Journal of Scientific Exploration, Vol. 17, No. 2, pp. 243–270, 2003 0892-3310/03

Problems of Reproducibility in Complex Mind-Matter Systems

Harald Atmanspacher

Institut für Grenzgebiete der Psychologie und Psychohygiene
Wilhelmstr. 3a, D–79098 Freiburg, Germany;
Max-Planck-Institut für extraterrestrische Physik
D-85740 Garching, Germany

Robert G. Jahn

Princeton Engineering Anomalies Research
School of Engineering and Applied Science
Princeton University, Princeton NJ 08544-5263
e-mail: rgjahn@princeton.edu

Abstract—Systems exhibiting relationships between mental states and material states, briefly mind-matter systems, offer epistemological and methodological problems exceeding those of systems involving mental states or material states alone. Some of these problems can be addressed by proceeding from standard first-order approaches to more sophisticated second-order approaches. These can illuminate questions of reference and validity, and their ramifications for the topic of reproducibility. For various situations in complex systems it is shown that second-order approaches need to be employed. Considering mind-matter systems as generalized complex systems provides some guidelines for analyzing the problem of reproducibility in such systems from a novel perspective.

Keywords: complexity—meaning—mind-matter systems—reproducibility—second-order thinking

1. Introduction

In a fundamental philosophical sense, the very idea of reproducibility derives from the presumption of ontically given, invariant, ordered structures of nature. In contrast to sense data or introspective data, these structures are assumed to be of universal character. Insofar as their epistemic manifestations are essentially governed by their ontic, invariant origin, any proper acquisition of empirical knowledge (perception, observation, or measurement) about those manifestations should reveal the same aspects of the ontic order. As a consequence, it should be possible to reproduce, at least in principle, manifestations indicative of the same invariant order that are independent of where, when, or by whom the perception, observation, or measurement is carried out.

Given the importance of the concept of reproducibility and the extensive use of the term in the literature, it is astonishing how rarely the term is precisely defined. In many texts, it is more or less assumed that the reader knows what reproducibility is, so it is not easy to find comprehensive or exhaustive characterizations. A compact general account of some basic features of reproducibility is presented in the German *Enzyklopädie Philosophie und Wissenschaftstheorie* (Tetens, 1995):

> Reproducibility means that the process of establishing a fact, or the conditions under which the same fact can be observed, is repeatable. A fact F_2 is a *repetition* of a fact F_1 if only a few designators in the description of F_1 must be replaced to obtain a description of F_2. Therefore it makes sense to speak of the reproducibility of a fact only relative to an explicitly formulated description.
>
> The requirement of reproducibility is one of the basic methodological standards for all sciences claiming lawlike knowledge about their domain of reference. In particular, reproducibility is an inevitable requirement for experiments in the natural sciences: each experiment must be repeatable at any time and at any place by any informed experimentalist in such a way that the experiment takes the same course under the same initial and boundary conditions. The reproducibility of an experiment in the natural sciences includes the reproducibility of experimental setups and measuring instruments . . ..
>
> While the requirement of reproducibility is largely uncontroversial in physics, chemistry, and biology, the possibility of reproducible experiments in behavioral science, psychology, and cultural science is questionable. The reason is that in experiments about the behavior and action of human beings the test persons *know* that an experiment is repeated. This reflexive knowledge changes the initial and boundary conditions of the original experiment in a way which is essential for its outcome. As a consequence, the original experiment cannot be repeated under the same initial and boundary conditions.

The commonly agreed primacy of reproducibility as a pillar of scientific methodology was explicitly challenged by Popper (1965). His main point of criticism was due to a switch of perspective from ahistoric, time-independent knowledge to historically evolving knowledge which crucially depends on the time at which it is acquired. While it most often turns out to be inconsequential to disregard historicity in the natural sciences, its consideration is obviously crucial in many problems of psychological and cultural areas of research.

Therefore it is not astonishing that the main body of literature on reproducibility is found in behavioral science, psychology, and cultural science, where the notion of replicability is often used instead of reproducibility. Some key monographs in these respects are Sidman (1960), Cronbach (1983), and Krathwohl (1985). Specific questions of reproducibility in psychological research have been addressed by Lykken (1968) and Smith (1970).[1]

The quoted characterization of reproducibility addresses a number of crucial issues such as the significance of a theoretical framework (description), the domain of referents, the role of initial and boundary conditions, and different variants of reproducibility in different sciences. These issues are discussed in some detail in Section 2. The epistemological framework for this discussion is based on the Cartesian roots of modern science, leading to the distinction between first-order and second-order schemes of thinking.

In Section 3, the general notions of Section 2 are applied to the field of complex systems. Concepts of second-order complexity are related to particular information theoretical concepts (3.1). Non-stationarity and non-ergodicity are described as typical features of complex systems. Large deviations statistics is suggested as a potential formal framework to deal with such features from a second-order point of view (3.2).

Section 4 extends the approach to complex mind-matter systems, i.e., systems whose description is assumed to require both a mental and a material component to be complete.[2] In this respect, the concepts of meaning and complexity can be understood to refer to mental and material domains, respectively, in a complementary manner (4.1). Several candidates for second-order features of such systems are presented and briefly discussed with respect to their reproducibility (4.2).

Section 5 contains a summary and conclusions.

2. Epistemological Remarks

2.1. First-Order and Second-Order Frameworks

All contemporary sciences share some basic characteristics with respect to their epistemological structure and methodological rules. Generally speaking, they are all based on the fundamental distinction between two domains: some kind of description and its set of referents. Although one can argue about the status of these domains as well as about the distinction itself, it is hard to imagine any serious scientific body of knowledge not embedded within such a scheme.

The specific appearance of the two domains depends on the particular science considered. In the natural sciences, more or less formalized kinds of description are desirable, denoted as theories or models. The set of referents, to which these theories and models refer, typically is a subset of elements of the material world. The historical roots of this framework go back to Descartes' distinction of *res cogitans*, thinking substance, and *res extensa*, extended substance. While the first is the domain of non-material theories and models, the second is the domain of their material referents.[3]

In the social sciences and the humanities, modes of descriptions are usually less formalized, largely due to the more complicated structure and function of most of their referents. Moreover, these referents are not restricted to elements of the material world. In fact, most of the concepts and models of those scientific areas refer to mental, psychological, categorial, conceptual and other elements belonging to the non-material domain of *res cogitans*. In such a situation, it is essential to distinguish between two subsets or levels of elements of the non-material domain, if one wants to avoid problems of self-reference.[4]

If models and conceptions are supposed to be more general than their referents, they are to be located at an upper, second level of description, hence the designation "second-order" description. A description is of second order if its set of referents includes another, first-order description.[5] In more detail, a second-order model or theory (sometimes also called meta-model or meta-theory) refers to second-order facts or data which usually include a first-order framework consisting of a first-order model, first-order data, and their mutual relationship. (Key elements of the reference structure of second-order approaches will be explicated in Section 2.2.) In principle, yet higher-order frameworks are conceivable as well, but we do neither discuss them nor the associated problem of infinite regress in this article.

Natural sciences do not meet problems of self-reference as long as they do not reflect their epistemological status explicitly. Under this condition, it is entirely sufficient to work with first-order approaches. However, there are examples for which the mentioned condition is not satisfied. In other words, there are situations in which a particular mode of description refers to elements of both the non-material and the material domain. A striking example, which will be discussed in Section 3, is the study of complex systems. Moreover, it will be argued that the study of conscious and unconscious aspects of mind-matter systems requires that similar considerations be taken into account.

In the study of complex systems, it has become increasingly clear that the complexity of a system cannot be defined without explicitly referring to the intentions or purposes to be realized by such a definition. This implies that definitions of complexity are not context-free. Among other issues, the concept of meaning becomes crucial in this respect. This situation is particularly interesting insofar as it evolved from within the physical sciences. It represents an example where explicit considerations of epistemological issues arise in specific physical problems.

In the study of mind-matter systems, it is clear that the mind-brain system cannot be understood without considering both non-material mental states and material brain states together. This is the arena in which sciences such as neuropsychology and neurophysiology meet. The corresponding problem of psychophysiological relations can be extended to mind-body relations and, more generally, to mind-matter relations beyond the boundaries set by individual living beings, e.g., particular aspects of their collective behavior.

All these examples presume that an appropriate material system (e.g., a central nervous system or brain) is a *necessary* condition for a functioning mental system. Relaxing this condition leads to broader questions of consciousness, e.g., questions as to the existence of "mentality" independent of a material body, or, from a different perspective, questions as to the relationship between mental and material systems which are not "bound together" within an individual living being. This broad scope pertains to the psychophysical problem or mind-matter problem in its most general sense. Particular aspects of this will be discussed in Section 4.2.

Since in both the study of complexity and of consciousness, including mind-matter research in its most general sense, models are required to refer to both material and non-material elements, they are paradigmatic cases for the necessity of second-order descriptions. It is obvious that this necessity can complicate the standard scientific methodology considerably. In the following, some of the ramifications implied by these complications will be discussed.

2.2. *Reference*

Standard, first-order approaches usually operate with clear-cut reference relations between elements of a model or theory and elements of the system to be described. Each theoretical term in a physical theory is defined in such a way that it refers as uniquely as possible to an equivalence class of elements of the material world. To give some oversimplified examples, the laws of classical mechanics refer to the behavior of solid bodies, the laws of hydrodynamics refer to the behavior of liquids, the laws of electrodynamics refer to the behavior of substances with electric and magnetic properties, and so on.

