

Measurement problems: contemporary discussions and models

V S Pronskikh

DOI: <https://doi.org/10.3367/UFNe.2019.06.038583>

Contents

1. Introduction	192
2. Accuracy of measurement and its characteristics	192
3. Approaches to measurement	194
3.1 Representational theory; 3.2 Realistic interpretation of measurements; 3.3 Bridgman's operationalism; 3.4 Conventionalism of Mach and Poincaré; 3.5 Information-theoretical approach to measurements; 3.6 Model approaches	
4. Conclusions	199
References	199

Abstract. The review discusses the basic concepts of measurement that have existed from the end of the 19th century to the present day, such as representational theory; realistic, operationalist, conventionalist, theoretical-informational, and theoretical-operational model approaches to measurement; and their advantages and weaknesses. Such characteristics of measurement as accuracy, precision, and various aspects of their application, related to agreement among measurement results, measurements and standards, measurement and ‘true’ values, as well as internal and external reproducibility of measurements, are considered. The theoretical-operational model approach to experiment (and measurement), based on V A Fock’s three-part scheme, proposed by him for quantum mechanics, is discussed in detail using an example from the physics of elementary particles. Advantages of the last approach over the others, as well as its possible applications for determining systematic uncertainty, are examined.

Keywords: measurement, accuracy, precision, reproducibility, theoretical-operational model

1. Introduction

Discussions about what should be considered a measurement and what its main features are have been ongoing since the end of the nineteenth century [1–9]. One of the fundamental quantities, which has long attracted the close attention of researchers interested in the accuracy of various methods for measuring it, is time [10, 11]. A separate set of problems is measurement in quantum-mechanical systems [4, 12–14], along with measurement in complex experiments in elementary particle physics [15, 16]. A consensus among most scholars was reached with the understanding that measure-

ment in simple systems can be operationally reduced to a comparison of some object or process to a connatural standard, whereas in complex cases the measurement must be distinguished in various stages. However, this article will show how, in complex particle measurements, the standard can be formed by the stage of data analysis. The analysis of modern discussions on measurement [9], in which different aspects (or meanings) of measurement are considered, is carried out such that the main traits of measurement and the definition of its key characteristic—accuracy—are highlighted. Notions of measurement that have existed since the end of the nineteenth century are examined, such as the representational theory, which defines measurement as a homomorphism between empirical and numerical systems; the realistic interpretation of measurement based on the idea of the existence of mind-independent ‘true’ values; Bridgman’s operationalism, which reduces measurement to a set of operations; Mach and Poincaré’s conventionalism, which predetermines the measurement result by the conventions of scientists; the information-theoretical approach, which connects the measurement theory to information theory; and the theoretical-operational model. The theoretical-operational model of a complex experiment [15] and its foundations and methods of application for evaluating measurement accuracy are elaborated through an example from the field of particle physics. It will be shown that the theoretical-operational model overcomes the shortcomings of many other models.

2. Accuracy of measurement and its characteristics

In the contemporary scientific literature on measurements, as a rule, two aspects are distinguished that characterize the reliability of the measurement: its accuracy and precision. A measurement result is accurate if it is consistent with some ‘true’ (independently established) or reference value; a measurement is precise if several measurements of the same value are consistent with each other. An example would be arrows hitting a target: the accuracy reflects how close the arrows land to the center of the target (which is set in advance), while the precision reflects how close the arrows

V S Pronskikh. Fermi National Accelerator Laboratory, PO Box 500, Batavia, Illinois, 60510-5011, USA
E-mail: vpnskikh@gmail.com

Received 12 February 2019
Uspekhi Fizicheskikh Nauk 190 (2) 211–219 (2020)
Translated by V S Pronskikh; edited by V L Derbov

land to each other. Indeed, accuracy and precision are different characteristics of a measurement, and a precision measurement will not necessarily be accurate, for example when there is an unknown systematic error and the ‘true’ value is not known in advance.

Contemporary discussions about measurement often distinguish five meanings in which the term ‘measurement accuracy’ is employed in science [9]: metaphysical, epistemic, operational, comparative, and pragmatic. It seems to us that in this case it is more relevant to speak not about the meanings but rather the aspects of accuracy, which are highly dissimilar and can characterize accuracy not only separately but also together, and, therefore, their list can hardly claim to classify accuracy. Let us consider the role and place of these aspects in characterizing the accuracy of measurements.

(1) The metaphysical aspect is agreement between measured and ‘true’ values of a quantity. This definition assumes the existence of some true value; for example, in the correspondent sense, as a correspondence between a value and some mind-independent reality. Such a formulation of the question requires adherence to realism with respect to measurements at the level of their interpretation. The problems of a realistic understanding of measurements and their theory-ladenness are analyzed in Section 3. However, we note that this aspect affects the stage of interpretation of the measurement; more precisely, it affects the interpretation of the standard that the measured value is compared to. In practice, the value obtained from the measurement is compared with the value obtained in another way (and taken as ‘true’).

(2) The epistemic aspect is the agreement between values reasonably attributed to a certain value based on their measurements. Such an understanding is an alternative to that set out in paragraph 1. Here, the interpretation does not require any assumptions about the existence of ‘true’ values but only the fact that there are theoretical grounds to expect that the readings of the device used for measurement will correspond to the measured value; for example, it is believed that the scale of a thermometer, the device, and the method of its use make it possible to reasonably attribute a certain temperature range to the body. This range is determined by the spread, which corresponds to the uncertainty (or error) of the measurement. Measurement error takes into account both statistical and systematic components; if each measurement is accompanied by the same unaccounted-for systematic error, the spread will not characterize measurement accuracy (in its epistemic aspect) but precision (discussed earlier). In the case of an epistemic interpretation of measurement, in contrast to the interpretation given in paragraph 1, measurements are compared operationally with each other rather than with some ‘true’ value that is determined in another way, though in both cases measurements can be characterized by an uncertainty.

