Nikola ILIĆ, LL.M., M. Econ.*

Kahneman, Daniel, Olivier Sibony, Cass R. Sunstein. 2021. *Noise: A Flaw in Human Judgement*. New York, Boston, London: The Little, Brown Spark, 464.

„Silence is foolish if we are wise, but wise if we are foolish."

Charles Caleb Colton

Imagine you went to see a doctor who suggested further medical examinations due to the possible heart disease. Following her recommendations, you consulted three specialists in the field and got three different opinions. The first cardiologist suggested immediate surgery, the second suggested intensive medical therapy excluding surgery, while the third cardiologist recommends a conservative medical treatment and further observations. You are surprised and confused, but you still have to make a final decision: what would you do?

Before answering the question, you could consider the „noise", a phenomenon Daniel Kahneman, Olivier Sibony, and Cass R. Sunstein analyse and explain in their recent book. The co-authors are the leading authorities in psychology, management, behavioural economics, and law,[1] which is one of the reasons this book is a candidate for the Amazon 2021 bestseller. Yet, the main reason for the increasing popularity of this book probably lies in the fact the „noise" is not relevant for the medical profession and doctors only, but also judges, lawyers, managers, insurers, and all other decision-makers in the public and private sector.

* Lecturer, University of Belgrade Faculty od Law, Serbia, *nikola.ilic@ius.bg.ac.rs*.

[1] Daniel Kahneman is a psychologist and economist, Princeton University Professor, winner of the Nobel prize in economics (2002), and author of the New York Times bestseller *Thinking Fast and Slow* (2011). Olivier Sibony is a HEC Paris professor, specialising in strategic decision making, and Cass R. Sunstein is a Harvard Law School professor, among other things, the co-author of *Nudge*, one of the best books according to The Economist (2008).

Excluding an engaging introduction with an illustrative „noise" example, the book has six main parts. In the first part, the co-authors explain the distinction between cognitive biases and the „noise" using the example of a judge ruling.[2] Many empirical studies have shown that judgements in the same or similar cases may significantly differ depending on the cognitive biases. For instance, the study of court decisions in Louisiana between 1996 and 2012 has found a statistically significant correlation between losses of the local football team on the weekend and harsher sentencing on Monday, and *vice versa* (Eren, Mocan 2018). Similarly, an analysis of 5 million court decisions in France suggests that judges are significantly more lenient when deciding on a defendant's birthday (Chen, Phillipe 2020), etc.

All these patterns of deviation in human judgment, i.e. in judges rendering, represent cognitive biases (Haselton, Nettle, Andrews 2005, 742). However, as opposed to cognitive biases, the „noise" occurs when decision-makers in identical cases (under the same cognitive biases) decides differently. In other words, the co-authors define „noise" as an „unwanted variability in judgement" that could create „rampant injustice, high economic costs, and errors of many kinds" (p. 24).

In the second part of the book, the co-authors discuss what kind of a method should be applied to identify and measure the „noise". In their opinion, in science, as in everyday life, measurement implies the use of instruments and the assignation of standardized values to given objects or events. In that way, one can measure the length of a road, air humidity, strength of an earthquake, etc. Similarly, judges determine sanctions, insurers specify risk, and doctors formulate diagnoses. In other words, these professional decision-makers use their mind as a measuring instrument to assign values on a scale to given events (for instance, for lawyers measurement implies subsumption, i.e. legal qualification of certain activities or events). However, the accuracy of the measurement is never perfect, even in natural sciences, much less in professional reasoning. There are always some errors, and the co-authors claim some of them could be biased, some „noisy". For instance, imagine a judge should sentence defendants to ten years in prison for a crime they committed. In three identical cases,[3] a judge may render different decisions: four, two, and six years of jail. On one side, the difference, positive or negative, between the mean length of the three sentences and ten represents a bias.

2 In addition to judge ruling, the authors also use insurance and other examples to show that noise is ubiquitous.

3 The co-authors always use the assumption that the cases are identical and then analyse judgement variability. The readers should be aware that such an assumption is highly unrealistic, at least in the judiciary, where every case is unique. It is similar in medicine, insurance, and many other fields.

On the other side, the variability of the sentences constitutes the „noise“. Since the variability in statistics is commonly measured by the standard deviation, the co-authors suggest the same for the „noise“. Naturally, when the standard deviation is higher, the „noise“ is lauder and *vice versa.*

The third part of the book further analyses predictive judgements. That is a specific type of judgement containing predictions, and since the co-authors can evaluate variability in these predictions, they can more accurately observe and explain the „noise“. Among other things, they analyse and compare predictions made by professionals and by machines. In the co-authors' opinion, machines are distinctly superior to humans in predictive judgements. Even though Meehl (2013) concluded the same a long time ago, the authors provide additional evidence supporting the superiority of mechanical predictions. Namely, predictive judgements are made based on the available information or predictors. For instance, different candidates may have similar qualities, such as education, relevant experience, etc., when applying for the same job. In this case, professionals would assess all the predictors of future job performance, consult their intuition, and pick the candidate that seems the best. In contrast to that clinical judgement, a machine could conduct a multiple regression and provide a predictive score, i.e., a weighted average of all the predictors. In that way, the mechanical judgement minimizes the MSE (mean squared error) of the predictions in the same cases and thus is superior to clinical judgement. In the co-authors' words, the mechanical judgement eliminates variability in judgement, i.e. the „noise“, while clinical judgement could be very „noisy“.

