Uncertainty of Science and Decision-Making – Problems with Evidence-Based Policy

Grzegorz M. Malinowski

Kozminski University, Faculty of Economics

Abstract

The purpose of this article is primarily to introduce the topic of scientific uncertainty to the wider context of economics and management. Scientific uncertainty is one of the manifestations of irreducible uncertainty and reflection on it should enable better decision making. An entity that bases its operation on current scientific research, which depreciates over time and ultimately leads to erroneous decisions, is referred to as the "loser". The text indicates estimation of potential scale of this problem supplemented by an outline of sociological difficulties identified in the analysis of the process of building scientific statements. The article ends with a sketch of the answer to the question "how to act in the context of scientific uncertainty?"

Keywords: science uncertainty, irreducible uncertainty, loser problem, statistical errors, precautionary principle

Introduction

The phenomenon of science is undoubtedly one of the greatest achievements of humanity. Technological progress and improving living conditions mean that not only the products of scientific ventures themselves, but also, and maybe even above all, the scientific way of thinking and the so-called scientific method have become the standard that various social groups try to meet – researchers, thinkers, politicians or entrepreneurs.

The dissemination of scientific method has continued since the Enlightenment and is occupying new areas previously dominated by myths, religions, folk imaginations and traditions. As a consequence, the phenomenon of **fact-based action** crystallised, which is manifested, for example, by evidence-based medicine or evidence-based policy. The basic assumption of this approach is that action must be based on credible scientific theories whose effectiveness and safety have been confirmed by appropriate empirical research.

Nevertheless, some difficulty arises. Namely, the methodological specificity of the "scientific approach" is directly related to its main goal, which is "learning the truth", not solving specific problems. In other words, the scientific method is aimed at determining the logical status (true/false) of the research hypotheses, and not at the effectiveness of action.

The essence of this difference becomes easy to understand when viewed through the prism of "**scientific uncertainty**", or the variability and evolution of scientific theories themselves. In historical terms, scientific theories are subject to transformations, sometimes quite significant, and sometimes they are fully or partially falsified. As a result, we have an incomplete picture of the world, and although we know more and more about it, from the point of view of the decision maker, what we do not yet know may have a much more significant meaning than what we have already established.

This leads to the phenomena that I call the "loser problem". Namely, the loser is someone who ACTS based on a theory, concept or scientific model, which eventually turns out to be incorrect or incomplete, and the effects of loser's actions are deplorable. The most appropriate, though slightly populist illustration of this phenomenon, is the example of those who took their mortgage loans in Swiss francs, as the risk estimation models indicated that such a solution is safe (and there were also experts emphasizing the optimality of such action). Those faced disappointment thereafter and huge financial misfortune.

In this context it is worth paying attention to the division of science, mostly disliked by methodologists, into *basic science* and *applied science*. The essence of this classification is the statement that the fundamental goal of some scientific disciplines is discovering the truth, and of others – transforming reality into the desired direction.

It should be noted that the "loser problem" practically does not apply to basic sciences! If it did apply then, as an example, the proponents of "Earth's flatness theory", "creationism", the thesis that the earth is at most several thousand years old, or contestants of "gravity theory", "quantum mechanics" or "continental drift" – on the basis of their obviously erroneous beliefs would make fundamentally wrong decisions. However, this is not the case. These people start "normal" families and earn “normal money”. Their scientific ignorance does not transfer into ignorance in action.

The case is differently interpreted in applied science. The concepts created on the basis of disciplines such as: psychology, medicine, dietetics, sociology, economics, management or finance are the basis for specific actions of governments, investors, economic politicians, financial institutions and ordinary consumers. In these areas, the “loser problem” potentially becomes prominent in its extent. It should be taken into consideration that the scientific basis of these decisions may change over time or be rejected, and that in the future side effects associated with the given proceedings, which were previously unknown, may become visible.

The above findings lead to the conclusion that great scientific theories, which arose on the basis of basic sciences such as the theory of evolution or quantum mechanics, develop our knowledge of the world and are a direct or indirect source of many innovations that improve human existence, but are not the basis for everyday decisions – the ones on which survival depends, being understood in both biological and business – economic categories.

Therefore, a significant problem arises that must be faced by every entity making autonomous decisions. On the one hand, it is widely believed that the decisions being made are the more effective the better their scientific basis, and on the other hand there is scientific uncertainty that the decision-maker must take into account.

The aim of this article is primarily to introduce the topic of scientific uncertainty to the broader context of economics and management. Scientific uncertainty is one of the manifestations of irreducible uncertainty (Koźmiński, 2008) and reflection on it should enable better decisions to be made. Therefore, the second goal of this article is an attempt to outline the answer to the question “how to act in the context of scientific uncertainty?”.

The first part of the article introduces the issue of scientific uncertainty. The first paragraph describes the essence of this issue at the statistical level, which culminates in the preliminary estimation of the net predictive value indicator. In the following part, basic restrictions and weaknesses in making decisions based on scientific concepts (evidence based policy) are indicated. At the end of this part of the text there is an attempt to take into account these constraints within the quantitative analysis of credibility of empirical research. In the second part of this thesis, the topic of scientific uncertainty is placed in the context of two other phenomena, namely “path dependence” and “disaster risk”, which makes it possible to draw a cautious conclusion regarding normative recommendations related to making decisions in conditions of scientific uncertainty.

Why one cannot rely on “the latest research results”?

Let us try to imagine verifying the truth of 1,000 research hypotheses. Neither the content of these hypotheses nor the issues they concern are relevant. Let us also assume that only 100 of these hypotheses are true. The other 900 are false hypotheses (The Economist, 2013). The hypothesis verification procedure is based on the use of statistical tools. In this process, the researcher is exposed to two types of errors. **Type I error** (α) can be made that the false hypothesis will be considered true[1] (false positive). A true hypothesis may be also considered to be false and, thus, one can make **a type II error** (β)[2] (false negative).