It is important to realize that the reference relations themselves cannot be explicitly investigated in terms of first-order approaches. Second-order approaches are needed to provide a reference structure more sophisticated than that of first-order approaches. In principle, the entire framework of first-order approaches can become the referent of a second-order approach. This is to say that a first-order model plus its first-order referent(s) plus the relation between the two can become the second-order referent of a second-order model (see Figure 1).

Although suggestive, it is not a triviality that first-order models and first-order referents can be decomposed with respect to each other. It is possible that the relation between them is of such a kind that, for particular issues, it is even illegitimate to conceive them as decomposable. This possible difficulty notwithstanding, the success of scientific practice seems to indicate that such a decomposition is at least very useful in many situations. The methodology of sciences such as physics, for instance, relies essentially on the assumption that first-order models and their first-order referents can be decomposed.

Based on this assumption, two special cases can be distinguished for second-order models. One of them deals with models about first-order models, such as in large parts of the social sciences and humanities. Epistemology is an example of a discipline in which knowledge about the tools for gaining knowledge is developed. The other, simpler case addresses models about first-order referents, such as in the natural sciences. Here, second-order models can typically be condensed into first-order models. As far as complexity is concerned, this will be discussed in detail below (Section 3). Such special cases are also conceivable for the study of certain aspects of consciousness.

Second-order models allow us to consider the relations between first-order models and their referents as a referent of second order. If the decomposability of first-order models and their referents cannot be presupposed or is questionable, i.e., in holistic approaches, this option is of vital significance. After all, second-order approaches introduce a major extension of scientific modeling, and along with this a distinct shift of perspective. Before discussing this in

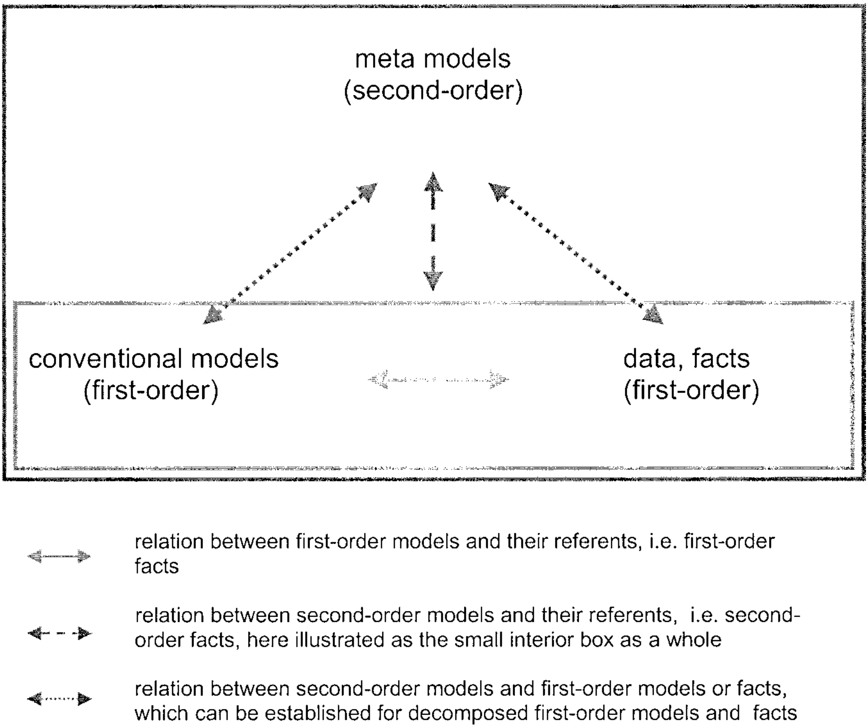

Fig. 1. Schematic illustration of second-order (meta-) models and their relations with first-order (conventional) models and their first-order referents.

more detail (Sections 2.3, 2.4), three short examples may serve to illustrate the basic idea.

- The relation between first-order models and their referents is crucial for any process of scientific discovery. Although the rational reconstruction of a first-order model and its application usually disregards this process retrospectively, the psychology of scientific discovery (and, more generally, of creativity) cannot dispense with it. The aspects of interpretation and understanding of a first-order model depend in large part on its relation to its referents. Typical questions of a second-order framework in this respect would be: Why is there a relation between mental categories (i.e., theoretical terms) and material referents (i.e., empirical facts)? Why is this relation not entirely arbitrary? How is it constituted? By contrast, first-order frameworks presuppose that such a relation exists and use it to answer questions referring to the structure and behavior of systems in the material world.
- A bit more specific is the problem of relationships between mental states of a mind and material states of a brain. In another terminology, this is phrased

in terms of so-called neural correlates of consciousness. Interestingly, a major part of current neuroscientific research focuses on specifying particular correlates of this kind. The problem of what constitutes the correlations is mostly ignored, rarely stated, hardly addressed, and absolutely unsolved. This situation is remarkable since the terminology of neural correlates would indicate consideration of the relation between mind and brain in a second-order sense. Actual research questions, however, refer more or less to first-order referents of first-order neurobiological models.

- In computer science and automata theory, frameworks of second order abound insofar as the key to almost everything in this field is the relation between hardware and software. This example is especially interesting since it refers to an area of investigation which is not only conceptually important but also of high practical relevance. Electronic chips and their elements (hardware) serve to implement algorithms (software) extremely efficiently. In this way, relations between hardware and software at a first-order level are established by sophisticated engineering tools. The engineer, however, is lost without a second-order perspective. For instance, developing a computing architecture means to determine beforehand which relations between hardware and software are required to design any intended devices for a specified purpose. Needless to say, this applies as well if human engineers are replaced by AI systems.

2.3. Validity

The issue of reference is of crucial importance for the validity of a model. In order to check whether a model is valid, e.g., is correct in a particular domain to be specified, the predicted behavior of some system due to the model has to be compared with the actual behavior of that system to which the prediction refers. At a first-order level, this is the core of the mutual relation between theory and experiment as it is properly applied in traditional areas of science. If theoretical predictions are correct, corresponding experiments provide results consistent with the predictions and thus confirm the theory. If theoretical predictions are incorrect, corresponding experiments provide results inconsistent with the predictions and thus disconfirm the theory and lead to its rejection or revision.

Although there are many complications to this brief and rough characterization if one looks at concrete situations of scientific practice in detail, this is a fair overall characterization for first-order frameworks. Considering second-order frameworks, the scenario becomes more involved due to the more complicated structure of reference relations. An experiment in the sense of a second-order framework deals with more than first-order referents of first-order models.[6] Second-order referents are in principle the entire complex of first-order models plus first-order referents plus their mutual relation with each other (compare again Figure 1).

As a consequence, a second-order model cannot be disconfirmed by first-order referents alone, except in cases where second-order referents can be shown to be reducible to first-order referents. In all non-trivial second-order situations this will not be possible. Hence, an experiment in the sense of a second-order framework is markedly different from a conventional first-order experiment. It has to include mental referents (e.g., psychological states) in addition to material referents (e.g., physical states), and, ideally, it should even include some explicit reference with respect to the relations between the two.

In non-trivial second-order situations, experiments must be designed in such a way that the more or less intricate relation between first-order models and their referents is properly taken into account. In general, this means that an experimental result capable of confirming or disconfirming a second-order model derives from a sophisticated combination of first-order models and their first-order referents. Only such a sophisticated combination as a whole can be a referent of a second-order model. Specific examples will be discussed below.

One can easily see how misleading an attempt to use first-order experiments to check the correctness of second-order models can be. Let us use two examples of the preceding subsection to demonstrate this. Consider a mind-matter system with correlated mental and material states. A second-order model would start with referring to such a system as a whole, and then proceed to its separate mental and material properties and their interrelations. Only in such a way would it be possible to address the issue of mind-matter correlations explicitly. Trying to confirm or disconfirm such a second-order model by first-order experimental results would amount to using material properties alone, e.g., the behavior of a particular neuron assembly, to check the correctness of the second-order model. The inappropriateness of such an approach is obvious.

It would be similarly inappropriate to try to check the validity of a proper implementation of a computing architecture by checking whether the transistors in the chips operate properly. Of course, the operation of transistors is a necessary precondition for the functioning of the computer, but it does not say anything about the proper operation of hardware and software together.

2.4. Reproducibility

Checking the validity of a model or a theory requires empirical results which are reliable. The reliability of a result clearly depends on whether it is a singular (chance) event or it can be reproduced under appropriate conditions. In this sense, the issue of reproducibility is a central methodological ingredient of contemporary sciences.[7] If an empirical observation cannot be reproduced, then the general opinion will be to ignore it, disregard it, or at least not take it as seriously as other results belonging to the established body of scientific knowledge. Non-reproducible results are incapable of confirming or disconfirming a theory.[8]

The criterion of reproducibility is not unconditional. First of all, to reproduce an empirical result means to observe the same result under circumstances essentially identical with those leading to its preceding observation. The essence of this condition is that the relevant circumstances must be known and controlled to such an extent that they can be adequately reestablished for any future attempt to reproduce an observation. If the circumstances are known well enough, the aspect of control is often achieved by suitable laboratory designs. In a given laboratory experiment, the setup is chosen in a way enabling a precise observation of a deliberately selected feature of a system.