(3) The operational aspect is the agreement between the result of the measurement of the quantity and its reference value. Comparison with the standard, as a rule, in physics and technology is considered the definition of measurement; when values are measured, they are compared either directly with standards (length or time) or with scales calibrated using standards (for this, calibration is used, i.e., matching the scales of devices and measurement standards). In our opinion, unlike classification [9], this representation does not reflect one of the meanings of the term ‘measurement’ alternative to those discussed in paragraphs 1 and 2 but is rather a definition of

measurement that is a constituent element of procedures 1 and 2. For example, body temperature is first measured by comparing the height of the mercury column of the thermometer with a temperature scale calibrated by the standard, and then the results obtained in this way are compared in turn with the error by definition or to each other (as in paragraph 2), or with a value obtained by another, for example, calculational, method (as in paragraph 1).

(4) The comparative aspect is the agreement between measurements of the same value obtained by different methods (reproducibility). This reproducibility should be checked under controlled variations of the conditions and measurement methods, because it is precisely such variations that make it possible to identify and take into account systematic errors; accounting for systematic errors allows one to evaluate the accuracy of the measurement but not its precision. At the same time, especially in measurements with complex instruments and at experimental facilities (instrumental complexes), it is important to distinguish the following types of reproducibility: (1) calibration as the reproducibility of measurements of reference phenomena by a device; (2) reproducibility of the measurement result of other known phenomena (as a type of additional calibration); (3) reproducibility when measuring with devices of the same principle of operation (by the same method); (4) reproducibility of the measured value when measuring with devices of different principles of operation (by different methods). Such approaches to control of reproducibility (not only the measurement but the experiment as a whole as the unity of preparation and measurement of phenomena) are found, in particular, in the work of I Hacking [17] (essentially, what is stated in paragraph 2 is an interpretation of his ‘manipulative strategy’) and A Franklin [18, 19] (see a review and discussion of these approaches in [20]). The reproducibility of the measurement of a quantity can be divided into internal (agreement among repeated measurements within the same experimental system) and external (agreement among measurements of the same quantity by instruments based on different principles of operation, i.e., on different instrumental theories). The reproducibility of the measurement should be fulfilled regardless of how the comparison of the value is carried out: with an independently obtained ‘true’ value (as in paragraph 1) or with another measurement (as in paragraph 2).

(5) The pragmatic aspect demands that the measurement accuracy be sufficient for applications in which the measured value is used. Here, the requirement is formulated for the accuracy of the measurement, regardless of how it was determined, in terms of what will satisfy practical requirements. For example, the accuracy of a temperature measurement with a medical thermometer should be good enough to distinguish a healthy person from a sick one.

Although the aforementioned aspects of measurement accuracy can hardly be referred to as the meaning of accuracy, they nevertheless correctly highlight the features that accuracy must possess and are consistent with the definition of both the measurement itself and its accuracy. On this basis, one can assume that, from the operational point of view, measurement (in simple cases) is a procedure of comparison of a quantity (or some characteristic of a body or process) with a congeneric one accepted as a standard. Measurements in complex experimental systems are discussed in Section 3.5. A measurement is characterized by precision, which is the measure of agreement between the result of the measurement and either the results of other

measurements of the same quantity or the ‘true’ value of the same quantity obtained by another method. A measurement of a value must possess both internal and external reproducibility; that is, its values, as obtained via the same method and other measurement methods, must be consistent with each other. The accuracy of the measurement should be sufficient for practical applications of the measurement results. We have given an operational definition of measurement above, but there are several approaches to understanding a measurement; in Section 3, we consider the main ones.

3. Approaches to measurement

3.1 Representational theory

One of the longest-developing theories of measurement is the representational theory [3, 5, 9], which is aimed at studying the mathematical properties of measuring scales and the conditions of their application. This theory follows from the so-called semantic approach to theories developed by P Suppes [21]. According to Suppes, scientific theory can be represented as a hierarchy of models, in other words, structures that satisfy certain axioms. In terms of representational theory, measurement is the construction of morphisms: mappings between empirical and mathematical (numerical) relational systems. An empirical relational system can be represented as a set of empirical objects, such as metal rods, and the qualitative relationships between them, such as ordering and connections. A numerical relational system, according to Suppes, can be a set of real numbers with the operations of relationships between them, such as ‘greater,’ ‘less,’ or ‘equal.’ The measuring scale in this theory is a homomorphism, that is, the mapping of a set of elements onto a single element (sets of empirical objects onto a single numerical one, for example, or sets of bodies of the same length onto a single length scale reading).

3.2 Realistic interpretation of measurements

A realistic interpretation of measurements, common among researchers and engineers, suggests that measurement allows them to establish empirically (experimentally) some objective properties of bodies or processes or their relations. Objectivity is understood to mean that these properties do not depend on the views and beliefs of the person taking the measurements, nor on the methods used. This approach implies that some true quantities exist, and during the measurement their approximate values are estimated [9, 22]. Such understanding is implicitly based on the ideas of logical positivists and empiricists that empirical knowledge is a neutral supplier of facts on which theoretical constructions are grounded, as well as on the provisions of scientific realism, which state that objects, processes, and their properties described by scientific theories exist objectively, regardless of one’s state of mind. The combination of the aforementioned views leads to the expectation that, on the one hand, the measurement values of quantities can be obtained independently of theories of phenomena (which develop on the basis of the use of measured quantities) and, on the other hand, the measured quantities correspond to objective entities and indicate their approximate values.

