ZWIRN
Center for Scientific Interdisciplinary Risk and Sustainability Management

Ostfalia University of Applied Sciences

# Prediction and Predictability

Survey on the State of Knowledge about
Foundations of Prediction and Limits of Predictability

**Achim Michalke**

Professor of Technical Management

Faculty of Supply Engineering

Ostfalia University of Applied Sciences

Wolfenbüttel, Germany

a.michalke@ostfalia.de

www.ostfalia.de/pws/michalke

ORCID-ID 0000-0001-7895-0901

April 2022

## Abstract

The prediction of future events is a fundamental component of our thinking and acting. Prediction is the basis for any directed action and any meaningful development. In view of its eminent role, surprisingly little research has been targeted on the methods and results of prediction. In particular, very little is known on the accuracy and reliability of prediction results. Limitations of prediction are mostly ignored in predictive situations, including scientific ones.

This survey intends to provide an overview on what is known about the targets, methods, and results of prediction, with a particular focus on quality and limitations of prediction. As knowledge on prediction is distributed across many scientific disciplines, from philosophy of science to information technology, from quantum physics to the theory of complexity, it is also designated to give a holistic picture of prediction across all these disciplines. This picture is by no means complete, it is rather intended as a starting point for further research on the topic.

The discussion is limited to scientific aspects of prediction. Aspects beyond rationality like divination and prophecy are excluded.

***Keywords:*** *Causality, Chance, Complexity, Decision, Direction of Time, Error, Extrapolation, Forecast, Future, Hazard, Induction, Inference, Planning, Prediction, Predictability, Probability, Randomness, Recursion, Risk, Uncertainty, Validity*

# Prediction and Predictability

Survey on the State of Knowledge about
Foundations of Prediction and Limits of Predictability


Motivation .... 2
- Importance of Predictions in Everyday Life .... 3
- Missing Awareness of Limitations in Prediction.... 4

Prediction Fundamentals .... 4
- What is a Prediction? .... 4
- The Reason for Predictions.... 5
- The Direction of Time .... 6
- Causality .... 6
- Determinism and Chance.... 7
- Scientific Investigation of Predictions .... 8

The Process of Prediction.... 9
- Predictive Inference .... 10
- Reduction to Relevant Parameters .... 10
- Types and Variants of Predictions .... 11

Limits to Predictability .... 12
- Chance.... 13
- Chaos.... 14
- Complexity .... 14
- Self-Influencing Predictions.... 15

Quality of Predictions.... 16
- Accuracy of Predictions.... 17
- Reliability of Predictions .... 17
- Uncertainty and the Degree of Confidence .... 19

Conclusive Remarks.... 19

References .... 20


# Motivation

I came across the issue of prediction and predictability when I tried to understand the concept of risk. The idea of risk is that planning has just a slight chance to go wrong – not enough to put planning into question, but important enough to consider it as a disturbance of the targeted development. As everyone knows, this might be true in some situations – as you press the gas pedal of your car, it most likely will accelerate

– but might be utterly wrong in other situations, like setting up a business plan for an innovative product. However, the distinction between both cases is diffuse, and outside some very specific applications (like risk assessment for insurance, or market volatilities in finance) the accuracy of planning is completely ignored.

This should not surprise us, because the basis is simply missing: Planning and predicting is considered as a primordial task in most circumstances, as something that one simply is capable to do, just out of nothing but experience. A good manager is expected to be a good planner, without explaining him how to plan, and without him being able to explain how he has made his predictions. Prediction is simply not scrutinized, neither by the sciences that thoroughly rely on it, nor by a science of itself: There is no consistent theory of prediction, and in particular there is no theory that accounts for the accuracy and reliability of predictions.

Of course, there is knowledge on prediction, but this knowledge is spread across a variety of scientific disciplines. The aim of my effort here is to collect, compile and integrate available knowledge towards a better understanding of prediction and assessment of predictability. As a result, we could gain a better understanding where further research would be worthwhile to improve our application of predictions.

## Importance of Predictions in Everyday Life

Predictions are fundamental in our life, although we rarely recognize them when we make them. Every action we do is based on a decision selecting one of the possible alternatives, and every decision is based on predictions how the available alternatives would develop into the future. We need these predictions to assess which of the alternatives are likely to develop closest to our targets. The process is always the same, whether we decide where to put our next footstep, or where to invest our fortune. Without predictability, no directed action is possible.

However, in most cases we don’t even recognize that we are making predictions. Our brain is a prediction machine, in both its conscious as well as its unconscious activities. Thoughts are not just for fun, they shall enable us to act in a target-oriented way, and action is possible only into the future. Nearly every thought that we have is intended to prepare or produce predictions, and to use them for assessments. Unconscious brain activities support our fast and intuitive actions, again using predictive methods to assess their future effects.

The final aim of any knowledge, of any science (with the possible exception of historical sciences) is to make more reliable predictions. As a consequence, the philosophy of science is also a philosophy of prediction in a way. Whereas our predictions in decision-making are mostly private and often even subconscious, predictions in science are public and open to scrutiny. As a consequence, predictions in science must be o fa rational kind, enabling investigation beyond personal experience.

In contrast to its omnipresence and its fundamental role, it is amazing how little we know about prediction. While epistemology and philosophy of science have extensively discussed how and how deep we can understand nature, little has been said explicitly about how this understanding can be used to predict the future. Induction – the primary tool for prediction – has been discussed as a way of cognition, but rarely as an element of prediction. There is no such thing as a science of prediction or even a theory of prediction. The only comprehensive treatise on prediction that I could find (Rescher, 1998) is a very recent one, together with some earlier works, that partially cover the issue (Reichenbach, 1938).

### Missing Awareness of Limitations in Prediction

The most striking incongruity in this context is probably the lack of awareness on the limits of prediction. Of course we know that predictions are never for sure, that our knowledge about the future is incomplete at best. We know that we are guessing when we predict the future, but we forget that fact in most daily applications, and we often expect that our own predictions as well as predictions from other people are virtually perfect. Except for some narrow applications (like weather forecasting, or insurance assessments), we rarely try to figure out the accuracy of predictions, and usually overestimate their reliability.

In fact, very little knowledge is available on the reliability of predictions. We know several mechanisms that threaten the reliability of predictions, but there is no comprehensive understanding about the sources of deviations and imprecisions in predictions. The reliability of predictions has been analysed *ex post* only in rare cases, as in weather forecasting, or forecasting in financial markets. The latter is a good example that even proven limitations on prediction do not prevent practitioners form continuously treating their own predictions as thoroughly reliable. Many of our predictions have a quality from vague to irrelevant, but we don’t care, and in most cases don’t even know it. The practical consequences of this ignorance are obvious – we put a great deal of effort into considerations that cannot deliver value, or at least not as much as we think.