Table 1. Types of statistical errors

Types of statistical errors		The hypothesis (H_0) is true	
		TRUE	FALSE
Decision on the hypothesis (H_0)	Reject	Type I error	Correct conclusion
	Do not reject	Correct conclusion	Type II error

Source: own study.

Among those who are not professionally involved in scientific research, there is a common view that the verification of scientific hypotheses is aimed at minimising the likelihood of these both errors. However, this is not the case. The probability of making a type I error and the probability of making a type II error are mathematically related in such a way that reducing the likelihood of making one type of error automatically increases the likelihood of a second type error. Thus, by minimising the probability of making a type I error, we increase the likelihood of making a type II error and vice versa.

Although each type of research procedure tries to avoid both types of mistakes, the primary goal of any scientific undertaking is to know the truth and refrain from falsehood. Although it sounds a bit lofty, it is expressed in a simple proceeding, namely, every reliable researcher is most afraid of making a type I error. This is because a researcher, who wants to expand his knowledge of the world, is much more afraid of a situation in which he adds “a brick” to the existing building of scientific statements – a hypothesis that turns out to be false over time than a situation

[1] This happens when the true H_0 hypothesis stating no relationship between the studied phenomena is rejected.

[2] This happens when the false H_0 hypothesis stating no relationship between the studied phenomena is not rejected.

in which he or she misses the truth of some hypothesis. In other words, it is better not to add anything than to add something that is not true. This approach is quite rational if we assume that the superior goal is to expand knowledge of the world. In the further part of the article it will be shown that the rationality of decision-making remains in some conflict with the scientific rationality.

Accepting 5% as the maximum allowable probability of type I error is a kind of a standard. On the other hand, the maximum admissible probability of type II error is 20%. These statistical boundary conditions mean that the testing procedure of 1,000 hypotheses can generate 45 false hypotheses that will be considered true (false positives) and 20 true hypotheses that will be considered false (false negatives). As a result, 125 hypotheses will be positively verified, and on this outcome the supporters of action based on scientific knowledge will make decisions. It should be remembered, however, that among these 125 hypotheses as many as 45 (36%) may be false.

Let us imagine that we are testing 2,400 substances for carcinogenicity. Assuming that 40% of them have carcinogenic properties, 36% do not have such properties, while for the remaining 24% we are not sure. Thus, following statistical standards discussed, when working on a sufficiently large sample, the researcher considers 43 substances to be carcinogenic, although they are not, and 192 substances to be safe, although in fact they are carcinogenic. The result: almost 200 hazardous substances will be in usage and 43 will be unjustly unapproved (Lemons et al., 1997).

Do these observations depreciate the achievements of science? Definitely not. There is no doubt, however, that the described regularity indicates that in the very foundation of the scientific study (using statistical inference) there is some irremovable, irreducible uncertainty (Weiss, 2003; Lo, 2009), and the identification of this fact leads to several conclusions:

- Scientific hypotheses must be verified several times;
- Time is the most reliable criterion for determining the truth and falseness of research hypotheses (continuous repetition of research ultimately leads to elimination of false hypotheses);
- Decision-making based on scientific knowledge cannot be the only criterion for effective action.

Why one cannot trust the results of empirical research?

Because most published scientific discoveries based on empirical research are false. J. Ioannidis (2005) comes to such an amazing conclusion in relation to medical sciences. However, his suggestion can also be transferred to social sciences.

When a scientist, basing on research findings, discovers some dependency within the field he or she represents, then, whether such a relationship really exists depends on:

A. **A priori probabilities** (before the fact) – reflects the researcher's knowledge of the analysed reality before performing the target experiments or observations;
B. **Power of statistical test** – a probability of not making a type II error. Therefore, power complements the probability of making a type II error, i.e., $1 - \beta$;
C. **Level of statistical significance** – the maximum acceptable probability of a type 1 error.

In order to prove that the majority of empirical research is false, the R-coefficient needs to be introduced, which is the quotient of true discoveries and false discoveries.

$$R = \frac{\textit{true discoveries } (T)}{\textit{false discoveries } (F)}$$

It is also worth emphasizing that the R coefficient is characteristic for a given discipline and may take different values. In general, however, the progress of science assumes that $R \leq 1$, and, therefore, the number of true discoveries is smaller than the number of false discoveries.

Table 2. Empirical research – true discoveries

Researcher's verdict	Reality		Sum
	The hypothesis is true	The hypothesis is false	
The hypothesis is true	$c(1-\beta)R/(R+1)$	$c\alpha/(R+1)$	$c(R+\alpha-\beta R)/(R+1)$
The hypothesis is false	$c\beta R/(R+1)$	$c(1-\alpha)/(R+1)$	$c(1-\alpha+\beta R)/(R+1)$
Sum	$cR/(R+1)$	$c/(R+1)$	c

Source: Ioannidis (2005).

On the basis of the above findings, and also assuming that the parameter "c" indicates the number of new discoveries contained in scientific publications, it is possible to indicate the following dependences (Table 2):

A'. $R/(R+1)$ – *A priori* probability that the hypothesis being tested is true[3].
B'. $(1-\beta)$ – The power of the test.
C'. α – Probability of making a type I error.
D'. $c(1-\beta)R/(R+1)$ – Number of scientific discoveries that are true.

[3] If $R = \frac{T}{F}$, then the probability that the given hypothesis is true P(T) is equal to $P(T) = \frac{T}{T+F}$, which after transformation gives the expression $P(T) = \frac{R}{(R+1)}$.

E'. $c\alpha/(R+1)$ – Number of scientific discoveries made that are untrue (type I error).
F'. $c\beta R/(R+1)$ – Number of scientific discoveries overlooked (type II error).
G'. $c(1-\alpha)/(R+1)$ – Number of rejected hypotheses that are in fact false.