An experiment provides conclusive answers only to questions for which it is properly designed. This is sometimes expressed by the sloppy notion of a "Procrustes strategy" governing laboratory-based science.[9] More generally speaking, there are always features which must be disregarded since they are considered irrelevant for the question to be answered. It is part of the art of experimental science to design experiments in such a way that a particular question is translated into relevant experimental conditions as precisely as possible. In simple words, the goal is to establish (relevant) facts within (irrelevant) noise. For the issue of reproducibility it does not matter if irrelevant conditions vary: only the relevant conditions must be kept fixed.

Many situations in large-scale systems (e.g., geophysics or astrophysics) do not enable any active control of empirical observations. Such situations are examples for an intermediate situation between typical experimental science with well-controlled boundary conditions and the notorious loss of precise control which is inherent in inanimate complex systems (e.g., the atmosphere of the earth), and even more so in living systems. Nevertheless, a large body of sound, important, and sophisticated knowledge has been collected for such systems. Of course, the same holds in those areas of the life sciences, e.g., behavioral or developmental biology, in which rigorously fixed laboratory conditions generally do not make sense.

To reproduce an empirical result also means that the quantitative value of an observable obtained in a measurement must be consistent with the values obtained in previous measurements. This is important since experimental results always have (epistemic) measurement errors. For this reason, a limit for an admissible scatter of individual results is required beyond which results are not considered as successfully reproduced.

At this point the significance of statistics enters; in many situations, the standard deviation of a distribution of measured values around a mean serves as a key quantity to distinguish individual results consistent with the expected distribution from those which are inconsistent (so-called "outliers"[10]). A result satisfies the criterion of individual reproducibility if it is consistent with the expected distribution. Such a strategy presupposes that the distribution of measured results is known or can be reasonably assumed, e.g., as a normal (Gaussian) distribution. Moreover, care must be taken that the measured results are not obscured by systematic errors, or the measured results must be corrected for such errors if they are known.

A basic assumption for such a strategy is the existence of a sharp, dispersion-free (ontic) value of an observable which would be identical with its measured value if there were no measurement errors or other distortions due to the measuring process. There are two possible complications, though, to this assumption. First, there may in principle be dispersion-free values of observables for individual events, but the events cannot be precisely and uniquely prepared (for all practical purposes). This implies variability, and one has to deal with a dispersed distribution of assumed dispersion-free values for an ensemble of individual events. Conventional areas of physics (such as classical point mechanics) consider the limit of a delta function for this distribution. In complex systems such an idealization is inappropriate, e.g., due to epistemic uncertainties introduced by uncontrollable influences in preparation and measurement.

Second, there are situations in which the assumption of a dispersion-free ontic valuation itself cannot be guaranteed. For instance, a quantum system with non-commutative observables cannot be in a state in which two such observables have dispersion-free expectation values. In this case, even observables for individual events cannot be valuated pointwise, but require valuations over sets of points. Corresponding non-vanishing dispersions are not a result of imperfect measurement or measurement distortions (epistemic problems), but they are intrinsic to the system as such (an ontic problem). They are the result of an interval-valued ontic observable rather than an epistemic uncertainty.

If the assumption cannot be maintained that measurement errors introduce only a scatter of results around a sharp value, reproducibility must be understood with respect to ensembles of events rather than individual events. In such cases, the identification of "outliers" is more complicated. It may then be appropriate to assume a whole distribution of values as the primary object of measurement rather than one sharp value. The overall scatter of results is then given by a convolution of this primary distribution with the distribution accounting for measurement errors. It is obvious that this can entail complications for increasingly complex systems.

Such complications can make it necessary to proceed from traditional first-order thinking to second-order approaches to describe complex systems more appropriately. An illustrative example typical for complex systems is the problem that an expectation value of an observable, defined in some limit $N \to \infty$ does not exist. In this situation, a possible second-order point of view can be realized by studying the way in which the mean value of the considered observable changes as a function of (finite) N. Any model predicting such a functional dependence, regardless of the existence of the limit $N \to \infty$, would qualify as a second-order model. The criterion of reproducibility would then have to refer to this functional dependence. It would be pointless to try to disconfirm such a second-order model by the lack of reproducibility due to the non-existence of an expectation value in a first-order framework. This and other examples will be discussed in detail in the following sections.

3. Complexity as a Second-Order Concept

The concept of complexity and the study of complex systems constitute an important focus of research in contemporary science. Although one might say that its formal core lies in mathematics and physics, complexity in a broad sense is certainly one of the most interdisciplinary issues scientists of many backgrounds talk about today. Beyond the traditional disciplines of the natural sciences, the "virus" of complexity has even crossed the border to areas like psychology, sociology, and ecology, among others. (For a rough overview concerning different approaches and applications see Atmanspacher [1997].)

It is the intention of the present section to provide arguments for the relevance of second-order frameworks in the study of complex systems (Section 3.1). Central to these arguments is the classification of complexity measures, which can be related to different classes of information measures. On this basis it is possible to outline formal ways of implementing second-order approaches for complex systems, in particular large deviations statistics (Section 3.2). Coupled map lattices are briefly addressed as particularly interesting candidates for concrete applications.

3.1. Convex Complexity and Pragmatic Information

A systematic orientation in the jungle of concepts of complexity is impossible unless a reasonable classification is at hand. There are several approaches that can be found in the literature: two of them are (i) the distinction of structural and dynamical measures (Wackerbauer et al., 1994) and (ii) the distinction of deterministic and statistical measures (Crutchfield & Young, 1989). Another, epistemologically inspired, scheme (iii) assigns ontic and epistemic levels of description to deterministic and statistical measures, respectively (Atmanspacher, 1994; Scheibe, 1973).

In addition to these approaches, a purely phenomenological criterion for classification can be given by the functional behavior in which a complexity measure is related to measures of randomness.[11] Within such an approach (for an early reference see Weaver [1968]), there are two classes of complexity measures: (iv) those for which complexity increases monotonically with randomness and those with a globally convex behavior as a function of randomness (cf. Figure 2). It turns out that classifications according to (ii) and (iii) distinguish measures of complexity in precisely the same manner as (iv) does: deterministic or ontic measures behave monotonically, and statistical or epistemic measures are convex. In other words, deterministic (ontic) measures are essentially measures of randomness, whereas statistical (epistemic) measures are not.

The class of monotonic measures of complexity contains, e.g., algorithmic complexity (Kolmogorov, 1965), various kinds of Rényi information (Balatoni & Rényi, 1956) (among them Shannon's information [Shannon & Weaver, 1949]), multifractal scaling indices (Halsey et al., 1986), and dynamical entropies (Kolmogorov, 1958). The class of convex measures of complexity contains, e.g., effective measure complexity (Grassberger, 1986), ε-machine complexity (Crutchfield & Young, 1989), fluctuation complexity (Bates & Shepard, 1991), and variance complexity (Atmanspacher et al., 1997). See also Landsberg and Shiner (1998) and Feldman and Crutchfield (1998) for further discussion.

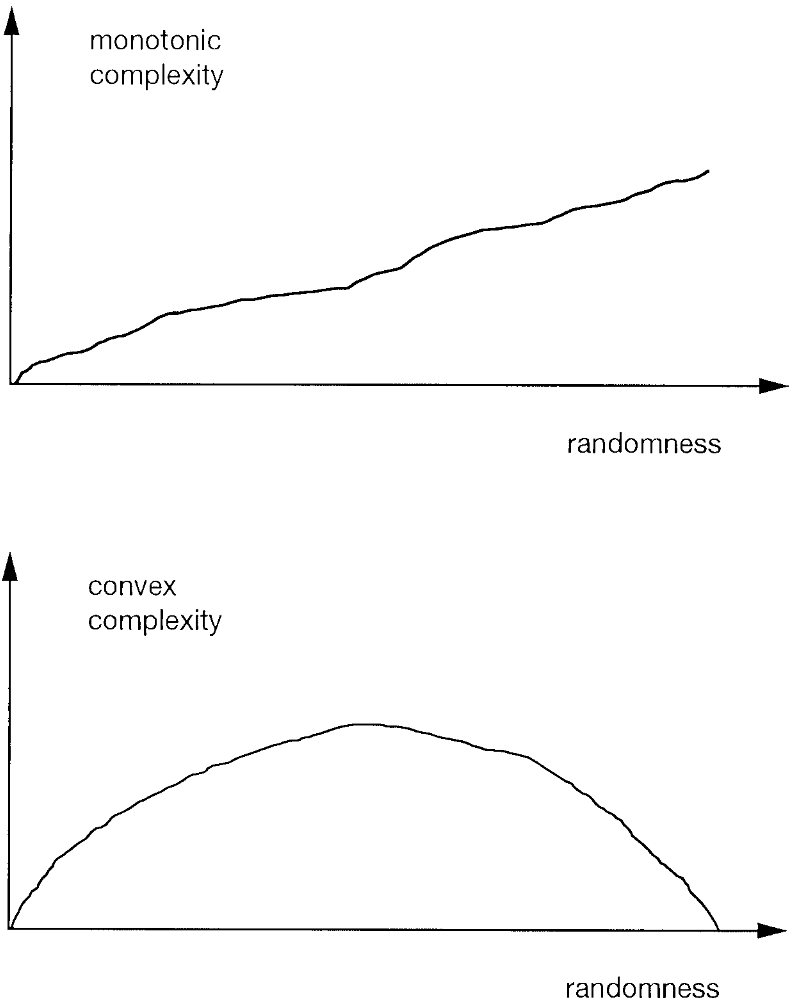

Fig. 2. Schematic illustration of two different classes of complexity measures, distinguished by their functional dependence on randomness.