Adherence to the claim about the existence of mind-independent objects and processes entails a metaphysical understanding of the measurement (discussed in Section 1)

along with the idea of the accuracy of the measurement as an agreement of the result with the ‘true’ meaning. With the realistic approach, the measurement error is also understood in this sense. To argue for their position, realists often point out that, with improvements in measurement technology, the accuracy of measurements increases and that different procedures for measuring the same value lead to consistent results (in our terms, external reproducibility is achieved) [6]. Realists also argue that the theories upon which the action of the instruments is based allow a causal relationship between the quantities and conclude that the existence of the quantities themselves precedes the measurement procedures. From the legacy of logical positivism, the claim is often borrowed about the difference between theoretical and observational (empirical) statements. The arguments for realism in measurement are critically analyzed in Section 3.6.1.

3.3 Bridgman’s operationalism

One of the early measurement concepts, the rise of which is associated with the work of P Bridgman, is operationalism [23]. Bridgman argued that behind any concept or scientific term is a set of operations, and the concept itself is nothing but a sequence of operations. For example, measurements of the length of the same rod with a ruler and an optical reflectometer give two different values, because they correspond to different sets of operations. Bridgman also believed that in specific applications, different measurement operations overlap to form connection points, and the results of such overlaps cannot be blindly trusted but must always be checked experimentally, especially when measurement methods are transferred to a new application. At the same time, Bridgman noted that to the extent that the results of measurements of the same quantity are consistent within the margin of uncertainty, the distinction between measurements obtained by different sets of operations is not necessary.

Operationalism implicitly relies on the empirical assumption that the operations underlying the measurements are independent of theories and thus make the measurement a supplier of measurement facts that precede the formulation of the theory; it also supports the belief of logical positivists that it is possible to independently verify theories by empirical data.

Although the relationship of measurement to the sets of operations is correctly identified by Bridgman, a well-founded thesis about theory-ladenness (see Section 3.2) requires us to put operationalism ‘right side up,’ which was done in [15] and is demonstrated in Section 3.6.2.

At the same time, one can agree [24] that there are no operational definitions for many important theoretical terms.

3.4 Conventionalism of Mach and Poincaré

Another influential trend in measurement theory is conventionalism [1, 2, 9]. The conventionalists argue that, while the result of a measurement cannot be fully reduced to a sequence of operations, any use of quantitative measurement contains an element of convention, in other words, an element of agreement among the parties involved. The conventionalists’ assertions should be understood in such a way that not only the obvious aspects of measurement, such as the choice of units of measurement, for example, metric or imperial systems, but also deeper, more substantial elements of measurement are determined by agreements (different participants in this trend consider different substantive aspects, to be dependent on convention).

One example discussed by Mach was the use of liquid thermometers, which are based on thermal expansion. Mach pointed out that not only do the different liquids used in thermometers expand differently, but the nonlinearities of expansions of different liquids are not the same. Therefore, the choice of a thermometer type determines the temperature measurement result. Poincaré paid attention to the fact that, when scientists choose Euclidean or non-Euclidean geometry to describe a particular system, they do so for the sake of convenience and sometimes even choose unconsciously, guided, for example, by ideas about the simplicity of the description, rather than the properties of natural phenomena. Conventionalism, as well as Bridgman's operationalism, is supported by logical positivists who propose the use of conventional rules of correspondence between theoretical and observational terms, especially where they believe the claims could not be empirically verified, such as a constant rod length during transportation [25], in which case such permanence should simply be postulated for conventional reasons.

Due to conventions, measuring concepts are not constrained to sequences of operations, which is a step forward from operationalism. However, as is evident from the concept of theory-ladenness, the division of terms by logical positivists into terms of theory and observation is not required, and Mach's argument about the different expansion of liquids in thermometers comes down to the claim that the use of thermometers should be based on developed theories of instrument that take into account the differences in expansion. Instruments should be calibrated so that even different thermometers give consistent measurement results (using instruments based on well-corroborated and developed instrumental theories is one of the strategies that researchers should follow [18]).

3.5 Information-theoretical approach to measurements

The basis of the information-theoretical approach to measurements is the resemblance of a measuring instrument to an information-communication system in which the message is encoded in the signal transmitted to the receiver and then decoded. In this approach, the measuring instrument interacts with the object in a certain state and sends a signal to a receiver in a coded form, while the accuracy of measurement depends on both the design of the measuring instrument and the noise (or background) level [7, 27]. This notion of measurement allows the application of the conceptual apparatus of the information theory to the analysis of measurements [26] or even the synthesis of the theory of measurements and information theory [8, 27].

One of the limitations of the application of information theory is that entropy and its uncertainty are insufficient to characterize the error of numerical values of random variables, and information theory in this respect is not sufficiently universal [7]. As with the representational theory described in Section 3.1, the information theory approach is based on the idea of mappings; however, in contrast to information theory, which studies mappings between input and output signals, measurement theory investigates mappings between objects and processes and the numbers characterizing their properties. Another limitation of the information-theoretical approach is that, although from the point of view of information theory the message sent (transmitted information about the measurement) can be known regardless of its transmission with arbitrary accuracy, the state of the measured object or

process cannot be known with arbitrary accuracy and will be determined by the characteristics of the measuring instrument [8].

Another version of the information-theoretical approach was proposed by van Fraassen [28], who distinguished between the physical level of measurement, at which the instrument interacts with the object during the measurement and produces readings, and the abstract level, at which the theory of the phenomenon represents a set of possible states of the object in the state space. The measurement thus limits the range of acceptable parameters of the theory and can be likened to collecting information about the system. Van Fraassen's approach does not require a reference to the concept of information [26]; however, in it, as in other versions of the information-theoretical approach, some empiricist residue can be distinguished. None of these theories is fully based on the concept of theory-ladenness, because dividing the level of the instrument reading and the one at which the theoretical interpretation of the initial signal is formed implies that the level of the device is independent of the interpretation. Such shortcomings of the information-theoretical approach have been overcome in the theoretical-operational model.

