Method of increasing the objectivity of empirical research

D N Galiahmetova, O A Feoktistova and Ya A Shchenikov
Saint-Petersburg State University of Aerospace Instrumentation, Bolshaya Morskaia str. 67A, Saint-Petersburg, 190000, Russian Federation

E-mail: yar2409@mail.ru

Abstract. The article describes types of cognitive distortions that can affect objectivity of empirical research and empirical data obtained. A self-assessment method for evaluating the experiment objectivity is presented, which can be used to self-assess the objectivity of empirical research. The developed method uses checklists at some stages of empirical research and allows reducing the research subjectivity by minimizing cognitive distortions that may arise due to deviations in the research or defects in the methods used.

1. Introduction
Companies involved in product development and production pay much attention to assessment and description of internal processes in order to improve their quality and objectivity. At the same time, one of the conditions for company successful development in the market is innovation. At the same time, minimal attention is paid to quality of experimental evaluation of innovative products, the quality of so-called “accurate experiments”.

According to a research published in the Harvard Business Review – 40-90% of new products don’t take top of the market. Consumers are usually conservative, and tend to overestimate advantages of an old product by 3 times, while manufacturers tend to overestimate the advantages of an innovative product by 3 times. Innovative products can be overestimated by a manufacturer, for example, the “survivorship bias” – if product test is unsuccessful, then it will most likely not be published and will not be taken into account in a presence of other more successful tests.

In an empirical research, an experimenter can act as both an “on” and a “not on” observer. In the second case, the observer claims to be objective, especially if statistical methods were used to increase the obtained data objectivity. Any empirical research begins with the fact that researcher fixes severity of properties of interest to the object or objects of research, usually with the help of numbers. Thus, it is necessary to distinguish between objects of research (processes, objects, phenomena), their properties – what makes up the subject of research and signs that reflect severity of the properties. At the same time, it is extremely important for the experimenter to be aware that the accuracy with which the characteristic reflects the measured property depends on the measurement procedure. But is it that simple?

Every second human brain processes trillions of different processes. It is not surprising that brain is looking for patterns of behavior that can reduce the number of own processes, thereby saving more energy. However, in scientific activity, these acquired patterns of behavior can adversely affect the objectivity of research results.
Cognitive distortions are errors of thinking that a person makes when processing information. They occur when a person: remembers and recalls, processes a large amount of information, quickly responds to a situation, a person doesn’t have enough information (table 1).

### Table 1. Types of cognitive distortion.

| Context                  | Cognitive distortion                                                                 |
|--------------------------|---------------------------------------------------------------------------------------|
| Remember and recall      | Store memories in different ways depending on the experience situation                |
|                          | Simplify events and lists to individual key points and elements                       |
|                          | We discard particulars for constructing and fixing generalizations                    |
|                          | Edit and enhance memories after events                                               |
| Quick response           | We prefer simple-looking and unambiguous choices to be more complex and uncertain     |
|                          | They tend to maintain personal autonomy and current status in the group avoiding irreversible decisions |
|                          | We prefer to complete what we already invested in time and effort                     |
|                          | We prefer to focus on immediate and close results                                     |
|                          | To act, one must be confident in an ability to change something and feel a importance of one’s actions |
| We process a large amount of information | We more readily notice what we previously remembered or often met Strange/funny/outwardly attractive/ anthropomorphic things attract more attention than usual / unfunny |
|                          | Notice when something has changed                                                    |
|                          | We are attracted to particulars that confirm existing beliefs                        |
|                          | It’s much easier to notice flaws in others than in ourselves                         |
| We lack information      | Discover stories and patterns even in poor data                                       |
|                          | We supplement information gaps with well-known attributes from stereotypes, generalizations and past experience |
|                          | We value more familiar/nice things and people                                        |
|                          | Simplify numbers and probabilities to make them easier to think about                 |
|                          | We think that we know what others think                                               |
|                          | Projecting our current mindset into past and future                                   |

