

Basic problems of scientific research in psychology

The article discusses the “immersion” of scientific and professional activities (assessment, consulting and psychotherapy) of psychologists from four perspectives: methodological, institutional, social practice and ethical. In scientific research, the primary importance is attached to testable empirical psychological theory. In particular, it has been emphasized that professional practice only makes

sense and is ethical if it is supported by the results of basic research conducted by research psychologists.

KEY WORDS

methodology; ethic; self-reports; study of behaviour; social practice

ORGANIZATION – Faculty of Psychology and Cognitive Science, Adam Mickiewicz University, Poznan, Poland

AUTHORS' CONTRIBUTIONS – A: Study design · B: Data collection · C: Statistical analysis · D: Data interpretation · E: Manuscript preparation · F: Literature search · G: Funds collection

CORRESPONDING AUTHOR – Prof. Jerzy M. Brzeziński, Faculty of Psychology and Cognitive Science, Adam Mickiewicz University, 89 Szamarzewskiego Str., 60-568 Poznan, Poland, e-mail: brzezuam@amu.edu.pl

TO CITE THIS ARTICLE – Brzeziński, J. M. (2020). Basic problems of scientific research in psychology. *Health Psychology Report*, 8(4), 373–384. <https://doi.org/10.5114/hpr.2020.97429>

RECEIVED 09.05.2020 · REVIEWED 12.06.2020 · ACCEPTED 12.06.2020 · PUBLISHED 21.07.2020

SCIENTIFIC RESEARCH IN PSYCHOLOGY IS NOT CONDUCTED IN AN IVORY TOWER

Jerzy M.
Brzeziński

Professional activity of psychologists is conducted in two areas. The first one is scientific research. The second one is professional activity, which includes: (a) diagnostic activity and expert counselling, (b) consulting and psychotherapy, as well as (c) education (primarily by providing the social practice with new psychologists and training professionals who are already on the market) – see Figure 1.

In order to discuss thoroughly the basic problems of psychology, which is defined twofold: firstly, as a scientific discipline within which basic research is carried out, the results of which (above all, new theories and new research instruments) enrich the canon of scientific knowledge, and secondly, as – precisely through the results of basic research – a method of exerting scientifically justified influence on individuals awaiting help, which is provided by professionals (not necessarily by psychologists themselves, but often with their participation), it is important to be aware of the fact that regardless whether it is the basic research itself or the application of its results in the sphere of social practice, applied psychology is subject to various conditions (also of a limiting nature). Metaphorically speaking, the research is not conducted in an *ivory tower*, in isolation from the external world. Its quality is not conditioned solely by internal methodological standards inherent to the methodological social awareness of a community of research psychologists (Brzeziński, 2013, 2016).

Research psychologists face various limitations. These are primarily related to institutional pressure exerted on them. However, a psychologist has to take the psychological nature of a psychological study into account – all in all, this is one of the peculiarities of this research (see: Rosenzweig, 1933; Rosenthal, 1963; Miller, 1972). A psychologist quite often interacts with the participants of his empirical stud-

ies. He may also, not necessarily consciously, alter the behaviour of the respondents. As a result, he will exert subtle unintentional pressure on the behaviour of the respondents, to make it congruent with the assumptions made in his research hypothesis (we are talking about the interpersonal expectancy effect or, simply, the Rosenthal effect). This statement has been confirmed by numerous studies conducted by Rosenthal (e.g., Rosenthal, 1963, 2002; Rosenthal & Jacobson, 1966, 1992) and the continuers of these studies in laboratories worldwide.

Psychologists are also prone to give in to temptations, because an original result may be rewarded with professional promotion and one sometimes lacks patience to check everything thoroughly, conduct at least one replication, before announcing the outcome to the public (e.g. by publishing an article in a prestigious scientific journal).

Similarly, development of a new method (related to counselling or diagnostics) to implement in social practice, based on scientific knowledge in the field of psychology and related sciences and improving the functioning of social practice, is also conditioned by various external non-scientific factors. It is the state of, let us say, unsatisfactory social practice (educational, rehabilitative, related to organizational functioning, etc.) as diagnosed (correctly) by the researchers, which poses a huge challenge for them.

These are the results of basic research conducted by psychologists, which become the scientific foundation for the development of new diagnostic tools and counselling programmes by professionals. In turn, these results of scientific research conducted by research psychologists make up the up-to-date definition of the methodological standard of psychological diagnosis and expert opinions developed by e.g. court-appointed psychologists – *Evidence-Based Assessment* (EBA) – as well as psychological practice – *Evidence-Based Practice in Psychology* (EBPP).

Finally, a psychologist – especially him! – both as a scientist and a professional, has to follow rigorous

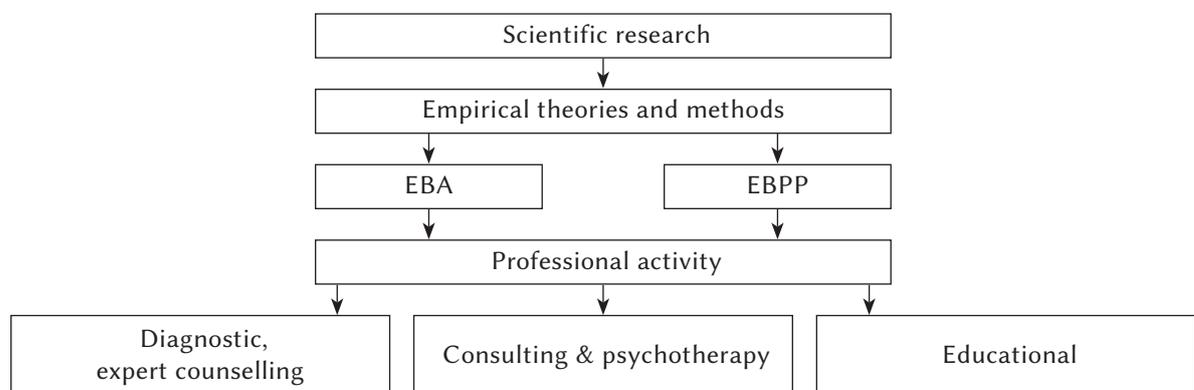


Figure 1. From scientific research to professional activity. Areas of applied psychology.

ethical standards. Both a scientist and a professional have to represent the highest level of ethical awareness.