3.6 Model approaches

Model approaches to measurement, namely approaches in which the task of measurement is considered to assign numerical values to the parameters of models, as a rule, distinguish the level of specific interactions of the objects or processes under study from the measuring device and the level of theoretical or statistical models of these objects or processes. The models in this case are abstract and local representations of these objects or processes, i.e., mathematical structures constructed using simplifying assumptions [9]. The advantage of model approaches is that, unlike the aforementioned approaches, which are representational or informational, they are not limited to only one aspect of measurement, such as the properties of scales or information transfer, but also explore the process of measurement as a whole, accounting for the design and calibration of devices, sample preparation, error processing, and much more [29]. The shortcomings of the existing model approaches include a lack of clear conceptualizations of the hierarchy of object levels and theories, which is obviously the legacy of logical positivism. These shortcomings are overcome in the theoretical-operational model.

3.6.1 Theory-ladenness of measurement. In discussing the basis for a realistic approach to understanding measurement outlined in Section 3.2, it is important to note that the main arguments of empiricism and logical positivism have been overcome by postpositivism, one of the cornerstones of which is the notion of the theory-ladenness of observation (measurement). In spite of the fact, that since the beginning of the twentieth century, theory and experiment (one of the key elements of which is measurement) have been developing separately [30], the evolution of theory and measurements is iterative, consistently passing through the stages of mutual influence and improvement of both [31]. The fact that without provisional (theoretical) assumptions about the nature and properties of the measured quantity neither measurement nor interpretation of its result would have been possible was indicated by P Duhem [32]. An important historical argument is the one made by T Kuhn, who observed that scientific theories are often accepted by the scientific community long

before they overcome the problems of empirical confirmation of their postulates, and measurement in physics does not serve as a test of a theory but, on the contrary, of its wider application and dissemination until many unexplained deviations appear, which will lead to a crisis of theory, followed by a new scientific revolution [33].

In many of the studies that have appeared since the 1960s, the notion of theory-ladenness has been developed on rich historical grounds. For example, A Franklin carried out a great deal of work with scholarly sources, showing, in particular, that while measuring the charge of the electron, R Millikan began with 175 drops of oil, which he studied in his experiments to determine the value of said charge; however, he used only twenty-three, excluding the rest on the basis of various theoretical considerations. Moreover, in modern experiments in particle physics, the selection and exclusion of data for theoretical necessity are applied directly at the stage of data accumulation and before their receipt for analysis (the measurement of particle properties) [18, 30, 34].

Thus, the rich history of experimental measurements in physics convincingly testifies to the fact that theoretical considerations precede measurements; moreover, theories offer a language in which the results are expressed and interpreted. Hence, the results of measurements are theory-laden or, in other words, based on the language and conceptual machinery of theories that precede the measurements. From this point of view, the realist arguments set out in Section 3.2 are not valid. The fact that measurements are becoming more accurate with the development of measurement methods and techniques can be explained by using the same theory of the phenomenon (this is illustrated in more detail in Section 3.6.2). The causal relationship between values is determined by the theory that not only precedes the measurement but also determines it. In addition, it is clear that the division used by realists, and accepted by logical positivists, into theoretical and empirical languages and terms is inadequate: language and terms can only be theoretical. A schematic way of representation of an experiment (and measurement as one of the stages of the experiment) is presented in Section 3.6.2.

The notion of theory-ladenness was initially met with caution, because it undermined the notion of the independence of the empirical from the theoretical along with the related belief in the primacy and preciousness of the empirical. However, it is necessary to separate the theories into two types that measurement can be laden with. These are, first, theories of the phenomenon under study and, second, theories of instruments (which are also called instrument theories, i.e., theories of measuring instruments). If the ladenness of the first type, i.e., theories of the phenomenon, indeed indicates an epistemic priority of the theory, then the ladenness of the second type is not only unproblematic but may, as was recently shown, serve to improve the measurement. For example, when the author of [11] makes a detailed description of measurement in terms of the instrumental theories on which it is based (like the one we suggested in [15, 16] and analyzed in Section 3.6.2), the ‘de-idealization’ of measurement suggests that understanding the theoretical assumptions made during measurement allows the correction of measurement errors and the assessment of their accuracy and uncertainties more carefully.

A possible problem which may arise during measurements because of their theory-ladenness with instrument theories is that, inasmuch as the measurement task is to check the

theoretical hypothesis, a certain independence of measurement from theory is mandatory. In this case, the best outcome a measurement can claim is to specify (or define) the theory parameters. An even greater freedom will be granted to a measurement if a theoretical model, rather than a fundamental theory, is tested, in which case only the language of measurement interpretation is specified. However, the value of the measurement is significantly compromised if the theory being tested is not separated from the instrumental theory but is included in it. This is a serious problem for a complex modern experiment in particle physics [35], but in some simpler cases it can be solved. For instance, the same problem of a vicious circle can arise if the thermometer, in which the principle of action is based on expansion of bodies (say, mercury), is used for the measurement of the temperature of bodies in an experiment testing a hypothesis about the thermal expansion of bodies [36]. As [36] shows, in such a measurement, a vicious circle will not occur (or can be evaded) if the mercury thermometer in this measurement is calibrated, for example, with a constant-volume gas thermometer (in which the principle of action is premised on the increase in pressure of a constant volume of gas when heated). Thus, the measurement ladenness with an instrument theory is not always an insurmountable problem; moreover, it can often be used to improve the accuracy of measurements, for example, within the theoretical-operational model.