Certainly, this survey cannot fill the gap. It is intended to obtain an overview over the available knowledge on prediction, as this is currently distributed among various disciplines, and to enable an informed assessment what we do already know about prediction, and where further research is required. The survey is limited to rational, or scientific prediction, where the statements of predictions can be reasoned by whatever rationality, and traced back to existing knowledge. Any form of prediction that is based on sources beyond rationality, like spiritual prophecy, religious divination, or other forms of enlightenment, will not be considered.

## Prediction Fundamentals

### What is a Prediction?

There is a bunch of words that have similar meaning of telling something about the future, but differ in detail. Among them are prophecy, divination, guessing, mantics, forecasting, and, of course, prediction. Their major differentiation is the source of insight: Prophecy is based on supernatural (divine) intuition, potentially based on an exegesis of divine rules. Divination relies on indicators that have no apparent relationship with the subject, like tea leaves, tarot cards, or astrology. As the connection between indicator and effect is not open to scrutiny, they do not qualify for scientific investigation.

A prediction is a statement on a future event, based on past experience. It claims a certain degree of correctness beyond sheer randomness, and a rational way of deduction. Prediction requires the ability of reasoning: it “must be cogent and substantiated independent from the predictor” (Rescher, 1998). Neither guessing nor prophecy qualify as prediction. Prediction is the rational way of looking into the future: trying to know about the future what can be known, given the experience from the past. A forecast is a special form of prediction, based on calculations. The word has been introduced and is nearly exclusively used in the context of weather forecasting (Orrell, 2007 S. 127)

Prediction does not depend upon predetermination: In order to predict something, there must be some underlying regularity where prediction can base upon, but determinism is not required. As we are never certain about anything in the future, prediction always has a probabilistic character, stating probabilities rather than certainties. Although some authors require an “affirmative statement with no probability left” (Rescher, 1998 S. 39), that does not remove the probabilistic character of the statement itself. The subsequent development may prove the statement to become true or false, without telling us something about the validity of our prediction – its intrinsic degree of correctness. This fact is closely linked to the question whether a probability can be objectively true, independent from subjective beliefs or individual sets of information (Savage, 1972).

The probabilistic character of prediction is independent from the question whether our uncertainty about the outcome is of ontological (due to fundamental uncertainty), or epistemological origin (due to our lack of knowledge). Both force us to bridge the existing uncertainty with a probabilistic statement. A certain element of chance together with some logical reasoning are ingredients of every prediction. A completely random event would be unpredictable, while a completely predetermined event would not require any prediction.

### The Reason for Predictions

We need predictions for our decisions. Whenever we make a decision, we intend to influence the future with a present action. Predictions enable us to act with a future target: They tell us how the possible (alternative) actions will expand into the future, allowing us to select the appropriate action according to its future impact. As we continuously make decisions even on the simplest actions, we also continuously make predictions on their potential future impact.

A particular characteristic of decision-related predictions is their conditional character. These predictions tell us what will happen *in case* we do a specific action. As we can always execute only one of several alternatives, most of these conditional predictions will never meet reality, as their conditions do not come true. We will never know whether these predictions might have come true or false, and in consequence cannot evaluate them against reality.

Predicting is the fundamental mechanism to make use of knowledge. Whatever we know, we can only make use of it by applying it to future situations, as this is the only sphere of time that can still be influenced. And the only way to influence the future in a targeted way is to predict what impact certain actions of us will have, using our knowledge for these predictions. In this respect, prediction (including the ability of prediction) is fundamental for the development of our knowledge society, and, in a broader sense, for any form of structured life. If meaningful prediction were impossible, life would be impossible as well.

## The Direction of Time

The specific character of predictions – that they always contain a certain degree of uncertainty, in contrast to experiences from the past – is connected to the fact that time has a direction. While space is completely isotropic – every direction has exactly the same properties – time is not symmetric: Past and future obviously are fundamentally different. While the past is unalterable, the future is always uncertain. The direction of time seems irreversible. This peculiarity is discussed extensively by various philosophers (Reichenbach, 1956) (Lucas, 1989) (Savitt, 1996), but still far from a convincing conclusion.

While it seems obvious from a macroscopic point of view that time has a definite direction and cannot be reversed, this property cannot be derived from microscopic structure. All known physical laws are symmetric in time, giving no reason for a directionality of time. The only exception is the second law of thermodynamics, which prevents entropy to decrease with time, but entropy is a statistical property, and the law holds only for macroscopic ensembles. The directionality of time thus seems to be an emergent property of macroscopic systems.

The asymmetry of time in thermodynamics and statistical mechanics, however, seems not to be originated from the second law of thermodynamics, as this law can be reformulated in a time-symmetrical way (Uffink, 2001). Rather, the definition of equilibrium states, which is fundamental to the definition of entropy, implies an asymmetry of time, as approximation to equilibrium is allowed, while deproximation is not. This couples the question about the direction of time to probability, as deproximation from equilibrium is highly improbable (the difference between probability of approximation and probability of deproximation is maximal at or near equilibrium), but not impossible. In a system near or at equilibrium, as long as it stays closed, there is no direction of time, as there is no distinction between past and present.

Strictly speaking, we do not have access to present and past when forming a prediction, but just to the present. Information from the past is available only through a memory that is read in present, while our methods to retrodict the past are symmetrical to those to predict the future. Under this pretext, the question arises why we do know more about the past than about the future. This seems to be the case exactly if the entropy of the system under investigation was lower in the past than it is in the present, and will be higher in the future than it is in the present (Barrett, 1992). It seems that an increase in entropy is necessary to allow access to information from the past. As a consequence, the direction in time seems to originate from a positive gradient in entropy.

## Causality

Causality is the way we use to explain the world. The fact that events have an effect on further events, and that there seems to be no effect without a cause, makes developments determinable. Already Kant (Kant, 1787) defined causality as an a priori property, that is a necessary element to experience the world. Reichenbach (1956) took causality as the reason for the directionality of time: there can be no effect prior to the cause. Pearl (2018) introduces causality into the toolkit of statistics, enabling us to investigate causality rather than just correlation.

However, causality also works only on macroscopic ensembles. Causality is thus an emergent property. Physical laws couple quantities, like force and acceleration in

Newton's law, but they do not distinguish between cause and effect. Albeit, Newton's third law (*actio = reactio*) explicitly denies it. In quantum mechanics, even effects preceding their causes become possible (entangled particles). It seems that physics does not need causality as a principle, but we do need it to better understand the world.

Causality enables us to make complex predictions, by linking events of different types to each other, and disentangle complex systems into simple chains of reaction. The concept of causality is so fundamental to our understanding of the world that we consider it to be an indispensable property of the world itself. Remarkably little has been published since on causality as a principle to categorize regularities and to derive predictions from them. Pearl investigated the verification of causal relationships by statistical methods (Pearl, 2018). However, further research shows that such verification is possible only for simple situations, where causal connections are sparse, and causal reactions immediate (Spirtes, 2000).