The subjectivity of empirical data acquisition tools was developed by such scientists as: Richard Gregory (the brain automatically completes the image according to his data and does the same with conclusions from incomplete information), R. Rosenthal (an experimenter, who believes in the hypothesis, unconsciously acts in such a way as to obtain results confirming the hypothesis) [1]. The topic was also developed by: Gerald Holton (a scientist was initially set up for what he wants to discover), Michael Polani (the concept of personal implicit knowledge), Mendeleev (measurement accuracy and the importance of metrology), Ernst Mach (data obtained by a person through experience – this is the only reality with which the person deals with, an analogy with the bus schedule), Moritz Schlick (principles of verification, tolerance and physicalism), Edward Thorndike and Phil Rosenzweig (Halo effect), as well as many contemporary authors: S.N. Tits, A.I. Khudyakov, A.D. Nasledov, L.V. Kopets, V.N. Druzhinin, P. Kline [2,3].

An empirical research is designed and implemented by a person – an experimenter, who, of course, can be affected by certain cognitive distortions. The effects encountered in the literature that affect the subjectivity of an empirical research are shown in table 2.

### Table 2. Effects affecting the subjectivity of an empirical research.

| Effect                                  | Demonstration                                                                 |
|-----------------------------------------|-------------------------------------------------------------------------------|
| The Rosenthal effect is a subconscious attempt to adjust their research results to fallacy; tendency to evaluate logical strength of an | Deformation of scale; the Texas sharpshooter fallacy; tendency to evaluate logical strength of an |
the desired theoretical scheme. It can appear in any science at any stage of empirical research.

| Error Type | Description |
|------------|-------------|
| Argument depending on belief in truth or falsity of conclusion; reassessment of value of a certain parameter in the experimental series; biased selection of data that contradicts the hypothesis, so as not to appear biased; sequence effect |
| Halo effect – the result of exposure to general impressions of something on perception of its particular features |
| Distortion in connection with the wording of a scientific law; framing effect – using too narrow approach to describe a situation or question; missing links between explanation and generalization; wrongful appeals to authority or to generally accepted fact; wrongful main cause allocation – a particular reason is taken as main |
| Excessive concern for success |
| Amplification – investing in achieving a goal more effort than necessary (attempt to kill a fly with a sledgehammer); bias towards information search; acceleration; advance |
| Reassessment of particular cases significance |
| Focusing effect; reassessment of a particular variable impact on overall result; deviation towards the result |
| Errors of experiment design |
| Lack of accounting for changes in an research object over time; associative confounding; selection (objects group nonequivalence in composition, which causes a systematic error); factors affecting sample size of objects; wrongful allocation of main cause – a particular reason is taken as main; statistical regression – a group of objects was selected on “extreme” indicators basis; zero risk preference; planning fallacy – the tendency to underestimate resources that need to be spent on tasks |
| Exaggeration of special cases probability |
| Contrast effect; the Baader-Meinhof phenomenon or the illusion of frequency; generalization of special cases |
| Reassessment of one's own opinion/position/choice/ opportunity significance |
| Congruence bias; irrational escalation; the phenomenon of expression of sympathy for the object of research only on the basis of existing acquaintance with it; professional deformation – look too narrow, discarding a more general point of view; retrospective tendency to ascribe positive qualities to an object or action that a person has chosen |

2. Known methods for increasing the objectivity of experimental results

Even in our time, the well-known methods for increasing the objectivity of research are often used only by psychologists, while representatives of other branches of science not only don’t use these methods, but may not be familiar at all with such a formulation of question as the significance of cognitive distortions impact.

Engineers typically consider the following errors: instrument errors, measurements, and processing. As a rule, their geometric mean is taken and they are considered statistically independent. The instrument error is determined by manufacturer and technical condition of the instrument. The measurement error is determined by observer eye. The processing error depends on how complex and multi-stage processing must be done with the data before the result is obtained. If the law is known or assumed, then it can be found from the formula. Thus, the more accurate measuring device experimenter has, the more accurate the results of an empirical research.