FOUR PERSPECTIVES

In my opinion, the activity of research and professional psychologists can be discussed from four different perspectives, differentiated on the basis of psychological knowledge:

- Perspective 1: methodological,
- Perspective 2: institutional,
- Perspective 3: social practice,
- Perspective 4: ethical.

Let us now elaborate on peculiarities of these perspectives.

METHODOLOGICAL PERSPECTIVE

It is the level of social methodological awareness in research psychologists (as defined by: Ajdukiewicz, 1974; Kmita, 1991; Brzeziński, 1993, 2016) which conditions the fact that knowledge generated by psychologists has the status of scientific knowledge (let us add: rational, intersubjective knowledge). Its research methods have the status of scientific methods. Only this kind of knowledge can form scientific grounds for professional psychological practice.

TESTABLE EMPIRICAL THEORY, PSYCHOLOGIST!

I do not think that someone would question the thesis that psychology is an empirical science, at least in the light of the main academic approach to conducting scientific research in psychology. The methodological consequences of this assumption were explained well by Hempel (1966, p. 1):

“The different branches of scientific inquiry may be divided into two major groups: the empirical and the nonempirical sciences. The former see to explore, to describe, to explain, and to predict the occurrences in the world we live in. Their states, therefore, **must be checked against the facts of our experience, and they are acceptable only if they are properly supported by empirical evidence**” (boldface by JMB).

The great Popper (2005, p. 7) characterized the tasks of researchers studying empirical sciences (physics, biology, medicine as well as, let us emphasize, psychology) very accurately: “[...] the work of the scientist consists in putting forward and testing theories”.

Obviously, not every idea, even the published one, is worth considering seriously. Developed scientific disciplines (I dare to categorize psychology as such) pose certain methodological requirements of all po-

tential theories (let us add: empirical ones). As I have mentioned above, these requirements are inherent to the methodological social awareness of the community of research psychologists. In turn, the individual methodological awareness of researcher X is a kind of embodiment of that social methodological awareness (Brzeziński, 1993, 2013).

There are three arguments which seem to support the need for theory in the research work of a psychologist.

Argument 1. Theory as a source of definitions of scientific terms (variables) and – in consequence – as the language of problems and hypotheses

Terms which appear in the wording of problems (and hypotheses) have to be defined. This can be done in two ways. One can either refer to the definition which is already present in the literature of the subject and is consistent with the empirical theory accepted by the researcher. An example of this could be the definition of the term “depression”, formulated on the basis of Beck’s cognitive theory of depression. Or it could be the author’s own definition of the given term, developed on the basis of the new, potential theory suggested by the author, which includes this concept.

Argument 2. The basis for giving an empirical meaning to variables (their operationalisation – also: scientific validation of psychological tests)

Psychological tests play a major role in the research and – in particular – diagnostic practice of psychologists. Importantly, the modern approach to the development of tests and interpretation of their results puts emphasis on the significance of the empirical psychological theory and the test validity measurement procedure – see the latest edition of the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014, p. 11):

“Validity refers to the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests. Validity is, therefore, the most fundamental consideration in developing tests and evaluating tests. The process of validation involves accumulating relevant evidence to provide a sound scientific basis for the proposed score interpretations. [...] Evidence of the validity of a given interpretation of test scores for a specified use is a necessary condition for the justifiable use of the test”.

This emphasis on the fundamental importance of psychological theory in the process of determining the accuracy of the test began with the publication in 1955 of the famous article by two prominent methodologically and psychometrically oriented psycholo-

*Scientific research
in psychology*

gists, Lee J. Cronbach and Paul Meehl: *Construct validity in psychological tests* (Cronbach & Meehl, 1955). This new method of measuring test validity has become increasingly popular despite the initially reluctant attitude of some psychologists (e.g., Bechtoldt, 1959) and nowadays nobody questions this, in my opinion, milestone in research on validity measurement. After several years, the opinion of specialists on the significance of this procedure is similar to the one expressed by Drew Westen and Robert Rosenthal (2003, p. 608 – there is also a description of two quantitative measures of construct validity):

“Construct validity is one of the most important concepts in all of psychology. It is at the heart of any study in which researchers use a measure as an index of a variable that is not itself directly observable (e.g., intelligence, aggression, working memory). If a psychological test (or, more broadly, a psychological procedure, including an experimental manipulation) lacks construct validity, results obtained using this test or procedure will be difficult to interpret”.

Therefore, such an “empirical” way of determining test validity, as confronting the results of the developed test with an external criterion, also requires theoretical foundations. The point is in the nature of the external criterion applied to the constructed test at the same time or in the near or distant future. If it is a different psychological test (this option is most often used in less theoretically sophisticated studies, because it seems that this is the easiest way to validate a test), then its construct validity should be established by linking it to a specific psychological theory. The problem is the use of psychiatric diagnosis as a criterion due to its not always satisfactory intersubjectivity.

On the other hand, determining the content validity is of utmost importance in the case of questionnaires constructed according to the “theoretical strategy” (a term taken from Bogdan Zawadzki, 2006). However, also in this case a reference to a psychological theory is essential in the procedure of constructing the questionnaire. It is characteristic for this strategy to “[...] make an assumption about the content value of the questionnaire answers” (Zawadzki, 2006, p. 85). In my opinion, only this strategy of developing questionnaires, which are so popular (as a matter of fact, too popular), meets the highest theoretical standards (see critical comments on the use of self-evaluation methods by psychologists: Baumeister, Vohs, & Funder, 2007; Doliński, 2018). I consider this true because this strategy stems from the verified empirical theory, which forms the scientific basis for the test validation process.