## Determinism and Chance

If our world is deterministic, then there is no room for a free will: Every action we do is inevitable, and we have no need for a prediction (as we make them nevertheless, also our predictions will be predetermined). If our world is not deterministic, there must be effects that cannot be completely explained by causality. Within the limits of the fundamental conservation laws, there must be processes that can have several different outcomes by chance alone, independent of their state. Although discussions are still ongoing, developments in quantum physics seem to manifest that there is in fact true randomness at the bottom of it (Colbeck, 2011).

There is a variety of definitions for chance in philosophy literature (Handfield, 2012). It is remarkable that chance is in most cases closely linked with probability, and most discussions about chance are considerations of probability, in particular on subjective probability, as defined by Savage (1972) and others. Chance in connection with complete uncertainty, without any connotation of probability, is rarely discussed, maybe because in this case there isn't anything left that can be known.

The oldest line of definition, which is associated to Aristotle, is a crossing of two independent lines of development, like accidentally meeting an old acquaintance. This is an event we would assign a low probability (which is difficult to estimate due to the multitude of factors involved), but by no means consider impossible. In terms of prediction, we typically would not consider events like these, because there is a huge number of possible constellations, each with a low probability. So, for practical reasons, we ignore such events in predicting, at least as individual events, while we might consider it quite probable that any (undefined) event of that sort will happen.

Another type of chance is based on our necessarily incomplete knowledge on initial conditions in connection with rapidly diverging developments, that blow up any initial ambiguity to complete unpredictability with time. This is the way we construct randomness (more precise: pseudo-randomness) in deterministic systems like computers. Systems that behave divergent in this way are called chaotic, and this type of chance is further discussed under that headline.

A third line of definition for chance entails events that cannot be linked to any regularity, or that follow a probabilistic regularity. In the first case, without any regularity, we have a fundamental randomness that is completely unpredictable. As per definition nothing can be known about such events before they happen, it is probably even difficult to identify them if they happen. There is little to be found in literature about fundamental randomness; randomness is mostly connected to probability. While we certainly cannot exclude fundamentally random events, there is no observation realized. Maybe we are not even able to recognise fundamental randomness, because we only can understand things through some sort of regularity.

Chance events that can be described in probabilistic terms are abundant both in the macroscopic as well as the microscopic view of the world. The probabilistic description entails random components, which are either due to limited knowledge of the initial conditions (like in statistical thermodynamics), or due to a fundamental random process, like in quantum mechanics (as far as we know). The decay of a state is naturally limited by conservation laws, but within these limits, the type, time, and spatial orientation of the decay is truly random. This randomness, however, is linked to objective probabilities, and their impact on macroscopic scale can be predicted with extreme precision.

While there is a lot written about chance, it is rarely described how randomness enters the world we observe. According to our experience, this occurrence of randomness will have to happen within the limits of conservation laws: Energy, momentum, spin, charge and other conserved quantities must not be violated by randomness. Remaining sources of randomness are the time and spatial direction of particle interactions, like the decay of a nucleus or the emission of a photon, as well as manifestations of quantum-mechanical states. All these random events have definite probability distributions, so they fall into the category of expectable randomness.

All these random events have in common that they occur in the microscopic states. Little discussion can be found under which conditions they may propagate into the macroscopic world. In most cases they don't: Singular microscopic events have a totally negligible macroscopic impact, and if there are large numbers of (similar) events, like in radioactive decay, their effect can be macroscopically described with high precision, due to the law of large numbers. However, there are settings where single events do make a difference: Otherwise, single energetic particles (as resulting from random events) could not be detected by devices like counting-tubes. Schrödinger's cat model is an example of a macroscopic situation depending on a single microscopic random event. Just under which conditions microscopic randomness may become macroscopic remains to be clarified.

## Scientific Investigation of Predictions

A scientific discipline of prediction does not exist. Of course there is a lot of knowledge on prediction, but it is scattered across various scientific disciplines and focused on details of the prediction process. Apart from a few quite recent exceptions (Orrell, 2007) (Rescher, 1998) there are no descriptions of prediction that bring all its aspects together. A complete and coherent theory of prediction is missing. As a consequence, our knowledge on prediction is unbalanced: Some aspects of prediction (like calculating probabilities) are fairly well explored, while others (like reliability of predictions) are still almost completely in darkness.

In epistemology, the process of induction that is fundamental to each prediction is investigated (Mill, 1872) (Skyrms, 1966), introducing the Principle of the Uniformity of Nature (Will, 1947). However, this line of investigation seems to remain a minor aspect in epistemology, maybe due to the fact that fundamental elements of induction cannot be covered sufficiently by pure logic.

Another issue covered by philosophy, in particular philosophy of science, is the nature of time, and its most particular feature, its directionality. Closely linked is the principle of causality and its relevance to understand the world (Reichenbach, 1930). Surprisingly, little is said about the role of chance in determining the direction of time.

The meaning of probability and the rules for calculating with probabilities are extensively discussed in statistics, or rather the philosophy of statistics. The Bayesian or subjectivistic approach to statistics (Savage, 1972) defines probability as a subjective concept and derives the rules that rationally confine subjective assumptions of probability. However, the subjective process of assessing the probability of a specific event is barely investigated (Good, 1966).

A relatively recent area of research is devoted to human behaviour under uncertainty, in particular in economic environments, called behavioural economics (Thaler, 2016). Research revealed how humans actually make predictions under uncertainty (Tversky, 1974), and how these deviate from pure logic, in favour of practicability (Ariely, 2009). While it is shown that the practical limits of our predictions fall short of rational possibilities, and that we are not aware of our shortfalls, nothing is said there about the fundamental limits of predictions.

Between about the years 1960 and 2000, a movement called futurology, or the science of the future, was active. While its target was to get a better understanding of the future, it did not precisely aim for predictions. Rather than to predict what will happen, futurology tries to describe scenarios that could happen, or may be likely to happen (Cornish, 1977).

Closely linked to futurology is the development and assessment of prediction techniques, like the DELPHI method (Dalkey, 1968), simulation techniques, or market-based methods (Bueno de Mesquita, 2009). These methods have in common that they refine and optimise existing subjective predictions, by providing appropriate incentives (market-based methods), or merging predictions from various individuals while preventing bias from group dynamics (DELPHI method and similar). However, none of these methods focuses on the generation of the subjective prediction itself – its existence is always presupposed.

Coming to the accuracy and reliability of predictions, research on this aspect is truly rare. Prediction techniques have been investigated and compared statistically (Mahmoud, 1984), but without any explanation why they may fail to predict correctly.