From the point of view of the purpose of this article, the most important are the cases of type I error, therefore, stating the existence of a certain relationship where it does not exist. From this point of view, the ratio $\frac{D'}{D'+E'}$ needs to be examined. The formula should be interpreted as the ratio of correctly recognised dependencies to all made (published true and false) scientific discoveries. It is, therefore, an indicator measuring the predictive value of a given discovery. In the following part I will refer to it as the Net Predictive Value (WPN) indicator. We receive:

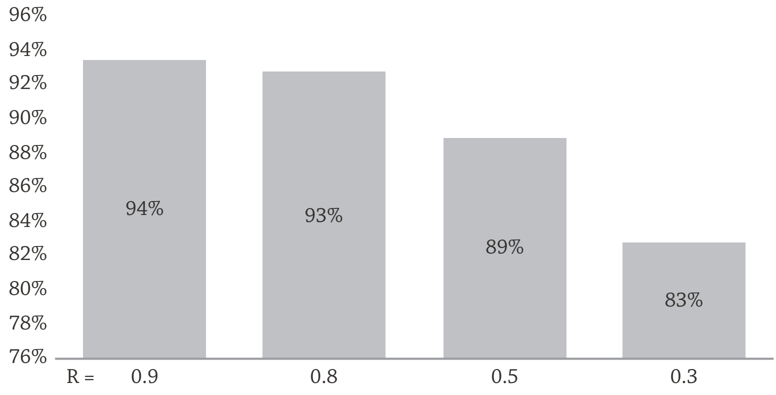
$$WPN = \frac{c(1-\beta)R/(R+1)}{c(R+\alpha-\beta R)/(R+1)}$$

And after transformation:

$$WPN = \frac{(1-\beta)R}{(1-\beta)R+\alpha}$$

The fundamental question to be asked concerns the "typical" value of the WPN indicator in social sciences research. And so, a kind of standard is the assumption that $\alpha = 0.05$ a $\beta = 0.2$. Depending on the value of the R factor adopted, the WPN value is between 94% and 83% (Figure 1).

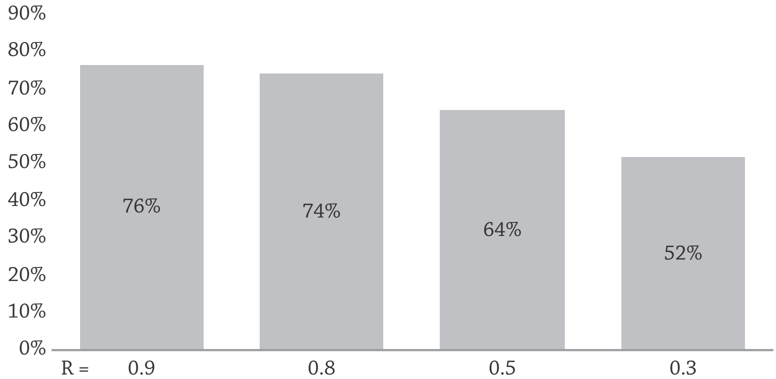
Figure 1. WPN value at different R-levels



Source: own study.

However, it should be noted that the β value is often not verified by researchers. An analysis of 159 meta–analyses of articles in the field of economics leads to the conclusion that the most typical level of statistical power that characterises them does not exceed 18% (Ioannidis et al., 2017). Taking this fact into account means that, depending on the adopted coefficient of R-value, the WPN value oscillates between 76% and 52%.

Figure 2. WPN value at different R-levels ($1 - \beta = 18\%$)



Source: own study.

At this stage of considerations, it can be concluded that the probability that the established statistically significant relationship is true may fluctuate between 95% and 50%. Further in the article there will be circumstances indicated which result in significant reduction of this value.

Problems with science

Scientific uncertainty is a fundamental problem for the decision-maker, because it casts a shadow over the effectiveness of fact-based action and, therefore, on scientific evidence and concepts, which is directly linked to the "loser problem".

However, not only statistical considerations cause the decisions based on scientific evidence to be accompanied with some fundamental, irreducible uncertainty.

Consequently, the decision makers who wish to base their behaviour on knowledge derived from published research face the following difficulties:

a. Science needs time – scientific evidence often appears too late;
b. Researcher's rational thinking – a scientist is most afraid of not what the decision-maker is afraid of;

c. System of incentives and publications – lack of reason can violate the ego of the researcher, lack of publication pushes him into oblivion;
d. Research bias – the research funder expects specific results.

The first point (a) is the relation of time and scientific proof. Scientific progress, in terms of falsificationism, takes place thanks to two factors. (1) Formulating bold scientific hypotheses and (2) falsification thereof. These two parallel processes take place while the time passes which becomes a kind of guarantee that the erroneous hypotheses will be rejected. In this context, it should be noted that almost all kinds of activities are accompanied by various side effects, the existence and intensity of which are difficult to immediately identify. It often takes many years to observe them.

In addition, there may be a situation in which the examined system is unique and it is not possible to create the necessary research sample. As an example, if the hypothesis that excessive carbon dioxide emissions can lead to catastrophic climate changes arises, then it must be stated that its truthfulness is impossible to prove, since there is only one possible test climate, and the historical data available is too scarce to provide clear conclusions. It should be emphasized that we will have objective, empirical evidence only when a climate disaster occurs. Therefore, scientific evidence will appear too late.

This issue becomes particularly relevant in the context of the considered “loser problem”. Historical reflection on cases where the lack of preventive actions was justified by the lack of scientific evidence, or existing scientific evidence suggested that caution is unjustified – leads to the conclusion that the decision maker should always ask questions “what if the theory turns out to be wrong?” and “can the effects of actions based on incorrect theory be worse than not following the recommended ones?”, and often has to make decisions despite the lack of scientific evidence, or even contrary to scientific evidence, if they are not sufficiently reliable (Ahteensuu, 2013; Malinowski, 2018).

There are many examples: margarine, thalidomide, asbestos, glyphosate, x-rays and recently even the issue of mammograms. Until recently, medicine had extensive scientific evidence showing the neutral or even positive effects of these substances/activities on health, but over time, many testimonies have emerged refuting previous findings. Similar difficulties are associated today with proving the dangers of smoking e-cigarettes. They are available on the market relatively recently, so any evidence of their harmfulness will be known in a few decades, when researchers will have an appropriate historical record.