As far as the operationalisation procedure – referring to psychological tests – is concerned, a psychologist decides to (1) use a test which is already available and meets the theoretical assumptions, or (2) construct a new test (also: order such a test from

a specialist test laboratory). “Operationalisation” is no longer discussed in terms of Bridgman’s meaning of *operationism* (Bridgman, 1927, p. 5), which is considered passé, but sometimes is referred to in journals and brochures publishing materials about “quasi-test products”. “[...] in general, we mean by a concept nothing more than a set of operations; the concept is synonymous with the corresponding sets of operations”. Operationism used to be popular in psychology some time ago, meeting the methodological standards of research work (Feest, 2005). Its “top” achievement in the field of psychological tests was the caricatural definition of the term “intelligence”, developed by Edwin Boring (1923, p. 35): “Intelligence is what the tests test”.

It is the correctly conducted operationalisation (let us repeat: referring to an empirical psychological theory), which largely determines the following: (1) the quality of a scientific empirical study, and (2) the quality of psychological practice: diagnostic (the EBA standard) and counselling (the EBPP standard).

Argument 3. Framework for psychological interpretation of research results

The result obtained from an empirical study has to be interpreted. The interpretation framework is provided by an empirical theory. The result itself is meaningless. It acquires meaning when we apply a relevant theory to it. For example, what do the results of Aaron Beck’s BDI questionnaire or Paul Costa’s and Robert McCrae’s NEO-FFI questionnaire tell you as such? To understand what these results tell you about individuals, the researcher has to reach for the cognitive theory of depression by Beck or the Five-Factor Model (FFM) personality model by Costa and McCrae. If we do not refer to a theory, then juggling with the acquired numbers, even supported by the most sophisticated statistical procedures, is useless.

WHAT PSYCHOLOGISTS ACTUALLY INVESTIGATE AND WHAT THEY SHOULD INVESTIGATE: SELF-REPORTS AND FINGER MOVEMENTS VERSUS STUDY OF BEHAVIOUR

Today’s university psychologists are moving away from the basic methodological model that was in force a few decades ago: researching people’s behaviour captured in their naturalistic settings, with reference to field experiments, interviews and observations. Let us recall, for example, the classical study conducted by Musefer Sherif and Carolyn Sherif (1953) on group conflict, or the study by Robert Rosenthal and Lenor Jacobson (1966) on the *expectancy effect* in teachers.

These were Baumeister, Vohs and Funder (2007), who drew the attention of the psychologist community (more precisely, that of social psychologists) to the worrying phenomenon of shifting the research approach from the study of the actual behaviour of the subjects in their naturalistic settings to the study of self-reports and finger movements. Their article, entitled provocatively “Psychology as the science of self-reports and finger movements: Whatever happened to actual behaviour?” brought forth a storm in the community.

The social psychologist Dariusz Doliński conducted similar research in Poland and published its outcome in the article entitled, equally provocatively, “Is psychology still a science of behaviour?” (Doliński, 2018 – see also the whole edition of the journal *Social Psychological Bulletin* where that article was discussed).

A further deviation from the study of behaviour in naturalistic settings is the transfer of some psychologists (especially social psychologists or sociologists) to a new, virtual reality that the Internet has created – increasingly invading the minds of social phenomenon researchers. Perhaps – and not at all in the distant reality – psychologists (mainly social) will function in the scientific substitute of *The Matrix* directed by the Wachowski Brothers. Without seeing the real subjects (because they are hidden *On the other side of the mirror...* that is, nowadays, on the other side of the computer screen), the scientists will interpret what they see on their screens: self-reports. Leaving aside the important questions about the possibility of committing abuse on the Internet, I will ask about the most important issue: the external validity of research conducted in this way, the similarity of Internet conditions to living conditions. The defenders of this new approach do not agree with its criticism, because, in their opinion, the Internet is also a “real” environment in which many people (especially the young ones) function. Nevertheless, professional Internet researchers (e.g., Jemielniak, 2020, Chap. 4: *Research ethics*) involved in social studies, draw attention to various threats related to that. Well, the future will show if my (and not just my) concerns were justified.

INSTITUTIONAL PERSPECTIVE

INSTITUTIONAL PRESSURE DESTRUCTIVE FOR SCIENCE

I wrote above that psychologists do not operate in an *ivory tower*. Their activity is regulated by various legal provisions. Unfortunately, the state – not always rightly and reasonably – interferes in the activities of universities (breaking their autonomy) and research institutes. And this must be a cause for concern.

The mechanical use of bibliometric indicators (e.g. the *impact factor*, IF) as part of employee promotion procedures, in order to assess the scientific achievements of individual scientists in place of a reliable peer review of their published research papers, shows how bureaucratization of science destroys its spirit. What is more, it shows young researchers – who are only at the beginning of their professional career – pathological patterns of effective professional functioning: the value of, let’s say, the IF factor is much more important than the content of the article. Participation in the pursuit of points (the famous cry: *publish or perish!*) becomes the point of academic life of young researchers – including psychologists. In Poland this pursuit has taken a pathological form.

The American Society for Cell Biology presented a critical assessment of bibliometric indicators used to evaluate individual achievements of scientists at its conference held in December 2012 in San Francisco. In a declaration issued and signed in 2013 – *San Francisco Declaration on Research Assessment. Putting science into the assessment of research* (DORA, 2013) – which was signed by many scientists (including Nobel Prize winners), university rectors and presidents of scientific organizations, we can read the following:

“Do not use journal-based metrics, such as Journal Impact Factors, as a surrogate measure of the quality of individual research articles, to assess an individual scientist’s contributions, or in hiring, promotion, or funding decisions”.