# The Process of Prediction

How do we form a hypothesis about the future? We observe patterns and correlations in the past. Starting from there, we form hypotheses about structures that may also hold into the future. In simpler cases, they are of a correlational character, either of the sufficiency type (when $\mathrm{A}$ happened in the past, often $\mathrm{B}$ followed. So if $\mathrm{A}$ happens, $\mathrm{B}$ may follow again) or the necessity type (when $\mathrm{B}$ occurred, often $\mathrm{A}$ was present. So if $\mathrm{A}$ is present, $\mathrm{B}$ may follow again). Correlational patterns don’t ask why, they just rely on repeated coincidences. Causal structures require more abstraction, but they enable us to predict more complex situations, where no direct observations are available, because causal elements can be concatenated. Causal reasoning is at the bottom of what we call intelligence.

When we observe a pattern in the present, we muster our portfolio of experiences and assumptions for matching ones, and use them to extrapolate the current situation into the future. We assign a probability to our proposal, related to (a) the degree of belief we have into the hypothesis and (b) the degree of match between observed pattern and hypothesis input. With confirming observations added to our experience we become increasingly confident in our hypotheses, while conflicting observations stimulate us to look for other, more promising patterns (Tetlock, 2015). However, research on behavioural economics suggests that we are inclined to prefer confirming against conflicting information, stabilizing our convictions even against evidence (confirmation bias).

### Predictive Inference

The patterns we observe are in the broadest sense correlations between two observables, though not necessarily linear ones. Patterns reduce the amount of information we need to describe a situation. They enable us to conceptualize experiences and to structure our memory. We are highly sensible to detect patterns in our observation – so much that we often detect patterns even when there are none (Silver, 2012) (Taleb, 2004).

The extrapolation from past observations to future expectations is an instance of induction. Induction leads from the properties of a sample to the properties of an entity, with the fundamental assumption that the sample is representative. For predictive induction, the sample is the past while the future is part of the unknown entity. All predictive inference is based on the principle of the uniformity of nature: „The future is like the past", or what has been valid in the past will also be valid in the future. If this assumption were false, any prediction would be useless (Skyrms, 1966).

Inductive reasoning can never lead to certainty, only to probability (Reichenbach, 1930). This leads to the vast literature on probability and its meaning, in particular its Bayesian and its Likelihood definition. As diverse as they are, they are united by a set of rules for consistency which they must fulfil (Good, 1966). For some special cases, like sequences with certain regularities, general rules for prediction have been defined, using the evidence of all past observations for prediction (prequential inference; Dawid 1984). In most practical cases, however, there are no fixed rules, and the process of developing a prediction statement remains vague.

### Reduction to Relevant Parameters

Prediction is a shortcut into the future: it can only fulfil its mission when determining and evaluating developments into the future is much faster than the development itself will evolve. Predictions must get ahead of the development; otherwise they are useless. This is only possible if the system in question is not described (and predicted) in all detail, but reduced to a set of relevant parameters. This reduction of a system to its characteristic properties is something that our brain is doing permanently. The character of categories has been examined in philosophy (for example, by Kant, 1787), but the mechanism our brain uses to define and distinguish categories on an appropriate level are still widely unknown.

There are two interrelated problems in reducing reality to a structure that can be characterized and predicted. One is the question of relevant parameters: Which macroscopic parameters deliver an adequate summary on the behaviour of all the microscopic properties of the system? How many of them do we need for a useful description? The answer will obviously depend on the scope of our prediction: Which level of detail do we need to make an appropriate decision – sufficiently fine to reveal relevant differences, but sufficiently coarse to enable a robust prediction? There will be no general answer, not even for a single system, as requirements may change from decision to decision.

The second problem is the question of the appropriate model for the selected parameters. No matter how many data are available (from the past): there are still infinitely many possibilities for rules to fit them, with any required accuracy. ("Give me four variables, and I'll fit you an elephant. Give me a fifth variable, and I'll let its tail wiggle.") All these rules coincide on the given events of the past, but of course most of them will not coincide when applied to the future. How to select the right one, that will also fit the future?

Of course, both problems are closely interrelated. "The future is like the past" is not valid for all properties of a system, but only for very specific ones, called laws. Only those can be carried into the future, all other properties will change. The identification of these invariants is at the core of the art of prediction, and the extrapolation of all other characteristics of the system around them. Finding and understanding invariants is at the heart of science since its beginning.

When looking for the most appropriate parametrization for a problem, we follow a principle that is known in science as Occam's razor: Take the simplest possible explanation. The idea behind Occam's razor: The simpler a rule is, the less probable it is that it fits to the past data just accidentally, and the more probable it is that it will fit to the future as well. However, simplicity as a property isn't easy to define or to measure. In practice, a balance between accuracy of fit (many parameters) and simplicity (few parameters) must be found (Handfield, 2012).

**Types and Variants of Predictions**

Within the scope of the principles described above there is a wide variety of prediction types that may be worthwhile for differentiation. Predictions may be conditional or unconditional, they may refer to a real or a fictitious future, or even to a fictitious past. They may claim a probability or a certainty, they may predict a single result or a range of possibilities, and they may refer to a certain event somewhere in the future, or to a chain of events that are predicted to develop over time. While the fundamental mechanisms of prediction are identical, the form of predictions have of course significant impact on their accuracy and reliability as well as on their verifiability.

Unconditional predictions give absolute statements on future situations, while conditional ones depend on certain events that also lie in the future, but earlier than the predicted ones. While the scope of unconditional predictions always will realize, enabling comparison of the prediction with reality, conditional predictions may remain completely hypothetical, in case their conditions never come true.

The most usual conditional predictions come with decisions: "what will happen, dependent on my alternatives for action?" There the conditions lie in the immediate future, while the subject of the predictions is positioned in the former future. In practice, all relevant predictions (leading to decisions) are conditional ones. Strictly speaking, all predictions are conditional, because all of them rely on certain intrinsic assumptions.

Similar to hypothetical conditional predictions are "postdictions" who address the hypothetical "what would have happened if" type of question (Rescher, 1998). Like conditional predictions whose conditions are not met, they are never confronted with reality, and thus cannot be checked for skill. Authors disagree, whether postdictions can be qualified as a type of predictions, but certainly they share some qualities with them (the uncertainty of the result), while they miss others (the potential of becoming real).

Postdictions cannot be used for decision-making, but they have their role in learning. Learning is trying to improve in making predictions with regard to decisions, by incorporating the experiences from past predictions. There it is not sufficient just to assess the prediction which conditions were met, and which subsequently became true or false. It is also necessary to assess the predictions for alternative conditions which were not met – these past predictions then become postdictions – in conjunction with the experience gained in the wake of the decision.

As the future is never certain, all predictions are in principle probabilistic: they give a probability distribution for the set of possible outcomes. However, in most practical cases, we don’t deal with probabilities, but choose the single most probable outcome as for granted, and ignore all other ones. That makes the prediction less accurate, but more easy to use. According to Rescher, a prediction has to be definitive, not probabilitstic: It requires a predictive question and an affirmative answer (Rescher, 1998). As predictive inference can only deliver probabilities, by this definition prediction also includes an element of decision that projects probability to an affirmative statement. Rescher however does not investigate or define how this decision should be done.