The second issue (b) concerns the researcher’s rational thinking. It should be said that scientific activity is accompanied by a slightly different type of motivation than is the case of making decisions in everyday life. This is important from the point of view of achieving the objective of certain activity.

Let us assume that in the vicinity of a housing estate a company was established – a factory that emits a certain chemical compound “X” into the atmosphere. It should be noted that this compound has never been used on such a large scale before. The residents of this estate are exposed to inhalation of vapours of this specificity, which is why a decision is made to test the substance ‘X’ for effects on human health.

The researchers who have undertaken this action follow the research procedure:

1. They put forward the hypothesis $\mathbf{H_0}$ which states that “substance X is neutral to human body”;
2. They formulate the alternative hypothesis $\mathbf{H_A}$ which declares that “substance X is harmful to human body”.

Then the hypothesis testing process begins, in which the $\mathbf{H_0}$ hypothesis may or may not be rejected. However, it is already known that this procedure is associated with the probability of making a type I and II error.

This is where the crux of the problem appears. The primary goal of researchers is to get to know the truth, not the safety of the inhabitants. These two goals are usually compatible with each other, but in this certain situation that is being under consideration, there is some tension between them. For this reason, they accept a higher probability of type II error. In the analysed example this means that **researchers are most afraid of considering substance “X” dangerous when it is in fact indifferent to health**.

However, the perspective of the residents of the housing estate looks different. Admittedly, they may also highly value scientific truth, but most likely they put their safety first and prefer to be careful. And if so, it means that they are **most afraid of making a type II error (and not type I), i.e., recognising the substance “X” as safe when it is not**. Therefore, that is why they would probably accept a higher probability of type I error in exchange for minimising the probability of type II error.

In this context it seems reasonable to state that, in practice, a researcher uses a slightly different rationality than the person for whom the erroneous result of the study is associated with serious danger. A researcher is a truth seeker, not a safety guard. An acting man usually puts his own or the community’s safety above everything else.

The research procedure characterises an emphasizes on test restrictiveness at the expense of test sensitivity. It is for this reason that weak or early warning signals are considered “statistically insignificant” and only over time it might be proven that the public health was exposed because statistical methods did not identify the risk at an early stage. In addition, if we assume that the typical statistical power of empirical research published by economists is 18%, then the probability of making a type II error turns out to be gigantic, and making decisions based on such research becomes an extremely risky undertaking.

Table 3. Statistical errors – example

Decision:	Actual state of the problem:	
	H_0 is true (substance is neutral)	**H_0 is false** **H_A is true** (substance is harmful)
Rejecting H_0 Acknowledging the substance to be harmful	type I error (*false positive*) Rejecting the zero hypothesis H_0, which is actually true Assuming the alternative hypothesis $\mathbf{H_A}$ to be true, which in fact is false	the right decision
Accepting H_0 Acknowledging the substance to be neutral	the right decision	type II error (*false negative*) Assuming the zero hypothesis H_0 to be true, which in fact is false

Type II error – more serious for residents.
Type I error – more important for researchers.

Source: own study.

Another issue (c) is concerned with the research process itself. The process of designing, performing and interpreting scientific research, which was developed as part of the so-called "scientific method", is one of the greatest achievements of humanity and at the same time a guarantee of objectivity and reliability. Science, being understood in this way, is also a sociological phenomenon, and as such – it can be analysed in terms of credibility or bias of research. Meta-analysis understood in this way allows to capture several problematic aspects of functioning of the broadly understood "scientific environment".

We should start with the fact that the incentives system used to assess scientific achievements of individual researchers places scientific publications at the very top of the hierarchy – what counts is their quantity as well as quality, which is measured by the position on the so-called Philadelphia list. It has been observed that scientific journals are much more likely to publish research that verifies new and original hypotheses than those that analyses and replicates existing hypotheses, concepts, etc. In other words, the magazine caring for its reputation wants to present the latest and most interesting discoveries. However, this approach raises at least one problem. Namely, it generates a strong stimulus for finding patterns and causality even where they do not exist.

As a consequence, a scientist is exposed to strong temptation already at the research design stage. The exploration can be designed in such a way that it is characterised by an increased probability of occurrence of statistically significant type I errors. In this way, the obtained result is original and suitable for publication. And that counts in an environment ruled by the “publish or perish” principle.

The media exemplification of this type of action is a famous provocation, which has gained wide publicity, and that is particularly important in the context of the “loser problem” considered in this text. In 2015, the world press was told that “chocolate helps to lose weight” (Bohannon et al., 2015)[4]. This conclusion was based on research performed by the “Institute of Health and Diet”. The study was conducted on three groups. One group was subjected to a low carbohydrate diet for some period of time, the other group followed the same diet with the difference that each participant had to eat a bar of chocolate every day. The third group – the control group – consisted of people who were supposed to continue their eating habits. The results of the study were extremely interesting, as it turned out that chocolate not only accelerates the process of weight loss, but also has a positive effect on cholesterol levels and life satisfaction.

Nevertheless, the study was mere manipulation. Its author consciously designed the study in such a way that it provided a statistically significant result and gave the impression of reliable science. However, careful analysis of the text raises doubts:

- A small research sample was used (each group of participants consisted of 5 people);
- As many as 18 different variables were verified for each participant in the experiment, including cholesterol levels, blood parameters, sleep quality, life satisfaction, etc.

If a small research sample is assessed against 18 different criteria, then the probability of obtaining a statistically significant type I error is about 60%. Taking into consideration such a small research sample and with so many criteria, it is always easy to find some random correlation.

