 51(308-314), 1-21. <https://doi.org/10.1098/rspl.1892.0001>

ASSOC. PROF. GEORGI PETKOV, PHD

Institute of Population and Human Studies
Bulgarian Academy of Sciences
georgi.petkov@iphs.eu

The research interests of Georgi Petkov (https://www.iphs.eu/n/images/cv-docs/G.Petkov/CV_en_Georgi_Petkov.pdf) are in cognitive science, statistics and methodology, relational thinking, analogy-making, memory, cognitive modelling. He had teach courses in Statistics, SPSS, Programming, Memory, Analogy-making in the psychology departments of New Bulgarian University, National Security and Defence Faculty, Bulgarian Academy of Science, and others.

Copyright (c) 2022 Psychological Research (in the Balkans)



This work is licensed under a **Creative Commons Attribution 4.0 International License**.