When the probability distribution for a prediction is given, then the functional to choose the best single value from the distribution depends on the scoring function that is applied to assess the accuracy of the prediction (Gneiting, 2011): When the accuracy is measured by mean absolute distance (between predicted value and realized value), the predicted value should be the median of the probability distribution; when the accuracy is measured by mean square distance, the predicted value should be the mean value of the distribution.

Caution has to be applied if the predicted probability distribution has a fat tail, as in this case the usual parameters (mean value and standard deviation) are misleading (Taleb, 2020). Although fat-tail distributions are widely present in our environment (for example in the volatility of stock prices), their particularities are barely considered in decision-making, leading to suboptimal decisions.

The subject of a prediction may be a single event somewhere in the future, or the development of a status into the future, in form of a time series. In case of a single event, the point in future time may either be fixed (stating the event for a definite point in time) or variable (stating it for some instance within a definite or indefinite future period). The prediction may target either a dichotomic (yes or no), a discrete (one of several values possible) or a continuous scope of possibilities. All these distinctions concern only the dimensionality of the prediction, but not its fundamental properties.

## Limits to Predictability

In the era of enlightenment, when more and more natural phenomena got to be understood, there was the ideal of a completely deterministic and predictable world. Causality was one of the fundamental principles, already postulated by Kant (1787). If one only knew the positions and momenta of all particles at a given time, then, applying the rules of nature, one could calculate all future developments with whatever precision desired. This principle was called Laplace’s demon, as Laplace was one of its foremost proponents (Laplace, 1902 S. 4). Determinism obviously negates any form of a free will.

While the discussion whether the world is deterministic, or contains elements of true randomness, is still ongoing, we know today that complete predictability is impossible, for several principal reasons. One, it is impossible to know the present state of the world beyond a certain level of accuracy, even for a small system and even for a higher “intelligence”, as Laplace calls his demon, because this is prevented by Heisenberg’s Principle of Indeterminacy. Two, the calculations required for a prediction may exceed the capabilities that are fundamentally available. Systems (deterministic or not) may be uncomputable in the sense that they cannot be reduced to characteristic parameters which can be calculated faster than the system develops itself. Three, the prediction of a system may have repercussions on the system itself that lead to irresolvable contradictions, as proved in Gödel’s Incompleteness Theorem.

Complete predictability would be contrary to our imagination of life, as well as complete unpredictability. Complete predictability would lead to a completely predetermined world, leaving no room for things like a free will, together with the opportunity for taking influence. At the same time, it is as important for our imagination of life that there is a certain level of predictability. A completely unpredictable world would be totally erratic, leaving no room for any directed development, human or other. Thus, it is the existing but limited predictability that makes the world bearing life as we know it. Monod (1970) pointed out that life and evolution require a coexistence and interaction of structure and chance. Structure (i.e., *nécessité*) is fundamental to obtain a directive development, and chance (i.e., *hasard*) is necessary to modify results for trial and error in evolution.

Apart from these principles that set fundamental limits to any prediction there are practical limits to predictability that set in much earlier: the current status of a system as well as its rules for development are only partially known, reduction to a few relevant characteristics may ignore elements that have influence on the development, limited capacities for analysis and calculation require approximations that may prove to be crude. There are three phenomena that are understood to limit predictions in theory as well as in practice: chance, chaos, and complexity. They are worth a closer look, together with the phenomenon of self-influence.

### Chance

Chance events are events that are not completely determined by the preceding state of the system, or, to be more precise, by the knowable facts about that preceding state (Loewer, 2001). They add a certain amount of randomness to the laws of system development. As far as we know, also chance events will have to obey fundamental laws, as those on preservation of energy and momentum. As a consequence, only some aspects of such an event can express randomness, while other aspects are predetermined. An example is the radioactive decay of a nucleus, where the instant and the direction of radiation are random, while the decay products and their energies are fixed.

All known events of this type have a fixed probability distribution which can be experimentally determined as these events have unlimited reproducibility. Their randomness is of a regular type: Whenever they are observed, they follow the same probability distribution. It is unknown whether also more fundamental chance events exist that cannot be predicted on a probability basis. Maybe we wouldn’t even be able to register such events, because our understanding of chance is so closely linked to probabilities (Edgeworth, 1922).

Chance events are not predictable *in concreto*, as for example the exact time of a decay, but they are well predictable in general, through probability distributions. The effect of chance on prediction depends on the view upon a system, as random behaviour on a microscopic scale may lead to a highly deterministic and precisely predictable behaviour on a macroscopic scale, due to the law of large numbers (Handfield, 2012). Examples are the activity of a radioactive lump, or the thermodynamic properties of a macroscopic body. However, it is also possible that chance on a microscopic level penetrates to the macroscopic level, as with a Geiger-Müller counter tube, or an electronic noise generator. Exactly which properties a system must fulfil that microscopic chance does not affect the macroscopic level seems to be still unknown.

### Chaos

The word chaos here shall not refer to the colloquial lack of order, but to systems with specific properties. Chaotic systems are deterministic systems that provide sensitive dependence on initial conditions, ergodicity, and a dense set of periodic points (Devaney, 1986). Well-known examples are eddies in fluids, weather development, or predator-prey systems. All chaotic systems are non-linear, but not all non-linear systems behave chaotic.

The property that affects predictability is the sensitive dependence on initial conditions. Tiny differences in initial conditions grow exponentially over time until the positions become completely delocalized. As the knowledge of initial conditions is always limited (at last by Heisenberg’s Principle of Indeterminacy), the inaccuracy of predictions grows with time until they become completely meaningless. Improving the accuracy of initial conditions does help very little, as errors grow exponentially, and the gain in prediction accuracy grows only marginally.

Chaotic systems thus may generate unpredictable behaviour from a deterministic structure, the unpredictability depending on the granularity (the resolution) of the view onto the system (Winnie, 1997). As the system is still deterministic, short sequences of development can be predicted quite accurately even in the future, provided the starting conditions are given. However, any longer period of development becomes stochastic under any choice of granularity, and also a long period of observation (in the past) does not help to obtain a better prediction for the future.

As differences typically develop exponentially in chaotic systems, a characteristic time (described by the Lyapunov exponent) exists for each specific system in which the error grows by a factor $e$. Compared with the resolution of the initial conditions (the relation between its accuracy and the full range of possible values), a maximum time can be determined until any prediction based on these initial conditions becomes meaningless.