Although the vast majority of scientific research is conducted in good faith, scientists often unknowingly use the so-called p-hacking[5] tools to obtain a statistically significant result. This is a huge problem, which combined with the aforementioned preference of scientific journals for publishing new and original concepts – causes that, first of all, making decisions based on the results of scientific research can be reduced to acting on the basis of type I errors, and thus leads directly to the “loser problem”. Secondly, it makes one of the fundamental properties of scientific research, namely *repeatability*[6], cease to function. As an example, it has been shown that

[4] This publication was essentially a scientific provocation. The article appeared on the website of the prestigious medical magazine. After revealing the authors of the text – the article was removed.

[5] P – hacking is a practice where the researcher makes a decision about how to analyse data based on the observation of the data itself, not on the design stage of the study.

[6] The question is whether the repetition of a certain scientific study gives the same results.

medical research in the area of cancer is repeatable only in 10% of cases[7] (Begley, 2013), psychological disciplines can boast a repeatability of 33% (Aarts et al., 2015). It is no better in economic sciences, Chang and Li (2015) stated the replicability of macroeconomic research to be at 48%, however, it should be noted that this value applies only to publications for which a set of data was available, and this availability is a separate problem in social science.

I will end this thread with an example of a study that perfectly illustrates the problem of data analysis, statistical errors, preferences of scientific journals and… the uncertainty of science. 29 teams of analysts (61 researchers in total) were provided with the same data and asked to answer the research question: do black-skinned players get red cards more often? The results (1) remarkably varied, (2) twenty teams found a statistically significant, positive relationship between skin colour and the number of red cards received, (3) nine teams did not find a statistically significant relationship, (4) two teams estimated that black-skinned players receive red cards three times more often (Silberzahn et al., 2017). The problem is that despite the fact that the conclusion (4) is a niche view, it is because of its originality, clarity and provocation – nevertheless, it has the greatest chance of publication.

The last point (4) is the issue of research bias. The traditional image of science forces us to perceive it as a process of building knowledge about reality whose flywheel is curiosity of the researcher pushing him to uncompromising confrontation with nature. This vision filled with pathos is very different from the reality of today's science (Ravetz, 2003). First of all, a modern scientist needs advanced and expensive tools to conduct research, which he or she usually cannot afford. There is, therefore, a need to fund research. This means that the research objectives do not come from pure research curiosity, but are determined by various interest groups who care not so much about the objective research procedure as on a specific result. Thus, science out of necessity becomes commercialised which is reflected by the fact that more and more research projects are financed by private companies operating in a given industry, and not by the state. It is worth realising how dangerous this phenomenon is from the point of view of the relationship: researcher – studied reality. Capital groups financing a given project are usually interested in a specific research result correlated with the purpose of their business activity. This creates unavoidable pressure on the researcher. In such conditions, the facts may be bent by researchers who want to generate a result in line with the expectations of the principals, as well as selective financing of the study (depending on the expected outcome), which gives the same effect. As an example, the obvious manipulations of the tobacco industry and the sugar industry can be recalled, which in the 1960s sponsored research proving the

[7] It is worth noting that the analysis concerned 53 studies published by leading scientific journals.

harmlessness of tobacco and sugar, or the recent glyphosate scandal, whose harmfulness was effectively concealed by the Monsanto group.

In natural sciences all manipulations can be relatively easily established and proven. The social sciences present themselves in much worse context. While in natural sciences the financing of a specific study does not guarantee successful results (from the point of view of the sponsoring company), in social sciences the co-financing of a given research project, arising within a specific "economic school", often provides predictable results (Malinowski, 2018).

Credibility of empirical research – even greater problems

In paragraph **3** it is shown that the reliability of empirical research is between 50 and 95%. It is worth emphasizing, however, that this conclusion applies to studies not burdened with some of the problems described in paragraph **4**. Ioannidis (2005) goes a step further and suggests to assume that some part of the research, which would ideally be rejected, goes to publication and becomes part of science.

In other words, he recommends recognising that some research is burdened with bias due to: research design, specifics of hypothesis verification, universally prevailing paradigm in a given discipline, dependence on sponsors, etc. Let us treat "*u*" as false observations that are recognised as real.

Table 4. Empirical research – taking into account the problem of bias

Researcher's verdict	Reality		Sum
	The hypothesis is true	The hypothesis is false	
The hypothesis is true	$(\mathbf{c}[\mathbf{1}-\boldsymbol{\beta}]\mathbf{R}+\mathbf{uc}\boldsymbol{\beta}\mathbf{R})/(\mathbf{R}+\mathbf{1})$	$\mathbf{c}\boldsymbol{\alpha}+\mathbf{uc}(\mathbf{1}-\boldsymbol{\alpha})/(\mathbf{R}+\mathbf{1})$	$\mathbf{c}(\mathbf{R}+\boldsymbol{\alpha}-\boldsymbol{\beta}\mathbf{R}+\mathbf{u}-\mathbf{u}\boldsymbol{\alpha}+\mathbf{u}\boldsymbol{\beta}\mathbf{R})/(\mathbf{R}+\mathbf{1})$
The hypothesis is false	$(1-u)c\beta R/(R+1)$	$(1-u)c(1-\alpha)/(R+1)$	$c(1-u)(1-\alpha+\beta R)/(R+1)$
Sum	$cR/(R+1)$	$c/(R+1)$	c

Source: Ioannidis (2005).