Finally, one more harsh opinion, formulated by Bruce Alberts, the editor-in-chief of the prestigious *Science*, who wrote in an *Editorial* (Alberts, 2013, p. 787):

“The misuse of the journal impact factor is highly destructive, inviting a gaming of the metric that can bias journals against publishing important papers in fields (such as social sciences and ecology) that are much less cited than others (such as biomedicine). And it wastes the time of scientists by overloading highly cited journals such as *Science* with inappropriate submissions from researchers who are desperate to gain points from their evaluators”.

One conclusion comes to mind. To evaluate the activity of a scientist, you have to read his articles and peer review them, instead of relying on surrogates. And there is no different – in particular, no better – method of doing that!

If we do not want to kill the spirit of the discoverer – especially in young researchers who are just starting their scientific work – then we should part – for good – with the imageries of officials who turned articles and books into countable points in order to make their work easier. I would like to endorse with absolute certainty the opinion of a biologist (Kubiak, 2019, p. 3):

“Parametric evaluation of the work of researchers, based on counting points, destroys science and annoys scientists. Especially the young ones. This

system cuts us off from the essence of our activity, that is, from acting to satisfy our own curiosity. Leonardo da Vinci, Galileo, Descartes, Darwin and probably Einstein would have thrown this system in the bin in a flash, because they would recognize what we do not see or try not to see: **young people who grow up in this system will not know the ethos of science, will not do science to satisfy their curiosity, but will fight for points instead! And it will be an increasingly ruthless fight.** They will find out that their future does not depend on the strength of their intellect, their imagination, their accumulated knowledge, their ability to cooperate, but on abstract points collected for a career” (boldface by JMB).

Unfortunately, psychologists, at least in Poland, have been seduced by this seemingly objective evaluation policy. I think that sooner or later they will regret it bitterly, because people whose achievements will consist of fantastically large sums of points, instead of original, inspiring scientific results, will take the floor. What will they be teaching the young generations, when they finally age? I think it will be the ability to collect points. Our scientific discipline will solely rely on the pen-and-paper method. For who will it find it “worthwhile” to conduct long-term field experiments, which are unprofitable due to a low or even negative “rate of return” calculated in points?

SOCIAL PRACTICE PERSPECTIVE

Psychology is one of those scientific disciplines in whose case society expects that the research results should become applicable in social practice almost immediately. As far as psychology is concerned, the social environment cares the most about various forms of providing support, e.g. psychotherapy or preventive measures. This pressure of the social environment is accurately recognized by various types of imposters, healers, shamans, etc. They are the ones who offer their “services” to naive individuals, including miraculous medicines, charms and removal of curses. Such expectations towards *normal science*¹ are unrealistic. Sometimes quite a long and bumpy road leads from the scientific result, announced in an article or monograph, to the proven method of action provided to a professional. There is no award awaiting us at the end of that road. If we take shortcuts, we are likely to make many mistakes. These, in turn, may lead to abuse and harm caused to the recipients of psychological services. My suggestion of correct (including from the ethical point of view) application of scientific achievements of psychology to improve the functioning of this area of social practice, which is supported by psychology, is presented in the next section.

FROM RESEARCH PROBLEM TO APPLICATION OF RESEARCH FINDINGS IN SOCIAL PRACTICE – FOUR PHASES

In order for social practice to be satisfactorily effective, it has to (under the sanction of its distortion!) draw on the results of scientific research, which is conducted by scientists (both those working in laboratories, ensuring a high level of control of all variables, and those carrying out “methodologically poorer” field research), and by professionals employed in schools, clinics, penitentiaries, workplaces. The latter often signal a worrying decrease in diagnostic practice (its low accuracy or predictive value) or the effectiveness, let us say, of psychotherapeutic effects carried out according to the current patterns, which were also scientifically sanctioned.

In my opinion, the right (i.e. scientifically correct, belonging to normal science and ethical) transmission of the results of scientific research from a laboratory to – for example – a psychological office or a clinic, is a process and consists of four phases:

- Phase 1: Conducting empirical research verifying the research hypothesis;
- Phase 2: Dissemination of research findings;
- Phase 3: Preparing and verifying efficacy of the practical action plan;
- Phase 4: Implementing the practical action plan in social practice.

A few remarks about each of the differentiated phases.

The first phase includes both contexts of scientific research that the philosopher Hans Reichenbach (1938) distinguished years ago: the *context of discovery* and the *context of justification*. When formulating his thesis about their distinctiveness², Reichenbach wrote that the domain of methodologists is the *context of justification*, and that the *context of discovery* should be left to psychologists and sociologists who specialize in the issues of creativity (in this case, the scientific one). However, as the critics of the distinction of both contexts have shown, it would be difficult to indicate this particular place in the research process where the “discovery” of some of the relationships between variables described in the hypothesis ends and the “justification” of this hypothesis begins. It is impossible to maintain, as the positivist methodology wanted, a specific division of research activities into those that belong to the context of discovery and those that belong to the context of justification. It happens that by justifying “something”, we discover “something new” and – in turn – we try to justify it. Instead of talking about the separateness of both contexts, we should rather talk about their unity – which is confirmed by research practice.

Most of the results of empirical research conducted by psychologists in order to verify the hypotheses are subject to the NHST statistical paradigm³

(which we teach our students about). This is related to the problem of choosing the right statistical test for testing H_0 (the null hypothesis), which states that there are no significant differences between, let's say, means of two populations, in opposition to H_A (the alternative hypothesis), which assumes that there is a difference between these means – we are speaking here about the *t-test for two population means*).

Considering the *statistical significance level* (p) as a measure of the importance of the independent variable X for a specific dependent variable Y is an unreliable criterion (this has been proven convincingly by, for example, Jacob Cohen, 1990, 1994). Therefore, in publications (of course, not only psychological ones), the *effect size* value (ES) is always provided (should be provided) in addition to the statistical value of a given test (e.g. t for Student's t test). In the case of Student's t test, the additional values will be: Cohen's d or Hedges' g (see King & Minium, 2003)⁴.

