

How Real Are Statistics? Four Possible Attitudes*

BY ALAIN DESROSIÈRES

REFERENCE to “reality” is a commonplace among both producers and users of statistics. This “reality” is understood to be self-evident: statistics must “reflect reality” or “approximate reality as closely as possible.” But these two expressions are not synonymous. The very notion of “reflection” implies an intrinsic difference between an object and its statistics. In contrast, the concept of “approximation” reduces the issue to the problem of “bias” or “measurement error.” Thus even these two common expressions, generally used without regard to consequences, tell us something important: a critical reexamination of this notion of “reality” is, for statisticians, an efficient way to reconsider the deepest-rooted but also the most implicit aspects of their daily work—precepts, tricks of the trade, justifications provided to users, and so on. This paper argues that the way in which producers and users of statistics talk about “reality” is informed by the fairly unconscious intermingling of several attitudes to reality. The mix of these attitudes and the links between them vary according to the circumstances—or, rather, according to the specific constraints prevailing in different situations. As it happens, the field of business statistics offers a representative spectrum of these various possible

(*) This paper, which has been translated by Jonathan Mandelbaum, is a revised version of an article originally published in French in *La Lettre du Systeme Statistique d'Entreprise*, the in-house bulletin of INSEE (the French national statistical institute). The “business statistics” discussed are those supplied by national statistical institutes to describe the activity of enterprises. The accounting and statistical data produced and used in firms are discussed here only with regard to their use by statisticians and, subsequently, by economists.

attitudes to reality. However, we cannot say—despite their differences—that some are better than others since each is so closely associated with situational constraints specific to particular phases of the statistician's technical, administrative, or managerial work.

Four possible attitudes to reality (among others) will be discussed. First, we will describe them in a "pure" (hence certainly exaggerated) form. Next, we will see how they are applied to concrete situations. Each attitude has its *language*, that is, a register of words, requirements, and arguments; these are consistent, but difficult to interlink if one shifts from one attitude to another. Here is a list of the four attitudes ranked by "obviousness" (at least for statisticians; for other communities the order would, no doubt, be different):

- metrological realism;
- pragmatism of accounting (which may be "national" accounting);
- use of material from a database for argumentative purposes in social life or in quantitative economics and social sciences; and
- the explicit admission that the definition and coding of the measured variables are "constructed," conventional, and arrived at through negotiation.

Of these four attitudes, the first three may be qualified as "realistic," but each has different *reality tests*—that is, ways of verifying and articulating the substance of that reality and its independence from observation. The fourth attitude, instead, emphasizes the conventional and social character of statistical variables, and may thus be labeled as "constructivist." It mainly comes into play in situations marked by discontinuity, controversy, and innovation. We will now examine each attitude individually, spelling out their languages, origins, and conventions.¹

Sampling and Confidence Interval

Metrological realism derives from the theory of measurement in the natural sciences that is complemented, in the social sciences,

by the sampling method. The object to be measured is just as real as a physical object, such as the height of a mountain. The vocabulary used is that of reliability: accuracy, precision, bias, measurement error (which may be broken down into sampling error and observation error), the law of large numbers, confidence interval, average, standard deviation, and estimation by the least-squares method (Stigler, 1986; Hacking, 1990). This terminology and methodology was developed by eighteenth-century astronomers and mathematicians, notably Gauss, Laplace, and Legendre. The core assumption is the existence of a reality that may be invisible but is permanent—even if its measurement varies over time. Above all, this reality is independent of the observation apparatus. In a sense, this is the dream of the statistician and the specialist in quantitative social sciences: the possibility of making the metrology of these sciences equivalent to the proven methodologies of the natural sciences. This may be seen as a benchmark, an ideal to which statisticians aspire, despite an awareness that their objects do not display all the properties assumed by the methodology. We could describe this as the lost paradise of the social sciences, which would have liked to have been endowed with the same persuasiveness as the natural sciences of the nineteenth century.

In this endeavor to connect the methodologies of the social sciences and the natural sciences, one element plays a crucial role: the law of *large numbers*, which is the basis for probability formulas and the resulting convergence theorems. The “law” serves as an operator for the transformation and transition from the world of observations to the world of generalization, extrapolation, and forecasting. Its hybrid nature is summed up in the famous and revealing quip dating back to the nineteenth century: astronomers and physicists believe the law to be a theorem demonstrated by mathematicians, whereas mathematicians think the law has been proved by the results of repeated testing. In fact, the possibility of conducting multiple and mutually independent observations of comparable objects is the foundation of a statistical methodology initially developed to study populations of human individuals or households.

Official statisticians in charge of business statistics are specifically exploring the gaps between the methodology of business statistics and the metrological ideal of the classical age. One of these gaps is due to the heterogeneity of the population, which is so great that the largest firms have to be “profiled” individually using monographic methods quite different from the established statistical method. There is also the difficulty of defining and classifying the “statistical units”: establishments (local legal units), enterprises, groups, and so on—in other words, the very units that need to be “counted” and “measured.” Last—and most important for our discussion—there *already* exists a quantification system internal to the world of enterprises: double-entry bookkeeping ledgers recording receipts and expenditures, along with claims and debts. The system, which was conceived in the sixteenth century, predates the age of statistical observation and is thus far older than the metrology of eighteenth-century astronomers (Hopwood and Miller, 1994).