Using comparable calculations to those from paragraph 3 it should be stated that:

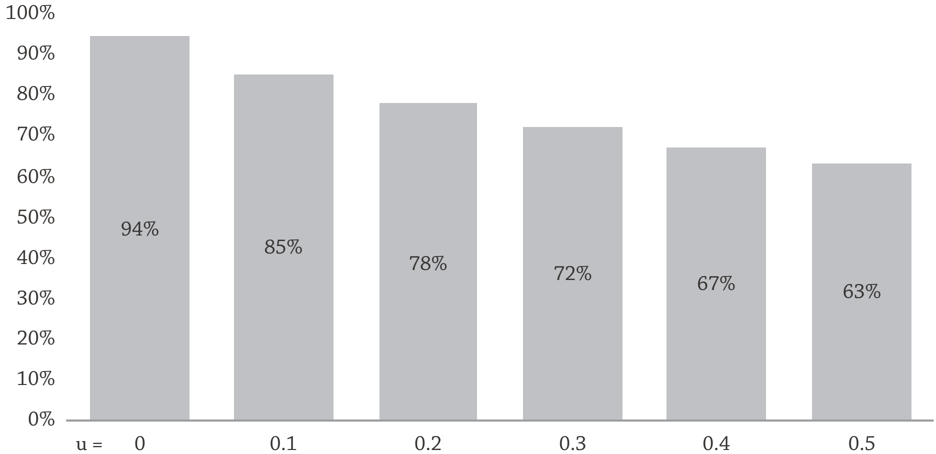
- $\mathbf{c}\boldsymbol{\alpha}+\mathbf{uc}(\mathbf{1}-\boldsymbol{\alpha})/(\mathbf{R}+\mathbf{1})$ – is the number of scientific discoveries that are false due to the probability of making a type I error and due to the bias of research;
- $(\mathbf{c}[\mathbf{1}-\boldsymbol{\beta}]\mathbf{R}+\mathbf{uc}\boldsymbol{\beta}\mathbf{R})/(\mathbf{R}+\mathbf{1})$ – is the number of scientific discoveries made that are true + studies which would be mistakenly considered false, but due to the bias of research were treated as true[8].

[8] Thus, it is a positive "side effect" of research bias.

- $$\mathbf{WPN} = \frac{\left(\mathbf{c}\left[\mathbf{1}-\boldsymbol{\beta}\right]\mathbf{R}+\mathbf{uc}\boldsymbol{\beta}\mathbf{R}\right)/(\mathbf{R}+\mathbf{1})}{\mathbf{c}\left(\mathbf{R}+\boldsymbol{\alpha}-\boldsymbol{\beta}\mathbf{R}+\mathbf{u}-\mathbf{u}\boldsymbol{\alpha}+\mathbf{u}\boldsymbol{\beta}\mathbf{R}\right)/(\mathbf{R}+\mathbf{1})} = \frac{(\mathbf{1}-\boldsymbol{\beta})\mathbf{R}+\mathbf{u}\boldsymbol{\beta}\mathbf{R}}{(\mathbf{1}-\boldsymbol{\beta})\mathbf{R}+\boldsymbol{\alpha}+\mathbf{u}(\boldsymbol{\beta}\mathbf{R}+\mathbf{1}-\boldsymbol{\alpha})}$$

It is critical how the Net Predictive Value (WPN) indicator reacts to changes in the value of the "u" parameter. In an ideal world R = 1, α = 0.05, 1 – β = 0.8 and the phenomenon of research bias does not appear (u = 0). In this reality, the WPN ratio reaches 94%.

Figure 3. WPN indicator (for different levels of the "u" parameter)



Source: own study.

However, if the "u" parameter begins to take higher values than zero (*ceteris paribus*), then the predictive power drops significantly. Probably that "u" parameter takes different values depending on the discipline of the research, but even setting it to 0.1 results in a decrease in the predictive power by 10 percentage points.

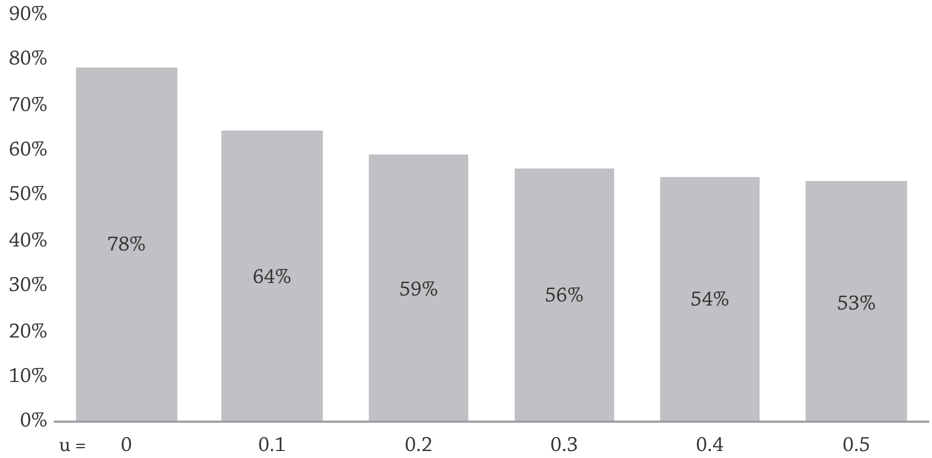
If, nonetheless, in the operation we use the typical value of statistical power characterising publications in the field of economics (1 – β = 18%), then even with a 10% level of bias, the predictive power drops to 78% and to 56% when u = 0.3 (Figure 4).

However, the biggest concern is the combination of 18% statistical power, different values of "u" and "R" (Figure 5), because the predictive power drops significantly and amounts to 78% in the most optimistic scenario and 25% in the most pessimistic scenario.

It should also be noted that the values u = 0.5 and R = 0.3, which in this article are presented as extreme, are not extreme at all. Depending on, for example, the politicisation of studies – the level of the "u" parameter may be much higher than 0.5