A most intriguing additional difference (v) between both classes can be recognized if one focuses on the way statistics is implemented in each of these measures. The crucial point is that convex measures, in contrast to monotonic measures, are *meta*-statistically formalized, i.e., effectively represent (in one or another way) second-order statistics in the sense of "statistics of statistics".[12] Fluctuation complexity is the standard deviation (second-order) of a net mean information flow (first-order); effective measure complexity is the convergence rate (second-order) of a difference of entropies (first-order); ε-machine complexity is the Shannon information with respect to machine states (second-order) that are constructed as a compressed description of a data stream (first-order); and variance complexity is based on the variance (second-order) of the mean of many individual variances (first-order) of a distribution of data. To our knowledge, there is no monotonic complexity measure providing such a two-level statistical structure. Although it would be desirable to have a theorem for the corresponding relationship between convex complexity measures and their two-level statistical structure, such a theorem is not yet available.

Since so many complexity measures bear an intimate relation to information theoretical concepts, it is interesting to see whether first-order and second-order complexity measures can be related to corresponding information measures. The way information theory has been traditionally applied in physics up to now is limited to its syntactical component, going back to the influential work Shannon published in a book coauthored with Weaver (Shannon & Weaver, 1949). Weaver's contribution in this book already pointed out that this syntactical component of information requires extension to semantic and pragmatic aspects (for more details see Atmanspacher [1997]).

An interesting approach in this direction was introduced by E. von Weizsäcker (1974) as a way to deal with the usage that is based on the meaning of a message in terms of pragmatic information. This concept relies on the two notions of *primordiality* ("Erstmaligkeit") and *confirmation* ("Bestätigung"). Weizsäcker argued that a message that does nothing but confirm the prior knowledge of a receiver will not change its structure or behavior. On the other hand, a message providing only material completely unrelated (primordial) to any prior knowledge of the receiver will also not change its structure or behavior, simply because it will not be understood. In both cases, the pragmatic information of the message vanishes. A maximum of pragmatic information is assigned to a message that transfers an optimum mixture of primordiality and confirmation to its receiver. For the limiting case of complete confirmation, purely syntactic Shannon information and pragmatic information vanish coincidentally. If primordiality is added, Shannon information increases monotonically.

Applying a proper algorithm in order to generate a regularly alternating, periodic pattern, the corresponding generation process is obviously recurrent after the first two steps, i.e., after the generation of the first two elements of the pattern. Considering the entire generation process as a process of information transmission, it presents only confirmation of its first two time steps once they have passed by. In this sense, a regular pattern, exhibiting no complexity, corresponds to a process of information transmission that has vanishing pragmatic information (or "meaning") as soon as an initial transient phase (the first two time steps) has passed by. This applies to both notions of complexity, the deterministic as well as the statistical one.

For a completely random pattern the situation is more involved, since deterministic complexity and statistical complexity lead to different viewpoints. Deterministically, a random pattern is generated by an incompressible algorithm which contains as many steps as the pattern contains elements. The process of generating the pattern is not recurrent within the length of the algorithm. This means that it never ceases to produce elements that are unpredictable, except under the assumption that the entire algorithm was known *a priori.* Such knowledge, however, would imply that the pattern itself was known, since the algorithm is nothing but an incompressible description of it. Hence, the process generating a random pattern can be interpreted as a transmission of information completely lacking confirmation, and, consequently, with vanishing pragmatic information.

As a consequence, there is indeed a strong conceptual similarity between complexity measures and information measures. Pragmatic information is as convex as second-order complexity, and syntactic information is as monotonic as first-order complexity. We will come back to this similarity in Section 4, where a generalization from complex systems to complex mind-matter systems will be discussed. At that point, the important role of pragmatic information as a measure of meaning will become crucial (Section 4.1).

3.2. Large Deviations, Limit Theorems, and Ergodicity in Complex Systems

In contrast to many areas of conventional physics, non-stationary, transient states play a significant role in complex systems. Before a complex system reaches a stationary attractor characterized by an invariant measure, it can show exceedingly long transients. Along those transients, time averages are not equivalent to ensemble averages, such that ergodicity must not be presupposed in general, and ergodic measures must be used with caution (Tanaka & Aizawa, 1993). Moreover, it is now well known that careless applications of limit theorems in statistical analyses of data from complex systems can lead to misinterpretations (Wolfram, 1984). We are only beginning to understand these complications in detail.

In addition to Crutchfield's highly developed approach in terms of ε-machine reconstruction, the framework of large deviations statistics (LDS) (Aizawa, 1989; Bucklew, 1990; Ellis, 1985; Oono, 1989; Seppäläinen, 1995) offers itself as a promising route of access. Relationships between LDS and ε-machine reconstruction have been indicated by Young and Crutchfield (1994). LDS is particularly attractive since it distinguishes explicitly between statistical (monotonic) and meta-statistical (convex) measures of complexity. For some introductory formal remarks on LDS, including their relation to Jaynes' maximum entropy principle, see Amann and Atmanspacher (1999).

A basic element in LDS is a switch of perspective from statistical moments of a distribution, e.g., expectation values, to the probability measure itself, e.g., moments of a distribution of distributions. Moments of a distribution provide first-order statistical characterizations of this distribution. They are defined in the limit of $N \to \infty$, where N can be the number of particles, of degrees of freedom, of subensembles, etc. The corresponding "law of large numbers" states that in this limit a distribution converges weakly to the unit point measure at the expectation value. LDS specifies that this convergence is exponential as a function of N (Ellis, 1985; Oono, 1989). The convergence rate is the so-called "large deviations entropy".

If an observable is defined in the sense of an expectation value, then the relevant framework is that of a so-called level-1 description in the terminology of LDS. For instance, the formalism of multifractal measures (Halsey et al., 1986; Paladin & Vulpiani, 1987) is based on the limit $N \to \infty$; hence it is a level-1 theory and uses only first-order statistical measures. A more restrictive limit theorem which (other than a law of large numbers) presupposes the existence of the second moment of a distribution is the "central limit theorem". It gives an estimate for the probability that the size of properly defined, i.e., normalized, fluctuations around the expectation value is of the order of $\sqrt{N}$.

If the limit $N \to \infty$ as a precondition for a law of large numbers in the sense of a level-1 description cannot be presupposed, one can consider a higher level at which the observed empirical distribution functions themselves (not single variables) are treated as stochastic objects. Measures on such a higher level are second-order measures; they characterize the fluctuations of the distribution functions as a function of N. Distributions in a purely structural (non-dynamical) sense then give rise to (second-order) level-2 descriptions. A good example is the behavior of histograms of scaling indices for finite N as a function of N (which become multifractal measures for $N \to \infty$). For distributions covering structural as well as dynamical elements it can be reasonable to proceed to meta-statistical descriptions that are called level-3 descriptions in the terminology of LDS (Ellis, 1985; Oono, 1989). The objects of these descriptions are trajectories or histories instead of level-2 distributions.

A level-(n-1) theory can in general be obtained from the corresponding level-n theory ("contraction principle"; Ellis, 1985; Oono, 1989). For instance, it is possible to infer the convergence rate toward an expectation value (assuming that it exists) from the convergence rate of its probability distribution. An analogous contraction principle does not in general apply to the *moments*. If the distribution function depends on time, averages over time and ensemble averages are not necessarily identical (cf. Feller, 1971). If this difference is not explicitly taken into account, pitfalls with respect to the validity of a law of large numbers are to be expected. Pikovsky and Kurths (1994) have recently clarified such a misunderstanding for a level-3 situation (see also Griniasty & Hakim, 1994). They have shown that properly defined higher-order fluctuations do not violate a level-3 law of large numbers, whereas such a law is irrelevant for fluctuations in a lower-level description. Briefly speaking, Pikovsky and Kurths demonstrated that stationarity and ergodicity must not be presupposed in complex systems such as coupled map lattices or, more specifically, globally coupled maps.[13]

This is interesting in view of the fact that under particular conditions such systems can provide long-living transients—a type of non-stationary behavior far from being explored exhaustively.[14] For a detailed study of so-called "super-transients" in coupled maps, see, e.g., Kaneko (1990). In addition to the significance of coupled maps in general (Kaneko, 1993) and in the life sciences (Kaneko & Tsuda, 2000), non-stationary and non-ergodic behavior in this sense is expected to play a significant role in cognitive systems (cf. Freeman, 1994; Nozawa, 1994). Moreover, coupled map lattices offer an interesting perspective for non-hierarchical "control" insofar as the behavior at each site in the lattice crucially depends on its environment (consisting of neighboring sites). It is not simply determined by an externally adjusted (set of) parameter(s), which is the central idea in hierarchical approaches such as controlling chaos (Ott et al., 1990).

In summary, LDS provides a formal framework for addressing systems whose long-living transients prevent a straightforward application of limit theorems. If the existence of the limit in which expectation values are defined cannot be presupposed, LDS provides an approach to characterize a distribution function by its behavior as a function of N rather than its moments. If an expectation value exists, the large deviations entropy is simply an exponential decay rate of fluctuations. In principle, other kinds of convergence are possible, and even cases with a non-converging size of fluctuations can be conceived.