One of the distinct features of scientific cognition is the reproducibility of research results. The internal complexity of the subject of scientific research in psychology, as well as difficulties in controlling interfering variables, makes it impossible to achieve a high level of reproducibility of studies. This was demonstrated by a very complex and extensive research programme conducted by the social psychologist Brian A. Nosek, together with colleagues from various countries (see: Open Science Collaboration, 2015)⁵. The aim was to determine to what extent repetition of the original tests by other researchers would lead to results comparable to the original ones. In a nutshell, this ambitious research project showed that while 97% of the original studies had statistically significant results ($p < .05$), in replication studies this percentage was significantly lower, at only 36%. The analysis of the *effect size* index values showed that only 47% of the index values obtained in the original studies lay within the 95% confidence interval for the values of these indices from the repeated studies.

So, what do we learn from these studies? Essentially, the novelty consists in the too low, alarmingly low level of reproducibility of results in the replicated studies. How can this worrying result be explained? If we assume (and I want to do so) that the studies conducted by Nosek et al. were carried out methodologically correctly (even if we might think that some improvements could have been made), then we have to (!) think about the answer to an extremely difficult question: Why, therefore, could such a low percentage of the studies be replicated precisely enough? The fact that there may also be problems with the repeatability of results in other scientific disciplines does not relieve us from answering this question. This answer does not have to be relevant only to the research practice of psychologists. I think that to some extent it may also apply to the research practice of scientists from other empirical disciplines: social, biological, medical, etc.

A study deserves to be called scientific if it can be repeated obtaining similar results, independent of the study author, by means of reconstructing the original conditions in which the research was conducted (i.e. the characteristics of the people who took part in it, the characteristics of the research situation, apparatus, measuring tools, statistical procedures, etc.). I do not insist on acquiring identical results, if we focus on studies conducted by psychologists, because some margin of error has to be assumed. Research is not conducted under ideal conditions, by ideal researchers who use the ideal apparatus. It is also necessary to accept the consequences resulting from the historical-cultural perspective and the specificity of the developmental period of the examined individuals. This error will probably be smaller in research conducted in chemical laboratories or cell biology, and larger in psychological research involving people who fill in personality questionnaires. As far as psychological research is concerned, it is easier to replicate successfully neurocognitive laboratory tests than studies on educational or clinical psychology. However, we do not expect the error to be too large. The reason is that only such a result is interesting from the perspective of the scientific knowledge accumulation principle, which can be reproduced with sufficient precision in various places of the globe by competent researchers – which is why Nosek et al. use the term *reproducibility* in their article. As far as scientific tools are concerned, psychology more and more often starts to resemble natural science and adopts its methodological patterns. In turn, *reproducibility* of the research results will increase.

A methodological recommendation of the research result is not only the statistical conclusion (level of $p < .05$ and appropriately high value of the *effect size* index), but also the research conclusion (King & Minium, 2003, p. 6). The second conclusion is a logical consequence of the statistical conclusion. But not only. It also refers to the experimental design, the physical conditions under which the test was carried out, characteristics of the respondents, measuring tools, researcher himself, etc. In other words, we must – as noted by King and Minium (2003, p. 6) – take into account that: “[...] a research conclusion is a conclusion about the subject matter”. This in turn means that we have to analyse carefully the external validity (and the factors impeding it) of the research result obtained in an empirical study.

The second phase seems to be easier. After all, the original problem and ingeniously designed research design, reliable empirical research, properly performed statistical analysis of the results, as well as meeting ethical requirements – they all predestine the obtained research result to be presented to the community of psychologists. It is best to do so by publishing the results in a solid (and recognizable) journal.

Third phase. The result of the empirical research published by the researcher, as well as its replications published by other scientists (because the result aroused – with its originality – great interest in the environment), can be used as a scientific basis for building a practical action programme aimed at solving a specific practical dilemma – for example, creating a new psychotherapeutic programme. I would like to point out that the result of the scientific research itself does not have any “therapeutic capabilities”. These can reasonably be expected from a practical action programme constructed by professionals. Such a programme – if it is to be more effective than the solutions applied so far (if such are in use) – must refer to objective scientific knowledge. It must be consistent with the *Evidence-Based Practice in Psychology* (EBPP) standards (see: APA, 2006; Brzeziński, 2016). Similarly, diagnostic or expert practice should remain in accordance with the *Evidence-Based Assessment* (EBA) standards – where *Daubert Standard* criteria are applied (see: Grove & Barden, 1999; Ritzler, Erard, & Pettigrew, 2002; Brzeziński, 2016). Before a programme is implemented in social practice,

it should undergo an empirical test verifying its efficacy. The test should be similar to the one described in the first phase.

In the fourth phase, the practical action programme is disseminated after its efficiency is confirmed. From that moment on, it is also subject to – not necessarily positive – “corrections” made by state institutions that operate under the dictates of the state’s social policy (including health policy – which may be particularly important for psychologists operating in the state segment of medical treatment and health care). Often these adaptations of programmes are dictated by politicians who manipulate funds from central and local budgets. However, the analysis of these peculiarities goes beyond the scope of this article.

A schematic path from the scientific problem to psychological practice is presented in Figure 2.

ETHICAL PERSPECTIVE

Last but not least, we shall discuss the context of ethical constraints (but also of positive recommendations) that are imposed on researchers (especially psychologists) by psychological scientific societies or, more generally, scientific institutions: international and national. Researchers learn some of these limitations in the course of their studies and read about the others in specialist codes of ethics. Moreover, specialist ethics committees at universities make sure that empirical research designed by psychologists does not cause quite broadly defined harm to its participants. We can also observe imposition of gradual limitations on arbitrariness of researchers over the past few decades. They are not allowed to do everything. And this is the way it should be. For example, nowadays the Stanford prison experiment would not be approved by the ethical committee. Two types of ethical violations can be distinguished in connection with the research activities of psychologists: (a) those related to research and publication procedures and (b) those related to improper treatment of study participants by the researcher (psychologist!) who designed and conducted the research.