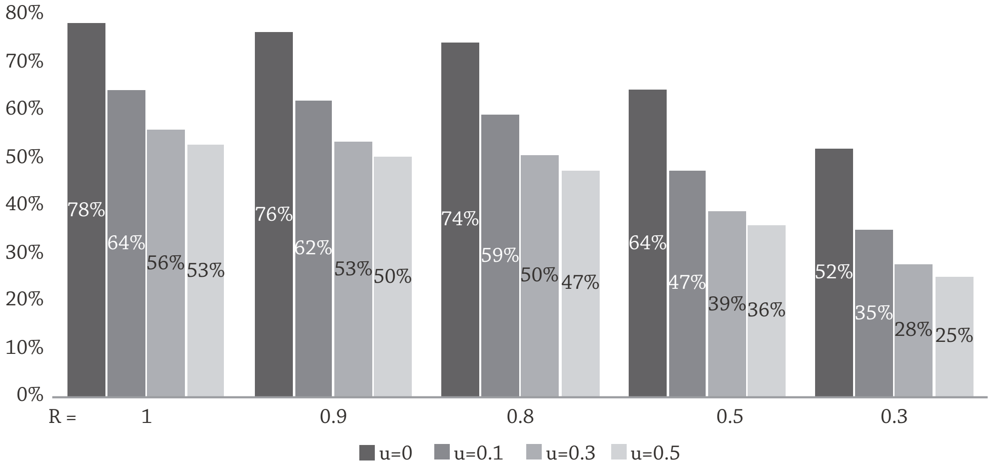
and the "R" factor, in turn, may take very low values (R < 0.1). This usually applies to research in which the context of the discovery plays a particularly important role.

Figure 4. WPN indicator (at different levels of the parameter "u" when 1 – β = 18%)



Source: own study.

Figure 5. WPN indicator (at different levels of the "R" factor and the "u" parameter, when 1 – β = 18%)



Source: own study.

The presented considerations lead to the conclusion that scientific disciplines in which proving is performed using statistical tools are subject to significant level of scientific uncertainty. Gradually, over time – incorrect hypotheses are eliminated, but for a practically oriented entity making decisions based on scientific research can be a very risky undertaking.

Science uncertainty in the context of “path dependence” and “disaster risk”

There would be no “loser problem” if it was not for the unique status of science in the modern world. Unfortunately, uncertainty of science is often not taken into account when making important decisions, especially of an economic/business nature. In other words, decision-makers usually assume that the opinions of experts, which, after all, are based on current scientific research – are objective and dependable, and approach them uncritically, or with insufficient level of criticism. Given the previous findings regarding the predictive power of empirical research, it seems that maintaining particular caution with respect to the credibility of research is by all means more rational.

However, it can be said that it is better to act on the basis of “uncertain science” since the alternative is to refer to premises of a methodologically unclear nature – such as intuition or heuristics. Such an approach, nonetheless, should be confronted with the well-known rhetorical question posed by Nassim Taleb “is it better to go through the forest with a wrong map or without a map?”

In addition, it is worth noting that the “loser problem” gains a special meaning in the context of:

- Dependence on path;
- Disaster risk.

The first phenomenon emphasizes the role of the sequence of events. A person playing roulette can first win a million dollars and then go bankrupt. However, one cannot go bankrupt first and then win a million dollars. This means that the final result of some process (it can be both the economic development of the country and the functioning of the company over time) is not a simple sum of successive partial results, because their order may be (although not necessarily) decisive (Taleb, 2016).

The second term refers to something that can also be labelled as “an unacceptable catastrophe” and, thus, some irreversible systemic destruction. The risk of a catastrophe is unacceptable because its materialisation is associated with irreversible disaster, bankruptcy, failure, etc. It is worth noting that making decisions related to this type of risk significantly changes the criteria for rational behaviour. It turns out that precise, quantified knowledge about the materialisation probability of this type of threat does not change the basis for making decisions. It can even be stated that in case of disaster risk, decision-makers often act as if the probability level of materialisation of catastrophic scenarios would not matter at all (Jablonowski, 2006).

In the context of these two phenomena, the problem of scientific uncertainty takes on a special nature due to the fact that there are certain types of mistakes that the organisation cannot commit, because they are associated with the definitive loss of the game. Normally, this is understood by entrepreneurs whose decisions are ruthlessly verified by the market and are usually the first to bear the consequences of wrong decisions.

An interesting reflection appears at this point, the essence of which differs from the issues of this article, but which is worth further exploration. Namely, it is worth paying attention to the fact that, taking into account the triad: state – enterprises – individuals, at the state level the risk of disaster is an unacceptable one. In other words, the state cannot make an existential mistake because its fundamental goal is survival. Survival is directly related to the avoidance of potentially catastrophic choices. In the long term, permanent, even minimal exposure to catastrophic risk will eventually materialise.

On the contrary, it is quite different with enterprises. Enterprises are affected by "creative destruction", which means that they are forced by the broadly understood society to make very risky decisions, because this is in line with the social interest. As a consequence, however, enterprises have a relatively short lifetime, as they are being constantly replaced by others whose risky decisions brought benefits, and potentially lethal risks have not (yet) materialised. In the context of the "loser problem" it can be concluded that, although entrepreneurs do not want to be losers, they often have to be, because otherwise the other company will make a very careless decision, which as a result of a happy coincidence will not lead to a disaster, but on the contrary – to success and achieving time advantage.

So how to make decisions in conditions of scientific uncertainty? The problem notified in the text is associated with irreducible uncertainty, in the context of which it is difficult to provide universal, methodologically justified recommendations. It seems that the most significant thing in these circumstances is the awareness of the problem of "scientific uncertainty". The decision-making entity must consider in its decision-making algorithms the fact that relying on current scientific research carries some risk, which, despite the estimates presented in this text, is non-quantifiable.

Perhaps a kind of compass in the world of uncertainty is the prudent use of "the wisdom indications" shaped over the centuries, which are essentially an acquired, multi-generational intuition (Gigerenzer, 2014). As an example, we can cite the Chinese maxim around which the Chinese economic development strategy was crowned with such a huge success (Góralczyk, 2018). The aphorism states that one should "walk across the river, feeling the stones under the feet". Interpretation of these words requires paying attention to two elements: (1) awareness of the stated goal and (2) caution. After all, you need to cross the river to get to the other bank, but you should do it gradually, if necessary you need to move to the side or step back, however, constantly examine the bottom you walk with full suspicion. We cannot allow ourselves to make a mistake because it means a disaster.