In the light of the discussions of Section 2, there is an important lesson to be learned from these deliberations. The validation of a first-order model with first-order data is usually considered to have failed if the distribution of data does not converge to an expectation value (of a relevant observable). From a second-order point of view, which may be necessary if first-order models and first-order data cannot be straightforwardly decomposed, such an attribution of failure is premature. In such cases, the non-existence of an expectation value can be an indication of the necessity of a second-order model, for whose validation first-order data alone are irrelevant. In the following section, this scenario will be discussed in more detail.

4. Mind-Matter Systems as Complex Systems

Why and in what sense can the psychophysical problem (i.e., the problem of how mind and matter are related to each other) be considered as a problem of complexity? Consciousness is often regarded as a property of a mental system arising when a particular level of complexity of its material correlate, the brain or the central nervous system, is reached. But there are subtler and more detailed aspects: for instance, several authors have emphasized that the concept of meaning, reference, or intentionality is essential to a definition of complexity (Atlan, 1991; Atmanspacher, 1994; Casti, 1992; Crutchfield, 1992; Grassberger, 1986; Haken, 1988). For instance, Grassberger (1989) wrote:

> Complexity in a very broad sense is a *difficulty* of a *meaningful task*. More precisely, the complexity of a pattern, a machine, an algorithm, etc. is the difficulty of the most important task related to it. (. . .) As a consequence of our insistence on *meaningful* tasks, the concept of complexity becomes *subjective*. We really cannot speak of the complexity of a pattern without reference to the observer. (. . .) A unique definition with a universal range of applications does not exist. Indeed one of the most obvious properties of a complex object is that there is no *unique* most important task related to it.

Addressing the concept of subjectivity in an unspecified way is tantamount to opening Pandora's box. Is a "science of the subjective" (Jahn & Dunne, 1997) possible at all, and, if it is, how can it be realized? In the following subsections, some specifications and ramifications of Grassberger's quotation will be outlined. First, it will be discussed how meaning and complexity can be related to each other from an epistemological point of view. This will be followed by some remarks concerning the issue of reproducibility within first-order and second-order approaches in this context. Finally, some indications of empirical second-order features will be noted.

4.1. Meaning and Complexity Are Complementary

Grassberger's quotation can be assessed in more detail if the two classes of complexity measures and associated information measures, as discussed in Section 3.1, are taken into account. Since monotonic, first-order measures of complexity are related to purely syntactic information, they can only be used to characterize systems in a way disregarding meaning.[15] If meaning is to be considered explicitly, one has to proceed to semantic or pragmatic information and associated convex, second-order measures of complexity. Corresponding definitions of complexity provide the validity domain to which Grassberger's quotation applies.

Two points should be stressed here. First, the fact that monotonic complexity is not related to meaning does not imply that corresponding measures are useless or ill-defined. It is obvious that there are many interesting applications of first-order complexity measures, and their benefit is that they do not lead to the complications which second-order complexity entails. Second, it should be kept in mind that, in contrast to syntactic information, semantic and/or pragmatic types of information are not defined as precisely as desirable. Hence, their relation to second-order complexity cannot be demonstrated as clearly as the relation between monotonic complexity and syntactic information. Nevertheless, their common feature of convexity is prominent enough to suspect an intimate connection between convex complexity and semantic/pragmatic information.

Complexity is a concept that has its origin in the study of physical properties of material systems. Meaning, on the other hand, originates in human concerns and has become a topic of philosophy and, more recently, cognitive science, and is discussed within a non-material domain. Assuming that the convexity of both second-order complexity and of pragmatic information are not accidental, it is remarkable how the perspectives of physics (complexity) and of cognitive science (meaning) show an explicit complementarity in this respect. "The impression of complexity often appears as something like the expression of an experience of meaning" (Casti, 1992).

From the viewpoints of the philosophy of mind and of cognitive science, dealing with the mental system and its mental properties (*res cogitans*), the concept of meaning is prior to the complexity of the brain as the material carrier of mental states. In the material domain of *res extensa*, on the other hand, the complexity of a system is prior to its capability to understand meaning. It seems in fact reasonable to expect that a certain degree and kind of complexity is a precondition for the capability of understanding meaning. Although it is still unclear what the exact criteria are in this respect, it would certainly be far too anthropocentric to fix them such as to exclude non-human beings.[16] It is even an open question to what extent meaning might be a reasonable concept for non-living systems. Atlan (1991) has proposed distinguishing different classes of complexity and assigning the notion of meaning only to a specific one among them. Other approaches, like those of von Weizsäcker (1974) or Crutchfield and Young (1989), do not restrict the notion of meaning in this manner.

Although complexity and meaning are subjects conceptually separated by a Cartesian cut between *res extensa* and *res cogitans,* it is evident that they have more to do with each other than such a cut would suggest. Moreover, the relations between the two are richer than one might expect naively. The nature of the interface between *res extensa* and *res cogitans* is almost unexplored as yet, and much work waits to be done to understand it better.

A key issue of corresponding approaches is the issue of decomposability. It is fairly obvious that the property of being complex is not appropriately treated by investigating systems in terms of decomposing them into parts. The same applies to the meaning of a message, a situation, or anything else. This does not merely restate the phrase that "the whole is more than the sum of its parts", but predicates a totally different perspective if the whole is to be studied instead of its parts. Some of the mathematical tools necessary for this kind of study might be structurally similar to those used in the quantum theory of entangled systems (see Primas, 1993). A corresponding attempt to describe cognitive functions in terms of non-commuting operators is due to Gernert (2000). Some formal ideas to generalize the quantum theoretical notions of complementarity and entanglement in and beyond physics have recently been outlined by Atmanspacher et al. (2002).

Beyond the decomposition of material systems, it is unavoidable to focus on the question of decomposability in the broader context of models and their relation to data. As discussed in Section 2, the reference structure of a second-order model is much more involved than that of a first-order model. For a second-order model, first-order models, first-order data, and their mutual relations are to be taken into account as referents. But in general it cannot be presupposed that these different classes of referents can be decomposed with respect to each other. It is not even known which conditions must be satisfied for such a decomposition.

Rephrased in terms of psychophysical (mind-matter) systems, the three classes of referents just mentioned translate into a mental component of a system, a material component of a system, and their mutual relation (see Figure 3). The most obvious example for such a system is the mind-brain of living beings. While cognitive, emotional, and even unconscious states belong to the mental component, electromagnetic or biochemical states belong to the material component.[17]

The mind-matter issue not only addresses the distinction of mind and brain but also, a bit more generally, that of mind and body, for instance in areas such as

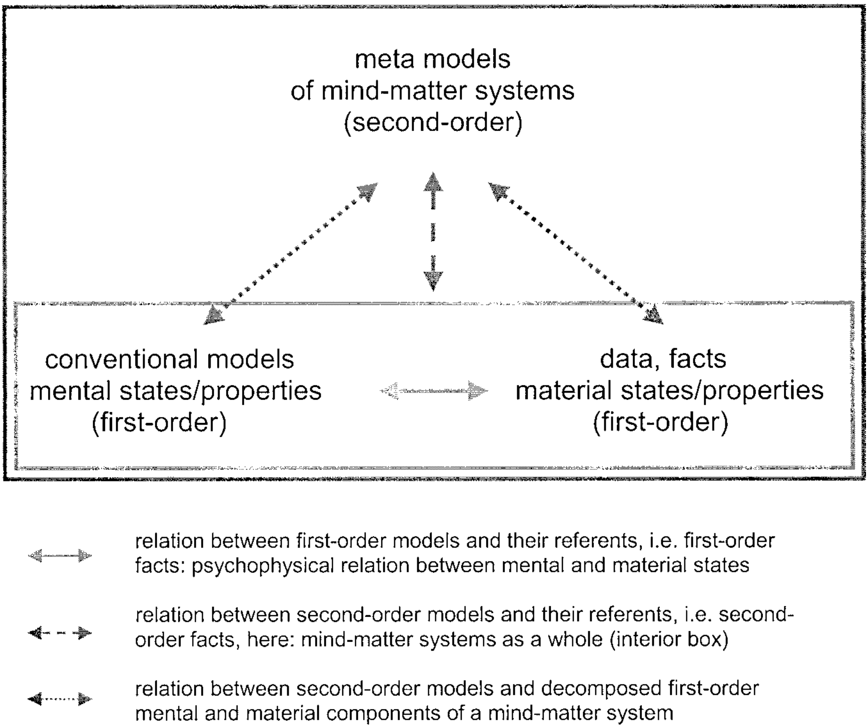

Fig. 3. Schematic illustration of second-order (meta-) models of mind-matter systems and their relations with mental states (first-order models) and material states (first-order data).

psychoendocrinology and psychoimmunology. Most counterintuitive, however, is the class of mind-matter questions dealing with relationships between conscious or unconscious states of the mental system of living beings and states of material systems *external* to them. Such topics of mind-matter research have traditionally been addressed in the off-mainstream fields of "psychical research" and "parapsychology", but those studies have rarely produced insights which could be incorporated into a consistent and acceptable scientific world view. Over the last decades, however, a number of more rigorous experiments have also been conducted, and a huge body of reliable data is now available. Meta-analytic evaluation of these sources remains difficult for a number of reasons (see Radin & Nelson, 1989; and Ehm, 2003), but is indicative of widespread and recurrent anomalous effects which must not be ignored, but are yet to be explained.