In 1994, standard ISO 5725-2:1994 “Accuracy (correctness and precision) of measurement methods and results” [4] was adopted. The standard is about methods for improving accuracy of measurements, about methods for statistical analysis of experimental data to assess precision, balancing results at uniform levels, but does not take into account the presence of cognitive distortions that can lead to experimental errors.
Science community recognizes two criteria on the basis of which it distinguishes scientific knowledge from pseudoscientific:

- The principle of verification proposed by Bertrand Russell – only that knowledge is scientific, which can be confirmed.
- The principle of falsification, proposed by Karl Popper – only that knowledge is scientific, which can be refuted.
- To counter experimental errors, you can also use a number of methodological techniques:
  - The "placebo" method, which allows you to overcome the observer-expectancy effect.
  - The field experiment is an experiment that is conducted under usual conditions of research object existence with minimal intervention by an experimenter in these conditions. The disadvantage of this method is uncontrolled factors that influence the result.
- The use of statistical methods for replicate observations.

3. Self-assessment of researcher’s objectivity in conducting empirical experiments

As mentioned earlier, the cause of experimental errors can be not only a chosen evaluation procedure, but also “cognitive distortions” – common, subconscious, thinking errors that we all systematically make. Such errors are inherent in most people and predictable.

The following is a method to increase the objectivity of experimental results, which helps to take into account the maximum number of known cognitive distortions that affect result of an experiment at all stages of empirical research lifecycle.

4. Method for increasing the objectivity of experimental results

Step 1. Design of experiment. Setting goal of the experiment. Choice of factors and response

After setting goal of the experiment, input and output parameters are identified and selected based on collection and analysis of background information. Factors can be controlled and uncontrolled (random), introducing a systematic or random error in the measurement results: instrumental errors, changes in properties of investigated object during experiment, the impact of personnel, etc.

Step 2. Application of checklist No. 1 and adjustment of the experimental plan

Table 3. Checklist No. 1, devoted to distortions arising during experiment organization.

| Statement 1 | Statement 2 | Distortion example |
|-------------|-------------|--------------------|
| An empirical research is planned in face of a lack of a priori data | An empirical research is planned to be carried out after a detailed research of all factors that can affect the result | Experimenter didn’t study literature enough on the problem under research, did not conduct a survey of experts and specialists related to object of research. Did not conduct preliminary univariate and screening experiments |
| When planning an empirical research, no analysis of noise factors was carried out and their possible impact on the accuracy of experiment was not evaluated | When planning an empirical research, noise factors were analyzed and their possible impact on the accuracy of experiment was evaluated | When studying impact on something different areas of animal brain on its behavior, the very fact of undergoing surgery to remove corresponding area can change the experimental animal behavior |
| Change in the characteristics of a | It was taken into account that over time, | When longitudinal studies of groups of animals should take into |
The object of research over time was not taken into account. Significant changes can occur with a object of research, and their life expectancy needs to be accounted for.

No attention was paid to ensuring the homogeneity of the research groups of objects. The uniformity of the groups of objects of research was ensured. Main and control group of objects should correspond to each other with the exception of those properties that the experiment aims to identify.

Only direct verification methods are included in the experiment. Indirect methods were used additionally together with methods of direct verification. Congruence bias. The subject was presented with two buttons, pressing one of which (left) opens the door. A direct test is a left-click; indirect test – pressing the right button. Since the door remains closed when you press the right button, you can indirectly conclude that you need to press left button.

Such experimental research methods were selected that guarantee a high probability of a favorable research outcome. To solve the problem, various methods were used, including those as a result of which probability of a research favorable outcome was unknown. The ambiguity effect suggests that people tend to choose a solution with known probability of a favorable outcome, compared with a solution where the probability of a favorable outcome is unknown.

It considered a model that was popular in a specific subject field, and not in an interdisciplinary field. The model was considered on the basis of a more general, interdisciplinary point of view.

| Total points | Objectivity  | Corrective actions                                                                 |
|--------------|--------------|-----------------------------------------------------------------------------------|
| 7-18         | Low          | Revision of the experimental design                                                |
|              |              | Selection of cognitive distortions with worst objectivity indicators and experimental design adjustment in order to improve these indicators |
| 19-29        | Medium       | Go to the next step                                                                |
| 30-35        | High         |                                                                                   |

Step 3. Research experiment organization (for example, choice of instruments, choice of personnel)
Clarification of the experiment conditions: instruments, timing, financial resources, executors, etc. Determination of required accuracy of measurement results (output parameters), possible variation of input parameters, types of effects clarification. Type of samples or objects to be studied is selected, taking into account the degree of their compliance with a real product in: condition, structure, shape, size and other characteristics. The accuracy of experimental data also substantially depends on the number of experiments – the more experiments (under the same conditions), the higher reliability of the results. The design of the experiment – the number and order of tests, a method of collecting, storing and documenting data.