3.6.2 Theoretical-operational model of the experiment. The theoretical-operational model of the experiment is predicated on the three-part scheme of the experiment in quantum mechanics proposed by V A Fock in his argument with Niels Bohr [4, 37]. In [38, 39], it was discovered that a three-part structure similar to Fock’s arises in Galileo’s classical experiments with falling bodies. Then, in [15, 16], this scheme was elaborated and applied to the description of a complex experiment in elementary particle physics, and it was demonstrated that any complex and branched description of an experiment in physics can be expressed in the form of a three-part Fock-type structure. The three-part structure of the experiment, according to Fock, can be schematically represented as follows:

$$\langle P|X(T)|M \rangle, \quad (1)$$

wherein the center is the studied phenomenon $X(T)$, T is the theory of the phenomenon, $\langle P|$ is the operation of preparing the phenomenon (the initial state of the system for measuring), and $|M\rangle$ is the operation of measuring the characteristics of the phenomenon. The scheme involves a certain hierarchy: the theory of the phenomenon is central to it and defines the sequence of operations of preparation and measurement. Of particular importance to Fock’s approach is that it was the first to explicitly state that the preparation of the phenomenon precedes the measurement: “... the measurement itself is only the last stage of setting up the experiment, and this last stage should be preceded by the preparation of the phenomenon itself in its purest form, that is, under precisely defined conditions ...” [4, p. 11].

In certain cases, the theory of the phenomenon may initially be absent (although a model or hypothesis will necessarily take its place), but subsequently the theory will be developed, and such development is the central aim of the experiment. If $X(T)$ is natural phenomena, then the operational parts, the preparation of $\langle P|$ and measurement $|M\rangle$, are

technical, artificial parts. Fock explained his views as follows. The structure of the experiment (not only in quantum mechanics but also in the proposed model and in arbitrary experiments) should be distinguished by three stages: “preparation of the object, the behavior of the object in fixed external conditions and the measurement itself. Accordingly, in the device one can distinguish between the preparation part, the working part and the registering part” [37, p. 468]. Because the technical procedures of preparation ($\langle P \rangle$) and measurement ($\langle M \rangle$) are implemented in devices and their complexes (installations), each of them can be represented as a combination of parallel and sequential operations, and at the heart of each of the operations is the corresponding instrument theory (or their set). The three-part scheme of the experiment (1) can also serve as a form of representation of the experiment in Hacking’s terms: his ‘representation’ and ‘intervention’ can be understood as similar to Fock’s theoretical (T) and operational (first and foremost, $\langle P \rangle$) parts of the system. In addition, the scheme (1) allows the elucidation of the distinction between ‘observation’ and ‘experiment’ by Hacking, because their essential difference is that the experiment unavoidably comprises the stage of preparation of phenomena (systems in certain states), while in the observation (for example, celestial bodies), the preparation of systems by the experimenter is absent (it occurs in nature without human participation) and can be viewed as a measurement guided by theory (the observation scheme, thus, can be simplified to $X(T)|M\rangle$).

Measurement is therefore a part, or stage, of an experiment, which in turn represents the form of relation among the three constituents: phenomenon, its theory, and preparation and measurement operations (laden by instrumental theories). If the measurement in a simple case (rod length measurement) can be presented as a comparison of a value with its reference (see Section 2), then in a more complex case, for example, measuring energy and determining the type of elementary particle, the measurement must be unveiled in a more sophisticated way.

In [15], we examine such a complex measurement as the sequence of operations of indication $\langle I \rangle$ and data analysis $\langle A \rangle$, and the measurement itself $\langle M \rangle$ as $\langle I \rangle * \langle A \rangle$, where the sign $*$ denotes the sequence of operations. The distinction between operations in a measurement is similar to the Maxwellian distinction between ‘indicators’ and ‘scales’ (such as thermometers, dynamometers, or voltmeters). Thus, indication, which is the separation of the object to be measured from the environment, precedes its juxtaposition with the reference. In the theoretical-operational model, it is evident that the very extraction of the measured object or process from its environment requires a theory (or theoretical model) preceding it. For instance, in the case of an elementary particle, its trace must first be visualized (when the analysis is performed visually); in other words, it must be singled out in the bubble chamber of the detector and photographed. This indication results in a photograph of the particle trace in the chamber. The stage of comparison with the standard, in a complex measurement following the indication, is data analysis.

Analysis in the case of particle tracks in the chamber comes down to selecting or creating a model of the possible particle track in the chamber (taking into account the properties of the substance in the chamber and the configuration of the electric and magnetic fields in it) and recognizing the track by adjusting the parameters of the model to best describe a particular track. This often requires a background

model, that is, natural phenomena different from the particles studied, which can also leave outwardly similar traces in the chamber. It seems important to note that the particle track analysis is a fully theory-laden stage and does not allow for empirical interpretation (like, for example, the statement in [40] that the example is ‘theoretically neutral’). It is not independent of the theoretical level, as is suggested in the earlier discussed measurement approaches, including the majority of the model ones; at the same time, it should be agreed with [40] that the ladenness, if properly taken into account, does not represent a technical problem. On the contrary, as discussed in the following, it can be used to increase the overall accuracy of measurements. The reason is that measurement (and data analysis) can be represented by an element of Fock’s three-part scheme, according to which the theory of phenomenon (e.g., the Standard Model) suggests the theory of the preparation of particles and determines the choice of instrumental theories (data indication and analysis) that can be applied in the course of the measurement. Therefore, the track on the photoplate is already predefined by theories of phenomenon and preparation, and at identification, when a particle is set by theoretical considerations, the model parameters can be calculated by fitting the model to the track. For example, the type of particle (from the set predicted by the theory) and the value of its energy (from the range predicted by the theory) can be extracted in this way.