4.2. Second-Order Features in Mind-Matter Systems

As mentioned at the end of Section 3, a proper validation of second-order models has to refer to second-order data, which consist of first-order models, first-order data, and their relations, if they can be decomposed. For applications to mind-matter systems, the referents can be generalized to states/properties of mental systems, states/properties of material systems, and their mutual relations. Trying to validate a second-order model by first-order data alone is ill-posed. Likewise, it is incorrect (or even irrelevant) to try to disconfirm a second-order model by first-order data alone. This applies, for instance, if an expectation value of an observable does not exist as a first-order measurement result. If a second-order model would predict substantial correlations between mental and material components of second-order data, then it is possible that first-order data, restricted to the material component alone, are not well-defined, e.g., are statistically unstable. This can lead to improper conclusions concerning the validity of a second-order model if the implications of a decomposition of second-order data are not explicitly taken into account.

A standard example for the decomposition problem at the level of second-order data is provided by the technique of so-called randomized clinical trials, used to test the effectiveness of pharmacologically active substances versus placebos. In order to do so in an unbiased manner, it must not be presupposed that the active substance has the desired effect on an organism. Therefore the key feature in such trials is that there are two groups of test persons, whose allocation to active substance or placebo remains unknown (‘‘blinded’’) to both participants and experimenters for the duration of the trial. In this way, material and mental states or properties are kept non-decomposed in an epistemological sense. Only when the condition of blindedness is lifted can both kinds of substances be assigned to the two test groups. If the effectiveness of the active substance is significantly established for the ‘‘correct’’ test group, the second-order data can be properly decomposed into mental and material components, and the material component can be used as a referent of a first-order model explaining the effectiveness of the active substance.

If there are statistically unstable, ‘‘irreproducible’’ results at a first-order level, the ideal situation is that a second-order model can predict the way in which those results deviate from statistically stable first-order results. For such situations, Lucadou (1994) has argued that, under particular circumstances, a specific decline effect will occur if experimental reproducibility in a first-order sense is attempted. Generally speaking, this decline effect can be regarded as a consequence of the decomposition of (holistic) second-order data. The detailed arguments are based on a model concerning the amount of pragmatic information extracted from the studied system under specific experimental conditions. They are somewhat sophisticated and non-trivial, and should be checked in the original reference (Lucadou, 1994).

Decline effects and other problems related to reproducibility with respect to first-order experimental results are not unfamiliar in behavioral studies where mind-matter questions may play a role. An interesting example from learning research is the publication by Dworkin and Miller (1986), who report a decline effect for the change in heart rate due to visceral learning in rats. More recently, Crabbe et al. (1999) reported failures to reproduce behavioral changes as a consequence of particular genetic manipulations of mice.

Beyond such observations, reproducibility problems are frequently met in mind-matter research of the more unconventional kind mentioned above (Lucadou, 1991; Shapin & Coly, 1985; Utts, 1991). Of course, it would hardly be reasonable to claim that they all result from misinterpretations of second-order data due to ignored decomposition problems. But there may be cases in which this is indeed the case. For instance, a study by Vassy (1990) shows that Lucadou's conjecture is consistent with a fairly large database with respect to precognitive timing. A huge body of material from mind-matter experiments collected by PEAR (and, more recently, their consortium partners) shows decline effects and other indications of statistical instability (so-called "series position effects"; see, e.g., Dunne et al. [1994]).

The mind-matter system studied in these experiments is based on the following components. First there is a material random process providing a distribution of random events as a function of time. Second, there is a human operator who is asked to apply his mental intention such that the mean of the random distribution either increases or decreases, depending on a prescribed instruction. The observables to be correlated are the shift of the mean of the distribution as compared to the expected mean and the intention of the operator (high or low). After a 12-year period of collecting data from 13 distinct experiments within this experimental paradigm, Jahn et al. (1997) reported a small but highly significant mean shift whose sign was correctly correlated with intention. Three independent attempts to reproduce this result all showed mean shifts in the intended direction, but at a statistically insignificant level, hence no corresponding correlations could be claimed (Jahn et al., 2000). (However, a number of other correlations concerning "secondary parameters" could be detected which were significant.)

Considered naively, this lack of reproducibility might be interpreted to indicate that mind-matter correlations in these experiments do not exist after all. However, in a second-order framework a declining effect size with respect to the material component of a mind-matter system must not necessarily be taken as such an indication.[18] As long as the mental component of the system and its relation to the material component remain weakly defined and largely undetermined, the system as a whole has so many degrees of freedom that reasonable implications for a second-order model cannot be drawn from a declining first-order result. On the other hand, a first-order framework, within which the mentioned data could be interpreted as evidence against mind-matter correlations, can be ruled out for a proper description of mind-matter systems.

Another interesting line of second-order thinking, related to the correspondence between second-order complexity and pragmatic information as a measure of meaning, derives from empirical data collected and analyzed by May et al. (2000). In a set of remote perception experiments, they tried to relate the success of remote perception to information theoretical characteristics of the images to be viewed. They found significant correlations ($r = 0.232$) between remote perception success ("figures of merit") and the gradient of Shannon entropies calculated for systematically shifted, small subimages of the image as a whole.

In contrast, calculations of the variance complexity of the images, a second-order complexity measure introduced by Atmanspacher et al. (1997), provides non-significant correlations ($r = -0.095$) with the corresponding figures of merit (May, 1999). This difference is not finally understood yet, but it might indicate that the success of remote perception is more correlated with syntactic rather than semantic (and pragmatic) features of images. On the other hand, recent results by Dunne and Jahn (2003) suggest the opposite.

A more far-reaching idea concerning the reproducibility of events in mind-matter systems arises in the dialog between Pauli and Jung. In the volume "Explanation of Nature and the Psyche" (Jung & Pauli, 1952), Jung presented his ideas about acausal connections ("synchronicities") between mind and matter in a comprehensive manner (Jung, 1952). He used empirical material from an astrological study about sun and moon positions in birth charts of partners as evidence for synchronistic relations. Pauli was quite unhappy with this particular example (which indeed showed a decline in significance when further material was added later on). In a letter to Fierz (Pauli, 1952), he wrote: "... synchronistic phenomena in a narrower sense cannot be grasped in terms of natural laws since they are not reproducible, i.e., singular. They are blurred by the statistics of large numbers. Just on the contrary, 'acausalities' in physics are captured by statistical laws (of large numbers)." In other words, if outliers of a distribution are due to singular events, they cannot be consistently discussed according to the usual laws of statistics. Simple as this statement reads, its applicability to concrete events is hard to check as long as there are no good criteria governing how to distinguish singular events from events which can be analyzed statistically.

In the same context, it is appropriate to refer to a proposal Pauli made concerning the relevance of meaning-related issues in mind-matter research. From a historical perspective, Pauli (1954) suggested to interpret Darwin's use of the concept of chance in order to model biological evolution as "an attempt to theoretically cling, according to the ideas of the second half of the 19th century, to the total elimination of any finality. As a consequence, this has in some way to be replaced by the introduction of chance." Pauli speculated that the concept of synchronicity might force science to revive the historically repressed concept of finality as a complement to causality. In *Die Vorlesung an die fremden Leute* (part of the essay *Die Klavierstunde*, Pauli, 1953), Pauli wrote about a "third kind of natural laws which consists in correcting the fluctuations of chance by meaningful or functional coincidences of causally non-connected events", in addition to deterministic and statistical laws of nature. But he hesitated to publish such thoughts (Pauli, 1953): "If one really would like to propose such ideas in public, it would be imperative to show something which is verifiable."

5. Summary and Conclusions

The Pauli quotation at the end of the preceding section leaves us with the problem of deciding what could be considered as a reliable verification. One possible meaning of verification refers to the reproducibility of experimental data. According to the main arguments in this article, complex mind-matter systems require us to consider types of reproducibility which are more sophisticated than those used in conventional scientific methodology. Since complex mind-matter systems in general require second-order models for their description, their validity cannot be reasonably tested by conventional first-order data alone.

Some recent ideas and techniques to study complex systems in the sense of second-order approaches were presented in Section 3. A central feature in this regard is the correspondence between complexity measures which are convex as a function of randomness, on the one hand, and pragmatic information as a measure of meaning on the other. The second-order character of these two concepts of complexity and meaning is related to the non-stationary, non-ergodic behavior of complex systems (e.g., coupled map lattices), for which limit theorems cannot be presupposed. Novel statistical tools of analysis such as large deviations statistics offer interesting perspectives in these contexts.

Section 4 extended the scope of the discussion from complex systems to mind-matter systems. A complementarity of complexity and meaning was proposed and discussed to reflect two different aspects of the same problem viewed from different sides: the problem of connecting the material and mental domains generated by the Cartesian distinction. The formal correspondence between convex complexity and pragmatic information outlined in Section 3 was generalized to systems to which, in addition to some complexity of their material component, mental properties such as consciousness, intentionality, and meaning can be explicitly attributed. Some second-order characteristics of such systems were indicated. It was shown that a declining statistical significance of empirical results does not necessarily indicate that such results can be disregarded. It can also indicate the necessity of a second-order approach or even validate a second-order model.