### Complexity

There seem to be fundamental differences between systems that consist of a large number of similar elements, like gas molecules (a gas), sand grains (a heap of sand), bacteria (a colony), or neurons (a brain). Assuming that we can describe the properties and predict the behaviour of individual elements statistically, some systems enable a highly precise description of their macro-properties – the more precise the more elements they contain (the gas), while others tend to behave rather unpredictably (the brain) and show behaviour (thoughts) that cannot be reduced to the behaviour of its elements. Sand grains show a macroscopic behaviour that is somewhere in between, as it is mostly regular (the repose angle of a heap), but sometimes accidental (a single grain more can cause an avalanche). The law of large numbers may render macro-properties highly predictable, but only in certain cases.

Just what distinguishes predictable systems form unpredictabe ones is barely understood. Identity of the individual elements is not necessary for macro-predictability, as the heap of sand or a multi-chemicals fluid like crude oil show. The degree of interactions between the elements seems to be an important factor, as research on networks suggests (Watts, 2004). As the network density (the relative amount of interactions between individual elements) increases, systems seem to undergo a phase transition from predictable to emergent behaviour. To be precise, not only the density, but also the regularity and locality of the network seem to play important roles.

However, network properties cannot explain the whole set of behaviour of large-scale systems. As Wolfram proved in his investigations on cellular automata (Wolfram, 2002), even exceedingly simple systems with strictly limited interactions can show a wide variety of possible behaviours, many of them quite unpredictable, depending on seemingly minor variations of the interaction rules between elements. Again, there is no predictability from rules of interaction to macro-properties: Only full execution of the system development delivers the information on its macro-behaviour.

Systems that show macro-properties which cannot be derived from the comparatively simple properties of their individual elements are called complex, the respective macro-properties are called emergent. Although complex systems have received a lot of scientific attention recently, they are still barely understood (Mitchell, 2009). While we have a good general understanding whether a system is a complex one, there is no generally agreed definition of complexity, as well as no unit of measure for it. Some authors measure system complexity in terms of system entropy, but even then there are various alternatives to extend the concept of entropy form statistical mechanics to abstract systems.

Complexity obviously is a threat on predictability, as understanding the individual behaviour of the elements does not help to understand the behaviour of the macroscopic system. Some complex systems show regularity in their macroscopic behaviour that enables their prediction without looking at their elementary level. But in many complex systems, this regularity is brittle, meaning that there are spots of regularity (in space and time), which however change to different patterns at their (fuzzy) borders. A simple example are cellular automata of Wolfram's Class 4: "interesting" behaviour is at the brink between regularity and randomness. Financial markets behaviour may be another example: Spots of regularity evoke the illusion of predictability, which is subsequently contradicted by unexpected behaviour.

### Self-Influencing Predictions

Publishing a prediction may have an impact on its validity, if the prediction relates to human behaviour. While publishing the prediction of an eclipse or an earthquake will have no influence on their occurrence, publishing predictions on a market crash or on a political election will have impact on their subjects. In fact, this impact is the very reason for publishing predictions – like the predictions on climate development from IPCC, or the predictions of lifestyle impacts on individual health.

People react on published predictions by changing their behaviour, either in order to conform to the prediction, or to prevent it, depending on its attractiveness. This feedback circle may lead to strange consequences. First of all, private (unpublished) predictions will be different from public predictions, provided they both shall be as close as possible to the truth. In case of moderate (in particular: linear) repercussion, the public prediction can be adapted to include its own impact. With stronger repercussion, the situation may be dominated by the impact of the prediction, leading to self-supporting or self-killing predictions.

In sociology, this effect is well-known as self-fulfilling prophecy (Merton, 1948), whereas there it is claimed that the original assumptions were false. More precisely, one should request that the prediction is false conditional to its non-publication, but true conditional to its own publication. As a consequence, a prediction can never be proven to be a self-fulfilling one, as one of the conditions always remains unmet. Its counterpart, the self-denying prediction, may lead to a situation that is logically impossible to predict, at least publicly: If the public prediction of $\mathrm{A}$ leads to non-$\mathrm{A}$, and the public prediction of non-$\mathrm{A}$ leads to $\mathrm{A}$, no true public prediction is possible, except the meta-prediction that every public prediction on $\mathrm{A}$ will be false.

A still more intricate reflexivity may come into action when predictability as a property (independent of the contents of an individual prediction) obtains importance. This happens in collaborative as well as in competitive situations, regarding the predictability of participants' actions. Collaborative situations happen in leadership, where predictability is a prerequisite for trust and leadership effectiveness: Participants will behave predictable, just for the sake of trustworthiness.

Competitive situations are dominant in markets, where any gain beyond normal market development is at the expense of someone else's loss. If participants behave predictably, other participants may use it to their advantage at the predictable persons' disadvantage. An example are short squeezes as on Volkswagen in 2008 or GameStop in 2021: Short sellers become predictable, as they have to close their positions at any cost. As a consequence, market participants will behave to some degree unpredictable, just for the protective effect of unpredictability. This way, markets become fundamentally unpredictable, at least to a certain extent. Again, this simple consequence seems not to have penetrated to market experts.

## Quality of Predictions

Given the omnipresence of predictions in decision-making, one should expect that the quality of predictions is thoroughly investigated, being of utmost importance. However, this is not the case. Predictions are rarely tracked whether they come true in the future, and our memory is notoriously unreliable on this task, emphasizing what has come true, while suppressing what has not. Scientific investigation on the potential as well as the real quality of prediction is rare as well.

In most cases predictions are given without a notion of their quality. In daily life, we sometimes add qualitative predicates to predictions, like "I am convinced that" …, or "I am in doubt that …", and in disputes about conflicting predictions, our degree of conviction is challenged. But we rarely undergo a true effort to assess their quality, or even quantify it. Quantitative forecasts are sometimes given together with a level of confidence, like in quality investigations ("99% of all produced items will be within the norm”), or with a range of accuracy, like in weather forecasts (“Maximum temperature will be at 22°C, plus or minus two degrees”).

While the fact of limited predictability is widely accepted, little is known about the effects of these limitations, as well as of the quality of predictions in general. With many predictions, in particular in social environments like business or politics, we have no idea how reliable they might be. The questions where the limits of our predictions are and whether we might be able to quantify them seems to be still open; little research can be found about it (Mahmoud, 1984).

Quality of predictions has two dimensions, which shall be treated separately: The accuracy of a prediction refers to its determination, to the amount of additional information it delivers provided it is reliable. The reliability of a prediction refers to its correctness, to the degree of certainty it will conform with reality. The two dimensions interact with each other: The more accurate a prediction shall be formulated the more challenging it will be to make it reliable.

### Accuracy of Predictions

The accuracy, or skill, of a prediction is a measure how strong the prediction focuses on a few of all possible events. The higher the accuracy of a prediction, the more directive and useful it is for decisions. A prediction that all possible outcomes are more or less equally probable (low skill) is in most cases much less useful than a prediction that selects one outcome to be the most probable (high skill). The skill is however independent from the question whether the prediction is reliable.