The idea of making decisions based on uncomplicated, cognitively accessible rules finds theoretical support in the so-called adaptive rationality. The researchers representing this approach prove that in conditions of irreducible uncertainty better decisions are the result of using suboptimal heuristics, not optimal models or algorithms (Potocki, Opolski, 2015).

Conclusion

As part of conclusion, attention should be paid to methodological specificity of management and economic sciences, which has a significant impact on the status of the “loser problem”. Herbert Simon divided the scientific disciplines into theoretical and engineering. The former ones strive to learn the truth, while the latter ones aim at transforming reality. Simon has classified such disciplines as medicine, management or economics in the latter category. Koźmiński and Latusek-Jurczak (2011) using this classification consider what is the relation of theorems arising on the basis of management theory (and thus also economics) to reality. They came to conclusion that the theoretical solutions are certainly not suitable for specific situations. The role of theory is rather to provide descriptions and explanations that contribute to a more in-depth understanding of reality and, in addition, theoretical concepts aim to bring practitioners closer to possible solutions of given problem situations. Therefore, ultimately, Koźminski comes to a conclusion that the primary function of research in economic sciences, in relation to practitioners, is an inspiration.

It is worth noting that the adoption of Koźminski and Latusek-Jurczak’s optics leads to a significant neutralisation of the “loser problem”. This is because the “only” practitioner is inspired by theoretical concepts derived from economic sciences. The findings of these areas do not determine actions, so the question of the final effectiveness of actions is somehow excluded from scientific analysis and left to such factors as intuition or experience. In other words – the scientific convincing power of a given concept/theory does not “force” the practitioner to follow it in the decision-making process.

However, it is hard to resist the impression that nowadays the status of scientific research is more than just “inspiration”. After all, the evidence-based medicine approach widespread in medicine does not leave medical intuition too much space, and the undoubted success of medical science causes that other disciplines implement a similar way of thinking in order to become more "scientific". In this way scientific research, or lack thereof, does not so much inspire, but determines the applicability of a given solution in practice.

It seems that the best illustration of this line of reasoning are the intellectual "inspirations" of economic politicians from recent decades. In the 1980 s, the treatment of privatisation, liberalisation and deregulation became a cure for all economic diseases. In the 90s, it was popular to think about the "end of history" or "dripping theory", a decade later the concept of "turbocapitalism" made a career, and today the so-called "middle income trap" is a famous phenomenon. It should be emphasized that these are not fictional concepts, but very strongly embedded in empirical research products of economic sciences. History shows that almost no economic politician could pass by this type of "winged concepts" indifferently, and the greatest favour of the media and the economic mainstream guaranteed making these ideas the keynotes around which the country's development strategy was built.

References

Aarts, A.A., Anderson, J., Anderson, C. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716-1-aac4716–8.

Ahteensuu, M. (2013). The precautionary principle and the justifiability of three imperatives. *Homo Oeconomicus*, 30(1), 17–36.

Begley, C.G. (2013). Reproducibility: Six red flags for suspect work. *Nature*, 497(7450), 433–434.

Bohannon, J., Koch, D., Homm, P., Driehaus, A. (2015). Chocolate with high cocoa content as a weight – loss accelerator. *International Archives of Medicine*, 8(55).

Chang, A., Li, P. (2015). *Is economic research replicable? Sixty published papers from thirteen journals say "Usually Not"*, Finance and Economics Discussion Series 2015–083. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.083.

Gigerenzer, G. (2014). *Risk savvy*. Penguin Random House.

Góralczyk, B. (2018). *Wielki renesans*. Warsaw: Wydawnictwo Akademickie Dialog.

Ioannidis, J.P.A. (2005). Why most published research findings are false. *PLoS Med*, 2(8): e124, 696–701.

Ioannidis, J.P.A., Stanley, T.D., Doucouliagos, H. (2017). The power of bias in economics research. *The Economic Journal*, 127(10), F236-F265.

Jablonowski, M. (2006). *Precautionary risk management*. New York: Palgrave Macmillan.

Koźmiński, A.K. (2008). *Zarządzanie w warunkach niepewności*. Warsaw: Polish Scientific Publishers PWN.

Koźmiński, A.K., Latusek-Jurczak, D. (2011). *Rozwój teorii organizacji*. Warsaw: Wolters Kluwer.
Malinowski, G.M. (2017). Renesans strategii – czyli o niemożliwości uprawiania evidence based policy. *Organizacja i Zarządzanie*, 1992, 265–284.
Malinowski, G.M. (2018). Zasada ostrożności, czyli heurystyka strachu oraz heurystyka odwagi w kontekście polityki gospodarczej. *Prakseologia*, 160, 291–332.
Lemons, J., Shrader-Frechette, K., Cranor, C. (1997). The precautionary principle: Scientific uncertainty and type I and type II errors. *Foundations of Science*, 2, 207–236.
Lo, Ch. (2009). Risks, scientific uncertainty and the approach of applying precautionary principle. *Medicine and Law*, 28, 283–300.
Potocki, T., Opolski, K. (2015). Decyzje w obliczu „niepewnych ryzyk" – rola heurystyk i nurtu racjonalności adaptacyjnej. *Finanse*, 1(8), 43–70.
Ravetz, J. (2003). The post – normal science of precaution. *Futures*, 36, 347–357.
Silberzahn, R., Uhlmann, E.L., Martin, D. (2017). Many analyst, one dataset: Making transparent how variations in analytical choices affect results. *Advances in Methods and Practices in Psychological Science*, 1(3), 337–356.
Taleb, N. (2016). *The logic of risk taking*. Retrieved from: http://www.fooledbyrandomness.com/rationality. pdf (access: 30.01.2019).
The Economist (2013). *Trouble at the lab*. October 19th, 26–30.
Weiss, Ch. (2003). Expressing scientific uncertainty. *Law, Probability and Risk*, 2, 25–46.

Grzegorz M. Malinowski

Ph.D., Assistant Professor at Kozminski University, economist philosopher. Special interests in policymaking under risk and uncertainty.
e-mail: gmalinowski@kozminski.edu.pl