ETHICALLY REPREHENSIBLE RESEARCH AND PUBLICATION PRACTICE

p-hacking, *HARKing* – *file drawer effect*. Aware of the reluctance of the editors-in-chief to print manuscripts in which it was not possible to reject the null hypothesis at the minimal required level of $p = .05$, the researchers commit two kinds of abuse:

Firstly, they compare various data (leading to different p values) using diverse statistical tests, until they finally find the required value of $p = .05$ (other popular names include *p-fishing* and *data dredging*).

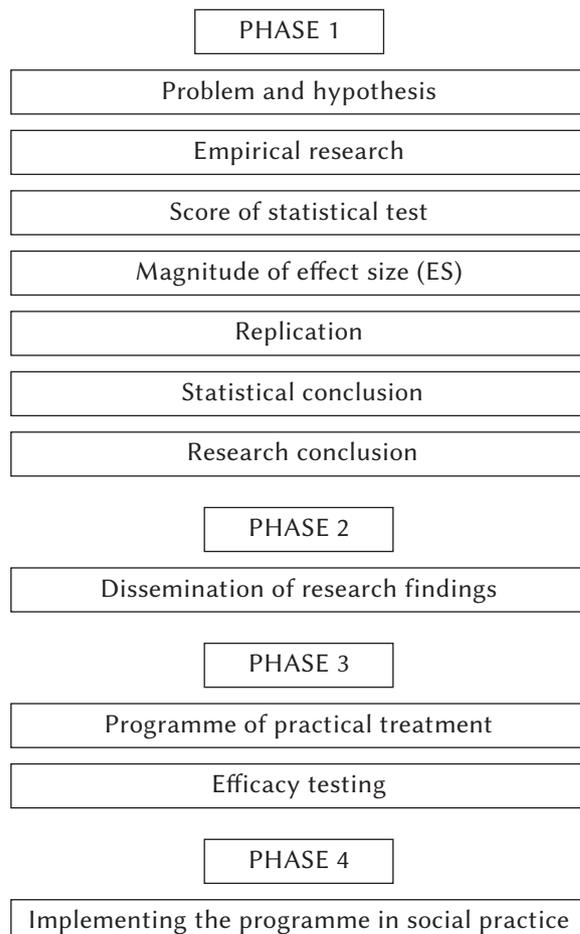


Figure 2. From research problem to application of research findings in social practice.

Secondly, they use a “strategy” in which at first the researcher makes various comparisons, and then, based on the statistically significant ones (at least at the level of $p = .05$), he formulates “hypotheses”, which he then verifies (PHASE 1 of the study – Figure 2), after which he announces a “success” due to the positively verified hypotheses. Today’s computer software available for even the most complex statistical analyses, which is also very easy to use (e.g. SPSS), makes it easier to carry out this abuse. These two types of abuse were described thoroughly by Karolina Karpe (2017).

Fabrication – Falsification – Plagiarism (FFP). According to the legal regulation of the research misconduct phenomenon adopted in the USA (and accepted worldwide), it is defined in the following way (retrieved from <https://ori.hhs.gov/federal-research-misconduct-policy> [accessed April 19, 2020]):

“Research misconduct is defined as fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results. [...] Fabrication is making up data or results and recording or reporting them. Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record. Plagiarism is the appropriation of another person’s ideas, processes, results, or words without giving appropriate credit”.

I have mentioned above the institutional pressure put on staff to deliver as many publications as possible. This makes less resilient researchers subject to this pressure (variously motivated: promotion, fees, fame, fear of losing their jobs, etc.). Paying attention to these misconducts at every opportunity (from the very first year of studies) and blunting them ruthlessly rather than “sweeping them under the carpet” are the best remedies.

Bad publishing practice – publishing in predatory journals. Beginner (but ambitious) young researchers may be tempted to print quickly in an English language journal, which belongs to the fraudulent category of *predatory journals*. Jeffrey Beall used to keep an updated online list of journals in this infamous category until 2017. Now this list of predatory journals is available on the Internet (unfortunately you have to pay for access) at: https://en.wikipedia.org/wiki/Cabell%27s_blacklist (this blacklist comprises more than 10 thousand positions).

UNETHICAL BEHAVIOUR OF SCIENTIFIC RESEARCH PARTICIPANTS

Reprehensible, ethically dubious or even borderline illegal treatment of persons participating in scientific research is – in a nutshell – related to the violation by the researcher (psychologist!) of the four

basic principles (these are inscribed in codes of ethics – e.g., APA, 2017; IUPS, 2008). Let us recall that the psychologist (as a researcher) should take care of: (a) protection of free and informed consent, (b) protection of privacy, (c) protection of confidentiality of personal information, (d) protection of anonymity. However, most importantly, he should respect people’s rights and dignity (an obligation which is derived from one of the most significant documents developed after the World War II, announced in Paris on 10 December 1948 by the United Nations: *The Universal Declaration of Human Rights*).

Experiments which were a shameful example of the violation of ethical principles of dealing with participants in a scientific psychological study were carried out by two prominent psychologists: Stanley Milgram (*the obedience experiments*) and Philip Zimbardo (the Stanford prison experiment). In my opinion, these are not studies worth following.

REMEDIES

What measures can be taken then? I would like to highlight five complementary ones.

The first remedy is to require the replication of research. Only results that can be replicated in other studies, conducted independently by other researchers, have scientific value. The requirement to replicate research also has a “sanitary” meaning. It is a safeguard not only against the dissemination of random results (obtained as an effect of a unique combination of conditions: the respondents, study location, researcher’s expectations, etc.), but – which is particularly important from an ethical point of view – it makes it possible to eliminate from the community of researchers a variety of imposters (in the community of social psychologists, such an infamous “scientist” was Diederik Stapel from Tilburg University, Netherlands, whose practices were disclosed thanks to this “weapon”).