As a prediction is a probability distribution over the set of all possible events, its skill can be expressed as the concentration of probability distribution on few events. With unordered events (like people elected), the Herfindahl index is an appropriate measure, while with ordered events (like a temperature or an amount of money) the bandwidth (standard deviation) is expresses its skill level. Error statements with predictions are usually defined as standard deviation, so that the total probability for the interval of mean plus/minus error is 65%. To evaluate the skill of a prediction the bandwidth must be related to the total bandwidth of possible events.

As predictions rarely operate in isolation, skill can also be a relative improvement compared to a current base of knowledge. When it is known that temperature in July always ranges at 20°C plus or minus ten degrees, a temperature forecast of 26°C with a bandwidth of one degree has a relative skill of ten. In dynamic systems, the occupied volume of phase space may serve as a reference to the basic range of possibilities.

As errors in the prediction add up with time, the skill of a prediction should decrease with increasing time into the future – in best case inversely proportional, in many cases however exponentially. In particular with chaotic systems (like the weather), errors grow exponentially (with the Lyapunov exponent), and the skill rapidly decreases with prospect in time. The same should be valid for predictions in complex systems with retarded feedback cycles (like earth climate), as the feedback causes exponential growth of errors.

### Reliability of Predictions

The reliability, or validity, of a prediction is a measure how close the prediction comes to reality. In order to assess its reliability, one must compare the statements of the prediction with the truth, what is of course only possible after the fact. But comparing a single prediction with the true result does not deliver any insight: As the prediction is a probability distribution, even a highly improbable event may obtain. An inference to the validity of the prediction is not possible. There is no way out: A prediction on a *singular* event cannot be assessed against reality (Reichenbach, 1930). Conditional predictions, whose conditions do not realize, as well as all postdictions, which are always fictitious, cannot be compared to reality at all.

The situation changes when we have multiple events under similar conditions, so that we can compare multiple predictions (of the same type) with their respective events. This may be the case when the prediction does cover a chain or development of events rather than a single event, or when similar events are repeatedly predicted by the same method. In this case we can compare the true distribution of events with the predicted one, gaining statistical evidence on the validity of the predicted probability distribution. This is in particular applicable on microscopic events that can be repeated without limits, like collisions of elementary particles, or decays of atomic nuclei. As expected, any evidence on predictions by comparison with reality is of statistical nature.

Comparing predictions with reality as described assesses how valid the predictions have been, but not how valid they could have been, as the comparison inevitably contains information that is only available after the event. What does a low validity tell us then? Maybe it was still the best possible method of prediction, as the event was barely predictable, maybe a different method of prediction would have yielded a much higher validity. In order to assess the *relative validity* of a prediction, we must compare it with the best prediction possible given the available information. It is however unclear whether such an ideal prediction exists at all (Handfield, 2012 S. 19).

If we would compare all possible methods of prediction (which is of course not feasible in practice, as there are infinitely many ones) in their validity on a series of events, there would always be at least one that fits perfectly, by sheer randomness. In order to distinguish systematic from random validity, it seems to be necessary to introduce another criterion, the simplicity of the prediction rule. This is exactly what we do when we are looking for appropriate rules for prediction. The ideal prediction is then the best possible combination of simplicity and validity. However, this replaces an old problem by two new ones: (a) how can we measure simplicity, and (b) how do we balance the measured simplicity with validity?

If we could identify an ideal prediction, we could compare its probability distribution with our real prediction, and the degree of overlap would deliver us an assessment how good we have made our job as predictors, and how close we came to the best possible prediction. The skill of the ideal prediction would be a measure of predictability: the degree to which an event is predictable at all. A completely unpredictable event (like the throw of a fair dice) would have zero skill (equal distribution) on its ideal prediction. This kind of unpredictability would of course be subject to the knowledge of the predictor, as the ideal prediction depends on the available information.

One could try to find an objective kind of benchmark for the prediction by comparing it with the “true” probability distribution at the time of prediction, given the complete state of the system. In case of a deterministic world, this would be simply the final event, whereas in a non-deterministic world it remains a probability distribution reflecting the impact of true randomness between prediction and event. As we don’t even know whether our world is deterministic or not, we are far from having a clue how we could obtain the true probability distribution.

### Uncertainty and the Degree of Confidence

There is an element in every prediction that we cannot express in the probability distribution, and that is the level of confidence we have in our prediction, or, conversely expressed, the level of uncertainty that is prevailing (Dequech, 2011). Consider an urn with 100 balls, black and white, and we shall predict the colour of a ball to be drawn. If we knew there were 50 white and 50 black balls, we would predict even chances, as well as we would if we had only seen two of the balls, one black and one white. The uncertainty however would be zero in the first case, and high in the second case (Ellsberg, 1961).

This uncertainty, also called ambiguity, can be expressed through the probability of the selected probability distribution. The uncertainty, or rather its reverse, the certainty, is related to the skill (concentration) of the probability distribution on the set of all probability distributions on the set of all possible events. In the first case above, the selected probability distribution is certain, while in the second case nearly all probability distributions (except 100% white, or 100% black) are possible. How they will be weighed in a meta-probability distribution depends on further assumptions one has to postulate. In case one considers every combination equally likely, the meta-probability of the 50/50 variant is just 7.9%.

It has been shown that calculating with meta-probabilities, or probabilities of probabilities, makes sense (Marschak, 1975). As the set of all probability distributions on $n$ possible events belongs to an $(n-1)$-dimensional space, mathematics become somewhat abstract and unsuited for practical purposes. In addition, there are situations with uncertainty where we even cannot form an opinion about probability distributions (Davidson, 1991).

Attempts have been made to form a theory of ambiguity that is not based on probability, which is for example relevant in game theory (Epstein, 1996). Ambiguity can be defined as a mathematical entity, like probability, but with different rules and attributes (Fishburn, 1993). Although ambiguity can be captured in an axiomatic system, it remains unclear how it can be determined quantitatively.

## Conclusive Remarks

A lot of research has been performed on aspects of prediction, although most of it does not explicitly refer to prediction. It is spread across a variety of scientific disciplines, from philosophy of science to theoretical computer science, from Bayesian statistics, to psychology, from thermodynamics and quantum physics to economics. This lack of focus makes it difficult to obtain a comprehensive overview on the current status of knowledge.

Several important aspects of prediction and predictability seem to be barely investigated at all. Among these are the subjective process of estimating a probability, the validity of predictions we regularly make, in particular in economics, and the fundamental question of predictability, in particular regarding complex systems. Further research in these directions could gain significant impact on the quality of predictions in our life and business.

# References

**Ariely, Dan. 2009.** *Predictably Irrational.* New York, NY : Harper, 2009. ISBN 978-0-06-201820-5.

**Barrett, M. and Sober, E. 1992.** Is Entropy Relevant to the Asymmetry Between Retrodiction and Prediction? *British Journal of Philosophy of Science.* 1992, Vol. 43, pp. 141-160.