Another, broader meaning of the topic of verification as addressed by Pauli refers to the consistency of experimental data with corresponding theoretical approaches. Although a full-fledged theoretical framework for the description of mind-matter systems is not available, many conceptual speculations with different degrees of specification can be found in the literature. Some recent lines of thinking, particularly emphasizing the role of the unconscious, were outlined in Atmanspacher (2003) or Jahn and Dunne (2001). Among other features, the speculative schemes addressed there provide room for a second-order approach toward decline effects.

After all, however, progress in mind-matter research requires more than novel conceptual approaches and models. "To show something which is verifiable" (Pauli, 1953) means that we need empirical data which can be compared with theoretical approaches. Due to the necessity of second-order thinking, it is expected that "verification" in mind-matter research is conceptually more complicated than in many areas of conventional science.

Notes

[1] In this article, we prefer to speak of reproducibility since replicability has the additional connotation of copying or cloning biological species. Moreover, reproducibility can be considered to characterize a special situation within the broader notion of repeatability.

[2] This assumption contradicts a radically materialistic position in which only material components are considered. We do not regard such a position as viable.

[3] It should be emphasized that throughout this paper the Cartesian framework is not taken so far as to indicate two ontologically different substances. (The notion of substance is here used as a philosophical term.) However, an epistemological separation of the two seems unavoidable in order to explore possible relations between mental and material domains, or between models and facts.

[4] Such problems arise whenever propositions and their referents are not distinguished in a clear-cut way. Typical examples are propositions (partially) referring to themselves, as in "This statement is false".

[5] Second-order thinking, i.e., thinking about thinking, is an illustrative example (cf. Elkana, 1986). Second-order thinking has gained major influence in cybernetics, where von Foerster coined the notion "second-order cybernetics" to address "the description of the 'describer' " (von Foerster, 1982). The journal *Cybernetics & Human Knowledge* has been devoted to critical discussions of second-order cybernetics, autopoiesis, and cybersemiotics since 1992.

[6] Referring back to the second example in the preceding subsection 2.2, material brain states are usually studied by first-order experiments, whereas an understanding of their correlations with mental states would be in the domain of a second-order experiment.

[7] It is interesting to note that Vico's famous verum-factum principle ("verum et factum convertuntur"; Vico [1990]) can be considered as a criterion for validity that is even stronger than reproducibility. Somewhat generously translated and interpreted, "the true and the made are the same" means that some model is considered valid only insofar as it can be utilized for engineering applications. In this article we will not further discuss Vico's principle.

[8] Needless to say the reproducibility of an empirical result is a necessary but not a sufficient condition for its acceptance in the sciences. An essential additional point, which cannot be emphasized enough, is the consistent incorporation and interpretation of such a result in a theoretical framework. In his definition, Tetens (1995) emphasizes this point as well: "... it makes sense to speak of the reproducibility of a fact only relative to an explicitly formulated description".

[9] Procrustes, a figure of Greek mythology, was a host who promised his guests a very special bed exactly matching their size. As soon as the guest lay down,

Procrustes went to work on him, stretching him on the rack if he was too short and chopping off his legs if he was too long.

10 The notion of an outlier is here used with respect to its deviation from a most probable value of a distribution, without regarding the origin of this deviation.

11 It is worth mentioning that randomness itself is a concept that is anything but finally clarified. In the framework of the present paper we use the notion of randomness in the broad sense of an entropy.

12 Note that the notion of "second-order statistics" has nothing to do with the second moment of a statistical distribution or second-order terms in a series expansion. Rather, it indicates a statistical approach which can serve to check the validity of a second-order model by an evaluation of second-order data.

13 An explicit large deviations approach to globally coupled maps is due to Hamm (2000).

14 A novel statistical approach toward the study of non-stationary processes was recently proposed by Galluccio et al. (1997) for an analysis of currency exchange rates. The key idea, similar to so-called random time change techniques, is to rescale the time axis of the process such that periods with high (low) activity are stretched (squeezed). This is done in such a way that the original non-stationary process is transformed into a stationary process which can be investigated by standard means. Galluccio et al. (1997) call their approach an "intrinsic time analysis", alluding to the fact that non-stationary behavior can be suitably characterized by a proper renormalization of time.

15 This view presupposes a certain kind of (analytical) bottom-up argumentation in the sense that information can be decomposed into its syntactic, semantic, and pragmatic components. From a top-down point of view one could argue alternatively that the phenomenological ("Lebenswelt") significance of information derives from the irrelevance of such a decomposition. From such a perspective, every element of syntax is inseparably linked to aspects of meaning and use, and it does not make sense to consider each of them separately.

16 Nevertheless, notions of meaning intended to apply beyond human beings (e.g., animals or AI systems) are often configured by analogies or similarities with our everyday notion of meaning.

17 In the framework of a more or less radical materialistic position only material states are considered, so there is no need at all to talk about mind-matter systems within such a framework.

18 It should be mentioned that declines of statistical significance can, of course, be due to much simpler reasons. Any laboratory scientist knows that complicated experiments can produce so-called artefacts until the experimental setup is properly under control. Such artefacts must not be confused with second-order effects as addressed here.

Acknowledgments

We are grateful to Robert Bishop, Hans Primas, and two anonymous referees for critical comments on earlier versions of this manuscript, and to Werner Ehm and Christian Scheer for communicating results prior to publication.

References

Aizawa, Y. (1989). Non-stationary chaos revisited from large deviation theory. *Progress in Theoretical Physics Supplement, 99,* 149–164.

Amann, A., & Atmanspacher, H. (1999). Introductory remarks on large deviations statistics. *Journal of Scientific Exploration, 13,* 639–664.

Atlan, H. (1991). Intentional self-organization in nature and the origin of meaning. In Rossi, C., & Tiezzi, E. (Eds.), *Ecological Physical Chemistry* (pp. 311–331). Amsterdam: Elsevier.

Atmanspacher, H. (1994). Complexity and meaning as a bridge across the Cartesian cut. *Journal of Consciousness Studies, 1,* 168–181.

Atmanspacher, H. (1997). Cartesian cut, Heisenberg cut, and the concept of complexity. *World Futures, 49,* 333–355.

Atmanspacher, H. (2003). Mind and matter as asymptotically disjoint, inequivalent representations with broken time-reversal symmetry. *BioSystems, 68,* 19–30.

Atmanspacher, H., Räth, C., & Wiedenmann, G. (1997). Statistics and meta-statistics in the concept of complexity. *Physica A, 234,* 819–829.

Atmanspacher, H., Römer, H., & Walach, H. (2002). Weak quantum theory: Complementarity and entanglement in physics and beyond. *Foundations of Physics, 32,* 379–406.

Balatoni, J., & Rényi, A. (1956). Remarks on entropy. *Publications of the Mathematics Institute of the Hungarian Academy of Science, 9,* 9–40.

Bates, J. E., & Shepard, H. (1991). Information fluctuation as a measure of complexity, UNH (Durham) preprint. See also Bates, J. E., & Shepard, H., Measuring complexity using information fluctuations. *Physics Letters A, 172,* 416–425.

Bucklew, J. A. (1990). *Large Deviations Techniques in Decision, Simulation, and Estimation.* New York: Wiley.

Casti, J. (1992). The simply complex: Trendy buzzword or emerging new science? *Bulletin of the Santa Fe Institute, 7,* 10–13.

Crabbe, J. C., Wahlsten, D., & Dudek, B. C. (1999). Genetics of mouse behavior: Interactions with laboratory environment. *Science, 284,* 1670–1672. See also the comment by M. Enserink in the same issue, pp. 1599–1600.

Cronbach, L. J. (1983). *Designing Evaluations of Educational and Social Programs.* San Francisco: Jossey-Bass.

Crutchfield, J. P. (1992). Knowledge and meaning... chaos and complexity. In Lam, L., & Naroditsky, V. (Eds.), *Modeling Complex Phenomena* (pp. 66–101). Berlin: Springer.

Crutchfield, J. P., & Young, K. (1989). Inferring statistical complexity. *Physical Review Letters, 63,* 105–108.

Dunne, B. J., Dobyns, Y. H., Jahn, R. G., & Nelson, R. D. (1994). Series position effects in random event generator experiments, with an appendix by A. Thompson. *Journal of Scientific Exploration, 8,* 197–215. See also the Technical Report PEAR 93002 (Princeton 1993), which contains a collection of effect sizes of 21 prolific operators as a function of series number which is not included in the published version.

Dunne, B. J., & Jahn, R. G. (2003). Information and uncertainty in remote perception research. *Journal of Scientific Exploration, 17,* 207–241.

Dworkin, B. R., & Miller, N. E. (1986). Failure to replicate visceral learning in the acute curarized rat preparation. *Behavioral Neuroscience, 100,* 299–314.

Ehm, W. (2003). Manuscript in preparation.

Elkana, Y. (1986). Die Entstehung des Denkens zweiter Ordnung im klassischen Griechenland. In *Anthropologie der Erkenntnis* (pp. 344–375). Frankfurt: Suhrkamp.

Ellis, R. S. (1985). *Entropy, Large Deviations, and Statistical Mechanics.* New York: Springer.