In other words, upon identification (recognition) of a particle, the standard is partially defined by the theories of the device and is partially *formed during the identification process* by the definition of parameters of the model by its adjustment to a particular trace. Such an interpretation of measurement is compatible with the aforementioned suggestions in [28] that in measurement, a limitation of a set of parameters from the space of parameters allowed by the theory occurs, but the essential difference is that we do not have a division into physical and abstract levels; instead, in the theoretical-operational model, there is a level of theories of the phenomenon and a level of operations (preparation and measurement), defined by both instrumental theories and theories of the phenomenon. In addition to [15, 16], the authors of standards and recent studies [41–43] have also arrived at the necessity of differentiation of indication (or counting) in the composition of measurement, which incorporates properties of the measuring instrument in the final state and results, that is, statements about the numerical values of measured quantities, including measurement errors. The theoretical-operational model not only distinguishes these concepts but also allows writing down the inclusion schemes of theoretical components at the stages where these concepts are defined. Section 3.6.3 discusses the practical use and methodological advantages of such schemes, including an example of the application of the theoretical-operational model.

3.6.3 Application of the theoretical-operational model. Scheme of the Gargamelle experiment. The Gargamelle experiment [30, 44] was an accelerator-based experiment in the physics of elementary particles, carried out at CERN, that confirmed the electroweak theory of Glashow–Weinberg–Salam, having discovered weak neutral currents, namely, the process of scattering of antineutrinos on an electron, which in the tested theory was carried out by means of Z^0 -boson exchange. In simple terms, the experiment consisted of the following

operations: accelerated protons collided with an aluminum (beryllium) target, forming charged hadrons; pions moved through an absorber, disintegrating to form muonic antineutrinos or neutrinos; and neutrinos and antineutrinos penetrated the chamber detector and interacted with its working substance, as the theory predicted. Much like other accelerator-based experiments, the problem lay in accounting for the background, that is, phenomena different from those being studied but, in the conditions of the experimental installation, indistinguishable from those under scrutiny. The major sources of neutronic background in the bubble chamber of the Gargamelle experiment were neutrons formed in the magnets and concrete walls by the same antineutrinos (or neutrinos).

Within the framework of the theoretical-operational model based on the three-part Fock setup, the theoretical components of preparation operations $\langle P|$ and measurements $|M\rangle$ can be detailed as follows [15].

Preparation of the phenomenon under study (scattering of antineutrinos on an electron):

$$\langle P| = \{ \langle P1| * \langle P2| * (\langle P3, 1| + \langle P3, 2|) * (\langle P4| + \langle P5|) \}, \quad (2.1)$$

where the character $*$ denotes the sequence and the sign $+$ denotes the parallelism of operations; $\langle P1|$ is preparation of a proton beam of a given energy in the accelerator; $\langle P2|$ is preparation of a stream of pions and kaons in the course of nuclear reactions under the action of a proton beam on a target containing Al and Be nuclei; $\langle P3, 1|$ and $\langle P3, 2|$ are preparation of a stream of muon antineutrinos at the disintegration of pions and muons, respectively; $\langle P4|$ is preparation of Z^0 - and W^\pm -bosons in a bubble chamber under the action of muon antineutrinos on the working substance of the chamber; and $\langle P5|$ is preparation of the neutron background under the action of an antineutrino beam on materials surrounding the chamber.

Proton beam preparation $\langle P1|$ can be presented in more detail as

$$\langle P1| = \langle P1\{dp1((T1)p(E1); p(E2))\}|, \quad (2.2)$$

where $dp1((T1)p(E1); p(E2))$ is the instrument-accelerator. The principle of operation of this device, which converts the initial proton flux with energy $p(E1)$ into a beam with energy $p(E2)$, is based on many theories of T1 (in particular, electrodynamics).

Schematic forms for preparation operations from $\langle P2|$ to $\langle P5|$ are given in [15]. In the measurement operation $|M\rangle = |I\rangle * |A\rangle$ the indication $|I\rangle$ can be represented as follows:

$$|I\rangle = |I\{dm6((T6)e^\pm, \mu^\pm, \text{charged hadrons}; \{\text{photo}_i\})\}|, \quad (2.3)$$

where $dm6((T6)e^\pm, \mu^\pm, \text{charged hadrons}; \{\text{photo}_i\})$ is the device, the bubble chamber, in which the indication of phenomena occurs. The many theories T6 on which the principles of the chamber are based are thermodynamics, mechanics, electrodynamics, and atomic physics, which together describe the formation of traces from bubbles of overheated liquid along the trajectory of the charged particles (electrons, muons, or hadrons). The result of the indication is a photo of the track.

However, as mentioned above, the photo is not yet the final measurement outcome, because the data analysis that

follows obtains the measurement result. First, in quantum field theory (on which the theoretical models tested in the Gargamelle experiment are based), the state is defined by the distribution of probabilities of measured values, which in theory binds to particles with certain properties. Because the measured value in the experiment was the ratio of the number of scatterings that occurred due to Z^0 -boson exchange to the number of scatterings that occurred due to W^\pm -boson exchange, a single measurement could not be sufficient, but it was necessary to analyze a whole set of such events, that is, a set of acts of measurement.

In addition, let us note that the scheme rules out the holistic indistinctness of the phenomenon and measurement, and there is none of Bohr's "totality of phenomena," which would prevent their differentiation [4, 45], although when the device (and its theory) does not possess sufficient selectivity with respect to the background, the theory of the phenomenon can be included as a component of the background, so the possibility of detecting the phenomenon from the background becomes problematic [35], especially in the case of insufficient statistics. In the course of data analysis, the specified particle tracking model is applied to photographs in order to extract the model parameters. In particular, the traces of particles of one type can be wider than the traces of another, and the curvature of the traces depends on the energy of the particle; it should be emphasized, however, that the number of particles that can be detected and the area of their permissible parameters are strictly limited by the theories of the phenomena under study and, consequently, by the methods of preparation and measurement (indication and analysis of data).