Step 4. Conducting an experiment using checklist No.2

| Table 4. The scale for assessing objectivity of an empirical research at planning stage. |
|------------------------------------------|------------------------------------------|------------------------------------------|
| Statement 1                             | Statement 2                             | Distortion example                      |
| We ask you to read 2 statements and mark on the scale which one and how much reflects your activity during the current experiment |

Table 5. Checklist No. 2, devoted to distortions arising during experiment.
It was allowed to transfer the characteristics of the sample to other samples or to the general population.

Statistical methods checked presence in other samples or population selected in a particular case of characteristics.

Based on a positive test result of a small sample of electronic components from a batch, a decision was made about suitability of the entire batch.

The impact on the experimental conditions was minimized after obtaining the experimental data after the first series of experiment.

When conducting a series of experiments, having received “good” and “bad” data in the first series of the experiment, I tried to “improve” the first and “deteriorate” the second in the next series.

The empirical research was conducted in accordance with the originally developed schedule.

“Walking through the minefield” when trying to perform an empirical research as soon as possible, however, attempts to work faster lead to frequent errors.

It was not allowed to influence one of the experimental conditions on the following experiment.

If object of research is a person or an animal, then they can react in different ways to same conditions. If during the experiment researcher changes order of its conduct, then adaptation of the subject to this experiment will take place in different ways.

Underestimation one of the factors significance with statistically proven significance or even its removal from model when a factor is detected whose impact is higher.

Substitution of one concept with another may occur. Researcher may uncritically assume that one situation recorded by him guarantees content of another and report only the last situation.

Table 6. The scale for assessing objectivity of an empirical research at the stage of its implementation.

| Total points | Objectivity | Corrective actions |
|--------------|-------------|--------------------|
| 6-15         | Low         | Conduct repeated experiments using other measuring instruments, experimenters, measurement methods. Increase number of considered factors. Choose measuring instruments with increased accuracy and noise immunity. Use randomization of the tests. Take additional measurements to test hypothesis about type of mathematical model. Increase sample size |
|              |             | Increase the number of experiments, additional measurements. Choose a more complex math model. Repeat experiment with another experimenter. |
| 16-24        | Medium      |                    |
|              |             |                    |
Step 5. Statistical processing of experimental results
Statistical processing of the experimental results is processed with identification and elimination of blunder and systematic errors.

Step 6. Application of checklist No.3 and statistical processing of the survey result

| Table 7. Checklist No. 3, devoted to distortions that occur during processing of results. |
|-----------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| **Statement 1** | **Statement 2** | **Distortion example** |
| Only experimental data were used that confirm the hypothesis of an experimental research | To confirm the hypothesis of the experimental research, all obtained experimental data were used | When processing measurement results procedure of statistical verification of “abnormal” data was not carried out, they were discarded as a “blunder” error |
| Methods convenient for hypothesis confirmation were used | The most objective methods were used to confirm the hypothesis | Using the same data, the use of different statistical methods can lead to different results. It is important not to fall to temptation to use methods convenient for confirming a hypothesis, but to give preference to most objective methods |
| To find the model parameters, only similar data were used and the “outliers” were ignored | To find the model parameters, not only grouping data was used, but also “outliers” | You can shoot at the wall, and only then, in the place where the largest number of holes appeared, draw a target |
| Intuition was used to assess the significance of factors | The significance of factors was estimated using standard statistical techniques | The experimenter neglected the preliminary and screening experiments |
| When analyzing the results, more importance was attached to data obtained after first series of experiments than subsequent | Information obtained after the first series of experiments and subsequent information are of equal value | Primacy effect – the experimenter unconsciously pays more attention to the data obtained after the first series of experiments, focuses on them and based on them builds a certain scheme in his head in advance |
| The tendency to look for common features among the "survivors" and to lose sight of fact that no less important information is hidden among the "dead" | There is no tendency to underestimate data that are not accessible to direct observation from the group of "dead" | Survivorship bias. For example, holes on returning planes show places where they can get damaged and survive, and those who got damaged in other places could not return to base |
| Assessing frequency or possibility of an event by the ease with which these examples or | Frequency or likelihood of event occurrence was estimated by results of processing statistical data | Availability heuristics. For example, a person estimates the risk of heart attack in middle-aged people, recalling similar cases among his friends |
cases are recalled