Feldman, D. P., & Crutchfield, J. P. (1998). Measures of statistical complexity: Why? *Physics Letters A, 238,* 244–252.
Feller, W. (1971). *An Introduction to Probability Theory and Its Applications* (Vol 1). New York: Wiley, pp. 152ff.
von Foerster, H. (1982). *Observing Systems.* Seaside, CA: Intersystems Publications, p. 258.
Freeman, W. (1994). Neural mechanisms underlying destabilization of cortex by sensory input. *Physica D, 75,* 151–164.
Galluccio, S., Caldarelli, G., Marsili, M., & Zhang, Y.-C. (1997). Scaling in currency exchange. *Physica A, 245,* 423–436.
Gernert, D. (2000). Towards a closed description of observation processes. *BioSystems, 54,* 165–180.
Grassberger, P. (1986). Toward a quantitative theory of self-generated complexity. *International Journal of Theoretical Physics, 25,* 907–938.
Grassberger, P. (1989). Problems in quantifying self-generated complexity. *Helvetica Physica Acta, 62,* 489–511.
Griniasty, M., & Hakim, V. (1994). Correlations and dynamics in ensembles of maps. *Physical Review E, 49,* 2661–2667.
Haken, H. (1988). *Information and Self-Organization.* Berlin: Springer, Section 1.6.
Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I., & Shraiman, B. I. (1986). Fractal measures and their singularities: The characterization of strange sets. *Physical Review A, 33,* 1141–1151.
Hamm, A. (2000). Large deviations from the thermodynamic limit in globally coupled maps. *Physica D, 142,* 41–69.
Jahn, R. G., & Dunne, B. J. (1997). Science of the subjective. *Journal of Scientific Exploration, 11,* 201–224.
Jahn, R. G., & Dunne, B. J. (2001). A modular model of mind/matter manifestations (M^5). *Journal of Scientific Exploration, 15,* 299–329.
Jahn, R. G., Dunne, B. J., Nelson, R. D., Dobyns, Y. H., & Bradish, G. J. (1997). Correlations of random binary sequences with pre-stated operator intention: A review of a 12-year program. *Journal of Scientific Exploration, 11,* 345–367.
Jahn, R., Dunne, B., Bradish, G., Dobyns, Y., Lettieri, A., Nelson, R., Mischo, J., Boller, E., Bösch, H., Vaitl, D., Houtkooper, J., & Walter, B. (2000). Mind-Machine Interaction Consortium: PortREG replication experiments. *Journal of Scientific Exploration, 14,* 499–555.
Jung, C. G. (1952). Synchronizität als ein Prinzip akausaler Zusammenhänge. In Jung, C. G., & Pauli, W. (Eds.), *Naturerklärung und Psyche* (pp. 1–107). Zürich: Rascher.
Jung, C. G., & Pauli, W. (1952). *Naturerklärung und Psyche.* Zürich: Rascher.
Kaneko, K. (1990). Supertransients, spatiotemporal intermittency, and stability of fully developed spatiotemporal chaos. *Physics Letters A, 149,* 105–112.
Kaneko, K. (Ed.). (1993). *Theory and Applications of Coupled Map Lattices.* New York: Wiley.
Kaneko, K., & Tsuda, I. (2000). *Complex Systems: Chaos and Beyond.* Berlin: Springer.
Kolmogorov, A. N. (1958). A new metric invariant of transitive dynamical systems and automorphisms in Lebesgue spaces. *Doklady Akademii Nauk SSSR, 119,* 861–864. See also Sinai, Ya. G. (1959). On the notion of entropy of a dynamical system. *Doklady Akademii Nauk SSSR, 124,* 768.
Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of complexity. *Problems of Information Transmission, 1,* 3–11.
Krathwohl, D. R. (1985). *Social and Behavioral Science Research.* San Francisco: Jossey-Bass.
Landsberg, P. T., & Shiner, J. S. (1998). Disorder and complexity in an ideal non-equilibrium Fermi gas. *Physics Letters A, 245,* 228–232.
von Lucadou, W. (1991). Some remarks on the problem of repeatability of experiments with complex systems. In Atmanspacher, H., & Scheingraber, H. (Eds.), *Information Dynamics* (pp. 143–151). New York: Plenum.
von Lucadou, W. (1994). Wigner's friend revitalized? In Atmanspacher, H., & Dalenoort, G. J. (Eds.), *Inside Versus Outside* (pp. 369–388). Berlin: Springer.
Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70,* 151–159.
May, E. C. (1999). Private communication in February 1999. The variance complexity for the images used in May et al. (2000) was calculated by Christian Scheer. The results were then correlated with the corresponding figures of merit.

May, E. C., Spottiswoode, S. J. P., & Faith, L. V. (2000). The correlation of the gradient of Shannon entropy and anomalous cognition: Toward an AC sensory system. *Journal of Scientific Exploration, 14,* 53–72.

Nozawa, H. (1994). Solution of the optimization problem using the neural network model as a globally coupled map. *Physica D, 75,* 179–189.

Oono, Y. (1989). Large deviations and statistical physics. *Progress in Theoretical Physics Supplement, 99,* 165–205.

Ott, E., Grebogi, C., & Yorke, J. A. (1990). Controlling chaos. *Physical Review Letters, 64,* 1196–1199.

Paladin, G., & Vulpiani, A. (1987). Anomalous scaling laws in multifractal objects. *Physics Reports, 156,* 147–225.

Pauli, W. (1952). Letter to Fierz of June 3, 1952. In von Meyenn, K. (Ed.), *Wolfgang Pauli. Wissenschaftlicher Briefwechsel, Band IV/1: 1950–1952* (p. 634). Berlin: Springer.

Pauli, W. (1953). Die Klavierstunde. In Atmanspacher, H., Primas, H., & Wertenschlag-Birkhäuser, E. (Eds.). *Der Pauli-Jung-Dialog und seine Bedeutung für die moderne Wissenschaft* (pp. 317–330). Berlin: Springer.

Pauli, W. (1954). Naturwissenschaftliche und erkenntnistheoretische Aspekte der Ideen vom Unbewussten. *Dialectica, 8,* 283–301. English translation in Enz, C. P., & von Meyenn, K. (Eds.), *Wolfgang Pauli, Writings on Physics and Philosophy* (pp. 149–164). Berlin: Springer.

Pikovsky, A. S., & Kurths, J. (1994). Do globally coupled maps really violate the law of large numbers? *Physical Review Letters, 72,* 1644–1646. See also Pikovsky, A. S. (1993). Comment on 'Noisy uncoupled chaotic map ensembles violate the law of large numbers'. *Physical Review Letters, 71,* 653.

Popper, K. R. (1965). *The Logic of Scientific Discovery.* London: Hutchinson. See also Popper, K. R. (1957). *The Poverty of Historicism.* London: Routledge and Kegan Paul.

Primas, H. (1993). The Cartesian cut, the Heisenberg cut, and disentangled observers. In Laurikainen, K. V., & Montonen, C. (Eds.), *Symposia on the Foundations of Modern Physics. Wolfgang Pauli as a Philosopher* (pp. 245–269). Singapore: World Scientific.

Radin, D. I., & Nelson, R. D. (1989). Evidence for consciousness-related anomalies in random physical systems. *Foundations of Physics, 19,* 1499–1514.

Scheibe, E. (1973). *The Logical Analysis of Quantum Mechanics.* Oxford: Pergamon.

Seppäläinen, T. (1995). Entropy, limit theorems, and variational principles for disordered lattice systems. *Communications in Mathematical Physics, 171,* 233–277.

Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication.* Urbana: University of Illinois Press.

Shapin, B., & Coly, L. (Eds.). (1985). *The Repeatability Problem in Parapsychology.* New York: Parapsychology Foundation.

Sidman, M. (1960). *Scientific Research.* New York: Basic.

Smith, N. G., Jr. (1970). Replication research: A neglected aspect of psychological research. *American Psychologist, 25,* 970–975.

Tanaka, K., & Aizawa, Y. (1993). Fine structures in stationary and non-stationary chaos. *Progress in Theoretical Physics, 90,* 547–567.

Tetens, H. (1995). Reproducibility. In Mittelstrasse, J., et al. (Eds.), *Enzyklopädie Philosophie und Wissenschaftstheorie Band 3.* Stuttgart: Metzler.

Utts, J. (1991). Repeatability and meta-analysis in parapsychology. *Statistical Science, 6,* 396–403.

Vico, G. B. (1990). *Prinzipien einer neuen Wissenschaft über die gemeinsame Natur der Völker.* Hamburg: Meiner. The original quotation is in *De antiquissima Italorum sapientia* of 1710.

Vassy, Z. (1990). Experimental study of precognitive timing: Indications of a radically noncausal operation. *Journal of Parapsychology, 54,* 299–320.

Wackerbauer, R., Witt, A., Atmanspacher, H., Kurths, J., & Scheingraber, H. (1994). A comparative classification of complexity measures. *Chaos, Solitons, & Fractals, 4,* 133–173.

Weaver, W. (1968). Science and complexity. *American Scientist, 36,* 536–544.

von Weizsäcker, E. (1974). Erstmaligkeit und Bestätigung als Komponenten der pragmatischen Information. In von Weizsäcker, E. (Ed.), *Offene Systeme I* (pp. 83–113). Stuttgart: Klett-Cotta.

Wolfram, S. (1984). Computation theory of cellular automata. *Communications in Mathematical Physics, 96,* 15–57.

Young, K., & Crutchfield, J. P. (1994). Fluctuation spectroscopy. *Chaos, Solitons, & Fractals, 4,* 5–39.