Thus, the theoretical-operational model not only allows the theory-ladenness of preparation and measurement of phenomena in the most consistent way but also permits researchers to write down a scheme of the theory-ladenness for an arbitrarily complex experiment. One of the advantages of this schematic representation is that it supports a determination of the factors that affect the accuracy of measurement. The schematic depiction of the experiment clearly unveils theories included in the individual stages of preparation and measurement due to the conveyor-like organization of large experiments; for example, the intensity of the proton beam, measured in some way at the output $\langle P1|$, is transmitted further in the form of a 'black box' to $\langle P2|$. However, while $\langle P1|$ uses only a set of theories T1, there is a high probability that the intensity measurement accuracy defined at stage $\langle P1|$ will take into consideration only the

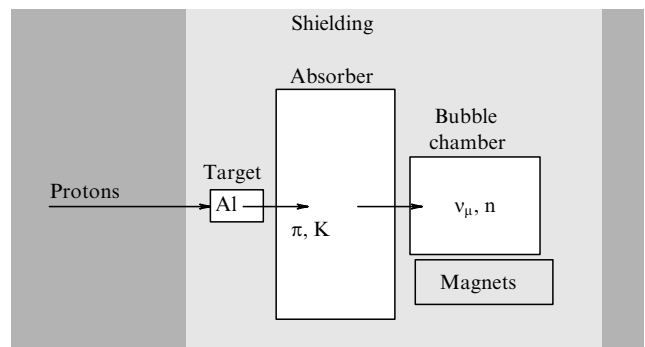


Figure 1. Simplified depiction of the Gargamelle experiment (see figure in [16]).

statistical error, not the systematic one (i.e., the measurement will be precise, but not accurate). Among the known methods for estimating the systematic error is the use of measuring instruments based on other principles of action (instrumental theories) [18].

Hence, the strategy for determining systematic measurement error, which can be obtained from the theoretical-operational model, may be the following:

(1) write out a theoretical scheme of the experiment (or a measurement);

(2) define for each T_n included in the scheme, $T_1 \dots T_m$, which can be used to implement the same $\langle P_n \rangle$ or $|M_n\rangle$;

(3) replace in a series the devices and units, dp and dm based on T_n , with the units and devices dp_1 (dm_1) $\dots dp$ (dm), based on $T_1 \dots T_m$;

(4) repeat the experiment or measurement with replaced instruments and assemblies;

(5) jointly reanalyze the measurement results, determining the accuracy, which will include the systematic error.

This theoretical-operational model can offer for any complex experiment or measurement a methodology for evaluation of its systematic error and accuracy in general.

4. Conclusions

This article seeks to formulate the concepts of measurement and their accuracy and precision, which differ by taking into account or not taking into account systematic error. We have analyzed the properties that measurements should possess, particularly the reproducibility of their results (in which we distinguish the internal and external). This article includes an overview of the major aspects of measurement, such as representational, realistic, operational (according to Bridgman), conventional, information-theoretical, and theoretical-operational. These approaches are not necessarily alternatives to each other, and in some cases they can be complementary; for example, the consideration of measurement as a mapping (homomorphism) of empirical systems on mathematical systems (typical for representation theory) can be premised on the notion of the existence of ‘true’ mind- (and theory)-independent objects and the quantities characterizing them.

It is suggested that a significant disadvantage of most approaches is the lack of consideration of the theory-ladenness of measurements, which, as a rule, in one way or another leads to claims about the independence of the theoretical and empirical levels inherent in logical positivism. The theoretical-operational model, which is discussed in more detail in the preceding sections, is devoid of such shortcomings.

The model represents the structure of the theory-ladenness of the experiment, which is presented as the unity of theories of the phenomenal and the operational parts, preparation, and measurement. The measurement is presented as a stage of the experiment (or observation) consisting of the indication of phenomena along with data analysis, during which a partial formation of the reference occurs. The model is based on a three-part scheme of experiment in quantum mechanics proposed by V A Fock, which we further develop and elaborate for the analysis of an accelerator-based experiment in neutrino physics and which relies on a clear distinction between the phenomenon and the measuring device.

The model allows the deduction of the scheme of theory-ladenness of an arbitrary experiment with theories of the

phenomenon and instrument theories; an example of its application to the description of the Gargamelle experiment in the physics of elementary particles is examined. As in [46], in measurement there is no room for a special role for the observer’s mind. On the grounds of the theoretical-operational model, the methodology of determining the systematic error of measurement in a complex experiment and of improving the accuracy of measurements is suggested; this methodology is based on the identification of instrumental-theoretical components in the experiment and the consistent replacement of the elements laden by them with those premised on other theoretical components, followed by a joint analysis of the results of measurements performed.

Acknowledgements

The Fermi National Accelerator Laboratory is operated by the Fermi Research Alliance, LLC under contract no. DE-AC02-07CH11359 with the US Department of Energy, Office of Science, Office of High Energy Physics.