When analyzing the results of an experimental research, already occurred events were perceived as obvious and predictable

Hindsight bias ("knew-it-all-along" effect). Physicians, as a rule, overestimate their ability to predict outcome of a medical case, stating that they knew the result in advance

With known frequency and specifics of some event there is a tendency to ignore the first event and focus on the second

John wears gothic clothing. How likely is he to be a Christian or a Satanist? If people ask this question, they will underestimate the likelihood that he is a Christian (although there are about 2 billion Christians), and overestimate the likelihood that he is a Satanist (according to statistics several thousand people are Satanists)

The procedures for constructing and verifying the model were carried out using the same experimental data

The experimenter decided to save on experiments and not split the sample into experimental and test

When processing the empirical research, the difference in the type of scales used was not taken into account, too high results of measurements were underestimated and too low ones were overstated

When working with values measured on different scales, certain rules must be followed, otherwise blunders are inevitable. The use of unreasonable methods for determining average values can lead to artificial overstatement or underestimation of average value of a system quality indicator

I emphasized a number of measurements in the experimental series

The value of certain data obtained experimentally wasn’t artificially underestimated / overestimated

Results of the experiment were evaluated on basis of belief in truth of hypothesis

As a sufficient or significant conclusion a statement is proposed that evidence supporting any point of view is an established or generally accepted fact. Instead of validated arguments for evaluating content of hypotheses, a reference is made to opinion of an authority.
Table 8. Scale for assessing the objectivity of an empirical research at the stage of constructing a model of the process/phenomenon.

| Total points | Objectivity | Corrective actions |
|--------------|-------------|--------------------|
| 14-35        | Low         | Redesign the experiment so that the same data is not used to build model and verify its predictive properties. Pay attention to the “anomalies” in the data, try to find the reason for their appearance |
| 36-54        | Medium      | Researched process/phenomenon may be more complex than you expected, and more complex models should be used to describe them |
| 55-70        | High        | Your empirical research is objective enough |

5. Conclusion
This article identifies a problem associated with the lack of a regular practice of taking into account experimenter subjectivity in conducting empirical research. A method is proposed for self-assessment of research objectivity, which includes checklists filled out by experimenter and allowing to reduce subjectivity of the research by minimizing cognitive distortions.

It can be concluded that the method proposed in this article is able in practice to facilitate work of any experimenter who has set himself task of conducting an empirical research that is as objective and free from cognitive distortions as possible. This technique allows you to take into account the most common cognitive distortions and minimize their effect on the results of the experiment.

The problem of the objectivity of empirical research still open, the version of method proposed in this article, of course, is subject to discussion.

As part of future research, further refinement of the scales is necessary, the development of standard norms for a larger number of samples. Prospective approach is the further development of diagnostic capabilities developed checklists in the method.

References
[1] Rosenzweig Phil 2007 The halo effect: the eight other business delusions that deceive managers (NY: Free Press) p 256
[2] Ernst Mach 2012 Knowledge and Error: Sketches on the Psychology of Enquiry (Springer Science & Business Media) p 394
[3] Michael Polanyi 2015 Personal Knowledge: Towards a Post-Critical Philosophy. (Chicago: University of Chicago Press) p 464
[4] ISO/TC69/SC6 Accuracy (trueness and precision) of measurement methods and results – part 2: basic method for the determination of repeatability and reproducibility of a standard measurement method (Moscow: Rosstandart) p 42