The second remedy is the requirement to disclose (at least to the editorial office of the journal where the article was submitted for printing) raw data in order to be able to carry out a data analysis. I do not accept the statement that the data are the property of the researcher and only he can evaluate them. Such an attitude is unacceptable, especially if the research was financed from public funds (the taxpayer pays for it), and this is the case in the state grant system.

The third remedy is a change in the publishing policy pursued by publishers of scientific journals. Today, journals do not want to publish articles that replicate previously published research results (they reserve the right to publish only original results!), or those that report “negative” results (the hypotheses have not been confirmed). The consequences of this are such that we do not know how many un-

published articles remain in the researchers' drawers because the researcher did not reach the "sanctified" value of $p = .05$ and did not want to "correct" the data. This is why we talk about the *file drawer effect*. A way out of this impasse is a new publishing initiative that some journals have started to follow. It is based on the fact that it is not the ready article, but the concept of the empirical research, which is the subject of evaluation – which is where the term *preregistration of research plan* comes from (Chambers & Munafò, 2013). If it is evaluated positively by the reviewers, then the editorial office assures its author that the results will be published in the journal (regardless of whether "something" was achieved or not).

The fourth remedy is the abandonment of bibliometric assessment of scientific achievements of individual researchers (e.g. via the IF index) in favour of *peer review*. This will also be effective in purifying science *sensu proprio* from junk science.

The fifth remedy is shaping the ethical awareness of researchers (from the first year of psychological studies) regarding the proper treatment of research participants.

SUMMARY

The results of the basic scientific research conducted by psychologists have two aims:

- They enrich the corpus of scientific knowledge with new empirical theories and new methods. As a result, the social methodological awareness of the psychological community is broadened: the existing methods of obtaining results and the applied methods of empirical verification of hypotheses are being replaced by new tools (e.g. years ago planning experiments according to ANOVA and MANOVA statistical models replaced the classical experiment model referring to the statistical model of Student's *t* test or its nonparametric equivalents). As a result, psychology acquires new theories which are justified better. Psychological instruments are also developed.
- They are adopted by social practice, providing it with scientific basis to develop professional practical action programmes and efficient means of testing their efficacy.

Precise acquisition of new results is very important, because in this way we are more certain that the image of reality created by psychologists allows us not only to understand the behaviour of others in different situations and their motives, but also to develop efficient methods of correcting abnormal behaviours. What is more, it enables us to predict accurately (and, in turn, prevent) the results of socially unacceptable behaviours being undertaken by individuals or even groups. The areas of social research and professional practice are closely connected, and

it is hard to imagine – at least in psychology – that it could be different.

Atomisation of research results – lack of syntheses.

ENDNOTES

- 1 The term was derived from Thomas S. Kuhn (1996) from his multiply republished monograph *The structure of scientific revolutions*.
- 2 "[...] I shall introduce the terms context of discovery and context of justification to mark this distinction. Then we have to say that epistemology is only occupied in constructing the context of justification" (Reichenbach, 1938, p. 7). "[...] We emphasized that epistemology cannot be concerned with the first but only with the latter; we showed that the analysis of science is not directed toward actual thinking processes but toward the rational reconstruction of knowledge. It is this determination of the task of epistemology which we must remember if we want to construct a theory of scientific research" (Reichenbach, 1938, p. 382).
- 3 NHST – null hypothesis significance testing.
- 4 These are the recommendations of, for example, the authors of *Publication manual of the American Psychological Association* (APA, 2020).
- 5 B. A. Nosek, professor of psychology at the University of Virginia and co-founder and Executive Director of the Center for Open Science in Charlottesville, re-conducted 100 psychological studies together with 270 researchers from all over the world. The studies were selected from 488 articles published in the three most recognizable psychological scientific journals worldwide: *Psychological Science*, *Journal of Personality and Social Psychology* and *Journal of Experimental Psychology: Learning, Memory, and Cognition*. As far as the scope of the studies is concerned, it included 43 cognitive studies and 57 social-personality studies. The results of the replicated research were analysed thoroughly using statistical tools, with reference to statistical tests of significance, confidence intervals, effect size indicators, correlation coefficients and meta-analyses. The research was conducted between November 2011 and December 2014. All the results are available in the open science system on the websites (in the quoted article there are links to these data sets – including each of 100 replications).

REFERENCES

- Ajdukiewicz, K. (1974). *Pragmatic logic*. Dordrecht, Netherlands: D. Reidel.
- Alberts, B. (2013). Editorial: Impact factor distortions. *Science*, 340, 787. <https://doi.org/10.1126/science.1240319>