References

1. Poincaré H “The measure of time” *The Value of Science* (New York: Dover Publ., 1958)
2. Mach E *Principles of the Theory of Heat. Historically and Critically Elucidated* (Dordrecht: D. Reidel, 1986)
3. Suppes P *Portug. Math.* **10** (4) 163 (1951)
4. Fock V A *Usp. Fiz. Nauk* **45** 3 (1951); Fock V A, Einstein A, Podolsky B, Rosen N, Bohr N *Usp. Fiz. Nauk* **16** 436 (1936)
5. Krantz D H et al. *Foundations of Measurement Vol. 1 Additive and Polynomial Representations* (New York: Academic Press, 1971)
6. Swoyer C “The metaphysics of measurement”, in *Measurement, Realism, and Objectivity. Essays on Measurement in the Social and Physical Sciences* (Australasian Studies in History and Philosophy of Science, Vol. 5, Ed. J Forge) (Dordrecht: D. Reidel, 1987)
7. Kirpatovskii S I *Meas. Tech.* **17** 655 (1974); *Izmerit. Tekh.* (5) 11 (1974)
8. Mari E *Measurement* **25** 183 (1999)
9. Tal E “Measurement in science”, in *The Stanford Encyclopedia of Philosophy* (Ed. E N Zalta) (Stanford, CA: Metaphysics Research Lab., Stanford Univ., 2015); <https://plato.stanford.edu/archives/sum2015/entries/measurement-science/>
10. Sheibe A *Usp. Fiz. Nauk* **18** 79 (1937)
11. Tal E *Philos. Sci.* **78** 1082 (2011)
12. Menskii M B *Phys. Usp.* **44** 438 (2001); *Usp. Fiz. Nauk* **171** 459 (2001)
13. Belinsky A V, Lapshin V B *Phys. Usp.* **60** 325 (2017); *Usp. Fiz. Nauk* **187** 349 (2017)
14. Belinskii A V *Phys. Usp.* **46** 877 (2003); *Usp. Fiz. Nauk* **173** 905 (2003)
15. Lipkin A I, Pronskikh V S *Investigated in Russia* 511 (2009); <https://philpapers.org/archive/LIPIOT.pdf>
16. Lipkin A I, Pronskikh V S *Vestn. Ross. Univ. Druzhby Narodov Ser. Filosof.* (3) 56 (2010)
17. Hacking I *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science* (Cambridge: Cambridge Univ. Press, 1983); Translated into Russian: *Predstavlenie i Vmeshatel'stvo. Vvedenie v Filosofiyu Estestvennykh Nauk* (Moscow: Logos, 1998)
18. Franklin A *Shifting Standards: Experiments in Particle Physics in the Twentieth Century* (Pittsburgh, PA: Univ. of Pittsburgh Press, 2013) <https://doi.org/10.2307/j.ctv80c9p7>
19. Franklin A *What Makes a Good Experiment? Reasons and Roles in Science* (Pittsburgh, PA: Univ. of Pittsburgh Press, 2016)
20. Lipkin A I, Pronskikh V S “Filosofiya estestvennonauchnogo eksperimenta” (“Philosophy of experiment in natural sciences”), *Filosofiya Nauki: Uchebnik dlya Magistratury* (Philosophy of Sciences: Handbook for Graduate Schools) (Ed. A I Lipkin) (Moscow: Yurait, 2015) p. 235
21. Suppes P *Synthese* **12** 287 (1960)

22. Trout J D *Measuring the Intentional World: Realism, Naturalism, and Quantitative Methods in the Behavioral Sciences* (New York: Oxford Univ. Press, 1998)
23. Bridgman P W *The Logic of Modern Physics* (New York: Macmillan, 1927)
24. Chang H “Operationalism”, in *The Stanford Encyclopedia of Philosophy* 2009 ed. (Stanford, CA: Stanford Univ., 2009); <https://plato.stanford.edu/archives/win2019/entries/operationalism/>
25. Reichenbach H *The Philosophy of Space and Time* (New York: Dover Publ., 1958)
26. Shannon C E, Weaver W *The Mathematical Theory of Communication* (Urbana: The Univ. of Illinois Press, 1949)
27. Finkelstein L *Kybernetes* **4** (4) 215 (1975)
28. van Fraassen B C *Scientific Representation: Paradoxes of Perspective* (New York: Oxford Univ. Press, 2008)
29. Frigerio A, Giordani A, Mari L *Synthese* **175** 123 (2010)
30. Galison P *How Experiments End* (Chicago: Univ. of Chicago Press, 1987)
31. Chang H *Inventing Temperature: Measurement and Scientific Progress* (Oxford: Oxford Univ. Press, 2004)
32. Duhem P M M *The Aim and Structure of Physical Theory* (New York: Atheneum, 1962); Translated into Russian: *Fizicheskaya Teoriya. Ee Tsel' i Stroenie* (Moscow: KomKniga, 2007)
33. Kuhn T S *Isis* **52** (2) 161 (1961)
34. Pronskikh V S *Filosofiya Nauki* **66** (3) 147 (2015)
35. Pronskikh V S *Filosofiya Nauki* **65** (2) 41 (2015)
36. Franklin A et al. *British J. Philos. Sci.* **40** (2) 229 (1989)
37. Fock V A *Usp. Fiz. Nauk* **62** 461 (1957); *Czech. J. Phys.* **7** 643 (1957)
38. Klyshko D N, Lipkin A I *Elektron. Zh. Issled. Rossii* (53) 736 (2000); <http://zhurnal.ape.relarn.ru/articles/2000/053.pdf>
39. Lipkin A I *Phys. Usp.* **44** 417 (2001); *Usp. Fiz. Nauk* **171** 437 (2001)
40. Franklin A J. *Gen. Philos. Sci.* **46** 155 (2015)
41. Joint Committee for Guides in Metrology, International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM) (Sèvres: JCGM, 2012)
42. Giordani A, Mari L *IEEE Trans. Instrum. Measur.* **61** 2144 (2012)
43. Tal E *Philos. Compass* **8** 1159 (2013)
44. Cundy D, Sutton C *CERN Courier* (2017)
45. Bohr N *Atomic Physics and Human Knowledge* (New York: Wiley, 1958)
46. Zheltikov A M *Phys. Usp.* **61** 1016 (2018); *Usp. Fiz. Nauk* **188** 1119 (2018)