**Bueno de Mesquita, B. 2009.** *The Predictioneer's Game.* New York, NY : Random House, 2009. ISBN 978-0-8129-7977-0.

**Colbeck, R., and Renner, R. 2011.** No Extension of Quantum Theory can have Improved Predictive Power. *Nature Communications.* DOI: 10.1038/ncomms1416, 2011.

**Cornish, Edward. 1977.** *The Study of the Future. An Introduction to the Art and Science of Understanding and Shaping Tomorrow's World.* Bethesda, ML : World Future Society, 1977. ISBN 0-930242-03-3.

**Dalkey, N.C. 1968.** *Predicting the Future.* 1968.

**Davidson, Paul. 1991.** Is Probability Theory Relevant for Uncertainty? *Journal of Economic Perspectives.* 1991, Vol. 5 No. 1.

**Dequech, Daniel. 2011.** Uncertainty: A Typology and Refinements of Existing Concepts. *Journal of Economic Issues, Vol. XLV.* Sept. 2011, pp. 621-640.

**Devaney, Robert L. 1986.** *An Introduction to Chaotic Dynamical Systems.* Menlo Park, CA : Benjamin/Cummings Publishing Co., 1986. ISBN 0-8053-1601-9.

**Edgeworth, F.Y. 1922.** The Philosophy of Chance. *Mind.* Vol. 31 No. 123, 1922.

**Ellsberg, Daniel. 1961.** Risk, Ambiguity, and the Savage Axioms. *Quarterly Journal of Economics Vol.75 Issue 4.* 1961, pp. 643-669.

**Epstein, L.G. and Wang, T. 1996.** Beliefs about Beliefs without Probabilities. *Econometrica.* 1996, Vol. 64 No. 6.

**Fishburn. 1993.** Axioms and Algebra of Ambiguity. *Theory and Decision.* 1993, Vol. 34, p. 119.

**Gneiting, Tilmann. 2011.** Making and Evaluating Point Forecasts. *Journal pf the American Statistical Association.* 2011.

**Good, I.J. 1966.** How to Estimate Probabilities. *Journal of the Institute of Mathematics and its Applications.* 1966, Vol. 2, pp. 364-383.

**Handfield, Toby. 2012.** *A Philosophical Guide to Chance.* Cambridge : Cambridge University Press, 2012. ISBN 978-1-107-01378-0.

**Kant, Immanuel. 1787.** *Kritik der reinen Vernunft.* Riga : s.n., 1787.

**Laplace, Pierre-Simon Marquis de. 1902.** *A Philosophical Essay on Probability (Translation by Truscott/Emory).* New York : John Wiley & Sons, 1902.

**Loewer, B. 2001.** Determinism and Chance. *Studies in Hostory and Philosophy of Modern Physics.* 2001, Vol. 32 No. 4.

**Lucas, J.R. 1989.** *The Future.* Oxford : Basil Blackwell, 1989.

**Mahmoud, E. 1984.** Accuracy in Forecasting: A Survey. *Journal of Forecasting.* 1984.

**Marschak, Jacob et al. 1975.** Personal Probabilities of Probabilities. *Theory and Decision.* 1975, Vol. 6, pp. 121-153.

**Merton, Robert K. 1948.** The Self-Fulfilling Prophecy. *The Antioch Review.* 1948, Vol. 8 No. 2.

**Mill, John Stuart. 1872.** *On Induction.* 1872.

**Mitchell, Melanie. 2009.** *Complexity. A Guided Tour.* New York, NY : Oxford University Press, 2009. ISBN 978-0-19-512441-5.

**Monod, Jacques. 1970.** *Le hasard et la nécessité.* Paris : Editions du Seuil, 1970. ISBN 3-492-01913-7.

**Orrell, David, Ph.D. 2007.** *The Future of Everything - The Science of Prediction.* New York : Thunder's Mouth Press, 2007. ISBN 978-1-56858-369-3.

**Pearl, J., Mackenzie, D. 2018.** *The Book of Why.* London : Penguin Science, 2018. ISBN 978-0-141-98241-0.

**Reichenbach, Hans. 1938.** *Experience and Prediction.* Chicago : The University of Chicago Press, 1938.

**—. 1930.** Kausalität und Wahrscheinlichkeit. *Erkenntnis.* 1930, pp. 158-188.

**—. 1956.** *The Direction of Time.* 1956.

**Rescher, Nicholas. 1998.** *Predicting the Future. An Introduction to the Theory of Forecasting.* New York : State University of New York Press, 1998.

**Savage, Leonard J. 1972.** *The Foundations of Statistics.* New York : Dover Publications, Inc., 1972. ISBN 978-0-486-62349-8.

**Savitt, Steven F. 1996.** The Direction of Time. *British Journal of the Pilosophy of Science.* 1996, Vol. 47, pp. 347-370.

**Silver, Nate. 2012.** *The Signal and the Noise.* New York : Penguin Random House, 2012. ISBN 978-0-141-97565-8.

**Skyrms, Brian. 1966.** *Choice and Chance: An Introduction to Inductive Logic.* s.l. : Dickenson Pub. Co., 1966.

**Spirtes, P. et al. 2000.** *Causation, Prediction, and Search.* 2000. ISBN 978-0262194402.

**Taleb, N.N. et al. 2020.** On single point forecasts for fat-tailed variables. *International Journal of Forecasting.* 2020, Vol. 38 No. 2.

**Taleb, Nassim Nicholas. 2004.** *Fooled by Randomness.* London : Penguin Books, 2004. ISBN 978-0-141-03148-4.

**Tetlock, P., and Gardner, D. 2015.** *Superforecastiong - The Art & Science of Prediction.* London : Penguin Random House, 2015. ISBN 9781847947154.

**Thaler, Richard H. 2016.** *Misbehaving - The Making of Behavioral Economics.* New York, NY : W. W. Norton, 2016. ISBN 978-0-393-35279-5.

**Tversky, A., and Kahnemann, D. 1974.** Judgment under Uncertainty: Heuristics and Biases. *Science.* 1974, Vol. 185.

**Uffink, Brown and. 2001.** Origins of Time Asymmetry. *Stud. Hist. Phil. Mod. Phys.* 2001, Vol. 32 No.4, pp. 525-538.

**Watts, Duncan J. 2004.** *Six degrees - The Science of a Connected Age.* London : Vintage Books, 2004. ISBN 978-0099444961.

**Will, Frederick L. 1947.** Will the Future be like the Past? *Mind.* Vol. 46 No. 224, 1947, pp. 332-347.

**Winnie, John A. 1997.** Deterministic Chaos and the Nature of Chance. [book auth.] J., and Norton, J.D. (ed.) Earman. *The Cosmos of Science - Essays of Exploration.* Pittsburgh : University of Pittsburgh Press, 1997.

**Wolfram, Stephen. 2002.** *A New Kind of Science.* 2002.