After identifying and explaining the „noise“ on practical examples, the co-authors turn their attention to the possible causes of this phenomenon. The fourth part of the book explains that „noise“ occurs due to the many differences among decision-makers, including significant differences in their personality and cognitive styles. Also, as one of the underlying causes, the co-authors highlight „idiosyncratic variations in the weighting of different considerations“. It seems that the co-authors would come up with every possible explanation except to consult the neo-classical economics, i.e. the theory of rational decision making, and admit that all decision-makers have their individual preferences. If the co-authors did that, the variability in judgement, i.e. the „noise“, would be self-evident, and then it could be further debated whether or not the identified variability could be harmful. However, after describing some of the causes, the co-authors turn their attention to the „noise“ reduction.

In other words, in the fifth part of the book, the co-authors explore how to decrease the undesirable variability and improve professional judgement in practice. They introduce several noise-reduction techniques under the collective label „decision hygiene“. The first technique is „information sequencing", and it implies that decision-makers should have only the necessary information to make a judgement. All other information should be unavailable to decision-makers because additional information could generate different impressions and increase the „noise". In theory, that may seem like an effective solution. However, in practice, it is not clear who, on what ground, and how will decide whether some piece of information is necessary or not. Namely, the procedure of information screening also implies judgement, and the co-authors did not provide an answer to the question of who should judge before the judge, i.e., who should eliminate „noise" at the information screening level. The second noise-reduction technique refers to predictive judgements and „noisy" forecasts. The co-authors point out forecasts in practice, from weather to financial markets, bond, and stock prices could be very „noisy". In their words, all these forecasts may be improved by choosing a better forecaster or aggregating multiple independent estimates. The first solution seems almost trivial, and on top of that, finding a more skilful forecaster also implies judgement. Furthermore, even if one forecaster is better than the others, that does not mean her/his very next forecast will be equally successful because no one (still) knows what the future brings. That is also the main reason why the second suggested solution is not watertight. Aggregating two or more forecasts to an average, i.e. aggregating multiple independent estimates, may eliminate the „noise", but that is not, in any way, a warranty of an accurate forecast. Once again, it seems that the co-authors refuse to consult neo-classical economic theory and use some of its fundamental findings. In this case, the asymmetry of information and the EMH (Efficient Market Hypothesis) could explain why forecasters have different estimates, and consecutive accurate forecasting (alpha generation) is impossible.[4] To put it simply, no one can possess all the information in order to predict the future to the tiniest details. Moreover, the only way to obtain higher returns or be more successful than others at forecasting is to take more risks, i.e. to forecast in a relatively more „noisy" environment.

In the sixth part of the book, the co-authors admit there is an optimal „noise" level. In their opinion, in some areas, it is not feasible, while in others could be too expensive to reduce the „noise" to zero. Nevertheless, they did not provide any parameters on how to specify the optimal level. Instead of providing some guidelines, the co-authors only claim that the current „noise" level is unacceptable and that it should be reduced. To what extent? No one knows.

[4] The EMH, in general, refers to financial markets and investment strategies for forecasting the prices of various financial instruments. However, the same logic could be applied to all other types of forecasts. For more details about the EMH and its different versions, see Bodie, Kane, Marcus 2013, 347–350.

Now, let us go back to the initial heart disease example and further medical treatment. You received three different opinions, and you are still thinking about what to do next. After reading this book, it is clear there is variability in professional judgement, and the co-authors simply call that variability the „noise“. The co-authors suggest that variability is unacceptable and should be reduced, but they are not sure to what extent. However, that is not an obstacle for them to further suggest some reduction techniques, such as: the three cardiologists should be provided with the necessary information only (i.e. information sequencing), you should find a better cardiologist (i.e. a better forecaster), or you should not listen to any of the cardiologists in particular, but try to find a golden mean (i.e. aggregating multiple independent estimates). In other words, if you felt upset because of variability in professional judgement before reading this book, then after reading the book you might feel upset and disappointed.

Finally, when the disappointment gradually disappears, new questions arise: what could be done? Should all three cardiologists have the same diagnosis and recommendations? In other words, should all professionals be silent instead of making noise? Well, it depends – „Silence is foolish if we are wise, but wise if we are foolish“. Perhaps this could be one of the parameters the co-authors could use to establish the optimal level of the „noise“ they create.

REFERENCES

[1] Bodie, Zvi, Alex Kane, Alan J. Marcus. 2013. *Essentials of Investments*. New York: McGraw Hill Higher.

[2] Chen, Daniel L., Philippe Arnaud. 2020. Clash of Norms: Judicial Leniency on Defendant Birthdays. *Social Science Research Network*. Available at: *http://dx.doi.org/10.2139/ssrn.3203624* (last visited 31 July, 2021).

[3] Eren, Ozkan, Naci Mocan. 2018. Emotional Judges and Unlucky Juveniles. *American Economic Journal: Applied Economics* 10(3): 171–205.

[4] Martie, Haselton G., Daniel Nettle, Damian R. Murray. 2005. The Evolution of Cognitive Bias. 724–746 in *The Handbook of Evolutionary Psychology*, edited by David M. Buss. New Jersey: John Wiley & Sons.

[5] Meehl, E. Paul. 2013. *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence*. Battelebro: Echo Point Books & Media.