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (2014). *Standards for educational and psychological testing (6th ed.)*. Washington, DC: AERA, APA, NCME.
- American Psychological Association, Presidential Task Force on Evidence-Based Practice (2006). Evidence-based practice in psychology. *American Psychologist*, 61, 271–285. <https://doi.org/10.1037/0003-066X.61.4.271>
- APA (2017). *Ethical principles of psychologists and code of conduct*. Washington, DC: American Psychological Association. Retrieved from <http://www.apa.org/ethics/code/ethics-code-2017.pdf> [accessed April 18, 2020].
- APA (2020). *Publication manual of the American Psychological Association (7th ed.)*. Washington, DC: American Psychological Association.
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, 2, 396–403. <https://doi.org/10.1111/j.1745-6916.2007.00051.x>
- Bechtoldt, H. P. (1959). Construct validity: a critique. *American Psychologist*, 14, 619–629. <https://doi.org/10.1037/h0040359>
- Boring, E. (1923). Intelligence as the test tests it. *New Republic*, 36, 35–37.
- Bridgman, P. W. (1927). *The logic of modern physics*. New York: MacMillan.
- Brzeziński, J. (1993). Research process in psychology in the context of the researcher's methodological consciousness. In J. Brzeziński, S. Di Nuovo, T. Marek, & T. Maruszewski (Eds.), *Creativity and consciousness: Philosophical and psychological dimensions* (pp. 155–183). Amsterdam, Atlanta, GA: Rodopi.
- Brzeziński, J. (2013). Methodological awareness and ethical awareness in the context of university education (on the example of psychology). In B. Bokus (Ed.), *Responsibility. A cross-disciplinary perspective* (pp. 261–277). Warszawa: Lexem.
- Brzeziński, J. M. (2016). Towards a comprehensive model of scientific research and professional practice in psychology. *Current Issues in Personality Psychology*, 4, 1–10. <https://doi.org/10.5114/cipp.2016.58442>
- Chambers, C., & Munafò, M. (2013). *Trust in science would be improved by study pre-registration*. Retrieved from <http://www.theguardian.com/science/blog/2013/jun/05/trust-in-science-study-pre-registration> [accessed October 13, 2014].
- Cohen, J. (1990). Things I have learned (so far). *American Psychologist*, 45, 1304–1312. <https://doi.org/10.1037/0003-066X.45.12.1304>
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997–1003.
- Cronbach, L. J., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302. <https://doi.org/10.1037/h0040957>
- Doliński, D. (2018). Is psychology still a science of behaviour? *Social Psychological Bulletin*, 13, 1–14. <https://doi.org/10.5964/spb.v13i2.25025>
- DORA (2013). *San Francisco Declaration on Research Assessment. Putting science into the assessment of research*. Retrieved from www.ascb.org/files/SF-DeclarationFINAL.pdf [accessed April 7, 2020].
- Feest, U. (2005). Operationism in psychology: What the debate is about, what the debate should be about. *Journal of the History of the Behavioral Sciences*, 41, 131–149. <https://doi.org/10.1002/jhbs.20079>
- Grove, W. M., & Barden, R. C. (1999). Protecting the integrity of the legal system: The admissibility of testimony from mental health experts under Daubert/Kumho analyses. *Psychology, Public Policy, and Law*, 5, 224–242. <https://doi.org/10.1037/1076-8971.5.1.224>
- Hempel, C. G. (1966). *Philosophy of natural science*. Upper Saddle River, NJ: Prentice-Hall.
- IUPS, International Union of Psychological Science (2008). *Universal declaration of ethical principles for psychologists*. Retrieved from <https://www.iupsys.net/about/governance/universal-declaration-of-ethical-principles-for-psychologists.html> [accessed April 7, 2020]
- Jemielniak, D. (2020). *Thick big data: Doing digital social sciences*. Oxford: Oxford University Press.
- Karpe, K. (2017). O pogoni za wynikiem istotnym statystycznie. Konsekwencje rozpowszechnienia testowania istotności hipotezy zerowej w psychologii [In chase of statistically significant result. Consequences of widespread use of NHST (null hypothesis significance testing) in psychology]. *Nauka*, 1, 143–156.
- King, B. M., & Minium, E. W. (2003). *Statistical reasoning in psychology and education* (4th ed.). New York: Wiley.
- Kmita, J. (1991). *Essays on the theory of scientific cognition*. Warszawa: PWN – Polish Scientific Publisher; Dordrecht: Kluwer.
- Kubiak, J. (2019). Walka o punkty zabija naukę [The race for „points” is killing science]. *PAUza Akademiacka*, 460, 3. Retrieved from http://pauza.krakow.pl/460_2019.pdf [accessed April 7, 2020]
- Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago, IL: The Chicago University Press.
- Miller, A. G. (ed.). (1972). *The social psychology of psychological research*. New York: The Free Press.
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, 349, aac4716. <https://doi.org/10.1126/science.aac4716>
- Popper, K. (2005). *The logic of scientific discovery* [Adobe Digital Editions version]. Retrieved from <http://strangebeautiful.com/other-texts/popper-logic-scientific-discovery.pdf> [accessed April 7, 2020].

- Reichenbach, H. (1938). *Experience and prediction. An analysis of the foundations and the structure of knowledge*. Chicago, IL: The University of Chicago Press.
- Ritzler, B., Erard, R., & Pettigrew, G. (2002). Protecting the integrity of Rorschach expert witnesses. A reply to Grove and Barden (1999) re: The admissibility of testimony under Daubert/Kumho analyses. *Psychology, Public Policy, and Law*, *8*, 201–215. <https://doi.org/10.1037/1076-8971.8.2.201>
- Rosenthal, R. (1963). On the social psychology of the psychological experiment: The experimenter's hypothesis as unintended determinant of experimental results. *American Scientist*, *51*, 268–283.
- Rosenthal, R. (2002). Covert communication in classrooms, clinics, courtrooms, and cubicles. *American Psychologist*, *57*, 839–849. <https://doi.org/10.1037/0003-066X.57.11.839>
- Rosenthal, R., & Jacobson, L. (1966). Teachers' expectancies: Determinants of pupils' IQ gains. *Psychological Reports*, *19*, 115–118. <https://doi.org/10.2466/pr0.1966.19.1.115>
- Rosenthal, R., & Jacobson, L. (1992). *Pygmalion in the classroom. Teacher expectation and pupils' intellectual development* (expanded edition). New York: Irvington.
- Rosenzweig, S. (1933). The experimental situation as a psychological problem. *Psychological Review*, *40*, 337–354. <https://doi.org/10.1037/h0074916>
- Sherif, M., & Sherif, C. W. (1953). *Groups in harmony and tension*. New York: Harper.
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, *84*, 608–618. <https://doi.org/10.1037/0022-3514.84.3.608>
- Zawadzki, B. (2006). *Kwestionariusze osobowości. Strategie i procedura konstruowania* [Personality questionnaires. Strategies and construction procedure]. Warsaw: Wydawnictwo Naukowe Scholar.