Double Entries and Balancing the Books

Business accounting is predicated on concepts of reality and proof that underscore its profound differences with the metrology of natural science. To begin with, the “equivalence space” is composed not of physical quantities (space and time), but of a general equivalent: *money*. Money allows the circulation of claims and debts (via bills of exchange); it serves to determine profits by measuring receipts and expenditures and by assigning a “probable” value to claims and debts. It should be noted that this “*subjective*” probability—used, for example, to assess a doubtful loan—is different from the “frequentist” or “objective” probability on which classical metrology bases its computations.² As we can see from this crucial example of the calculation of doubtful loans, business accounting is a rich and dense social practice that seeks to achieve consistency and coordination in evaluations, actions,

and decisions, either for a single player over time, or for several players whose relationships need to be regarded as fair and hence reproducible. In this test of an accounting reality, double-entry bookkeeping plays a role similar to the repetition of observations in classical metrology. The requirements and tests involved in "balancing the books" are analogous to the regularities and "normal" distributions of repeated observations of the same object.

The tension between these two forms of quantification—one derived from scientific metrology, the other from business-accounting practices—has been sensitively analyzed by Oskar Morgenstern in his famous work, *On the Accuracy of Economic Observations* (1963) (translated into French with the unfortunate title of *The Statistical Illusion: Precision and Uncertainty of Economic Data*). Morgenstern is utterly dedicated to establishing a measurement system for economics that is just as rigorous as that of the other sciences. For this purpose, he examines the information provided by business accounts. He studies the status of "errors"—often regarded as "falsifications" or "lies"—that are to be found in these documents. Morgenstern distinguishes—for example, in balance sheets—between the items that are verified and identified without ambiguity (such as a cash position) and those that are merely estimated and shrouded in uncertainty, a practice justified by the need for prudence:

It will, however, be noted that a "lie" is, in this context, not a simple and obvious concept. It is unmistakable when a false cash position is willfully given or physical inventories are reported that do not exist. But when deliberately a more optimistic attitude is taken in interpreting the success of a year's operation—for example by small amortization—it will be hard to classify this statement as a "lie." Instead, it may be viewed as an error in judgment and as such be proved or disproved by later events (1963).

This analysis effectively indicates the constraints that weigh on the preparation of "fairly presented" (United States) or "true and

fair" (United Kingdom) accounts, to use the stock phrases. The "reality" thus described is connected to a set of "wagers on the future" that do not qualify as "lies" or even as "bias" (in the sense in which the term is used in classical inferential statistics). In the case of business accounting, which is a quantitative tool that underwrites social links, "reality" is inseparable from the *trust* inspired by the numbers compiled in the accounts—what Ted Porter calls "trust in numbers" (1995).

Another aspect of this "accounting" realism is surprising from the standpoint of scientific metrology: these documents display no radical discontinuity between the past, present, and future—between the closed accounts, the current accounts, and the budget forecast. Rather, we find an incremental shift from one to the other, since all are informed by the same conceptual frameworks, and are designed as tools for assessment, action, and decision. A strikingly similar continuity is found in the daily work of national accountants: they too concurrently manage economic budget forecasts, followed by preliminary, semirevised, and revised accounts.

The distinctive density and specificity of accounting practices are well known to—yet sometimes forgotten by—the business statistician imbued with a metrological culture. Since the 1970s, business statistics has gradually imported the sampling method, previously tested in social statistics. The result has been a juxtapositioning, followed by a close interlinking, of these two forms of quantification, despite their different origins and principles. One of the problems in using business accounts for statistical purposes is the impossibility of checking the uniformity of accounting procedures. In particular, the numbers reported in business accounts are already the result of an initial aggregation of myriad elementary operations that are entirely beyond the statisticians' grasp. The verification of this "first-level" accounting work is the job of professionals such as certified accountants or "auditors"—and they too, but in another way, have introduced the probabilistic method of verification on samples of accounting documents

(Power, 1992). In this case, however, the sampled unit is an accounting document among those of a single enterprise, not an entire enterprise among all those in an industry.

The issues raised by the linkage between the two methodologies—one statistical, the other accounting-based—were clearly visible in other circumstances: the establishment of *national accounts*, for example in France in the 1950s and 1960s. National accounting has partly inherited the reality tests derived from business accounting: its variables were defined (and, more important, interdefined) a priori; they were recorded in consistent, comprehensive, and theoretically balanced tables, where they were arranged in rows (transactions) and columns (agents). Disparate sources were reconciled (often for the first time) to compile these tables. The final—but not least significant—resemblance between national accounting and business accounting is that both tools were action- and decision-oriented: the national accounts were intended as “monitors” of macroeconomic policies, in the same way as the balance sheet and income statement provide guidance for the company executive. The accounts form a whole, explaining why the so-called reliability constraints are not identical to those of a pure “metrological” measurement of an isolated variable, whatever it may be.

These interrelated specificities of the reality forms, the rationales used to quantify them, and their applications explain the ritual debates between statisticians and national accountants in the statistical institutes in the 1950s and 1960s. The statisticians, trained in the methodology of fine-tuned sampling, were wary of what they regarded as the sometimes cursory practices applied by the national accountants to estimate some of the variables in the accounts (such as changes in inventories or trade margins). The only justification offered by the national accountants was pragmatic, based on *purpose*: their approach, they argued, was needed for policymaking. Even low-quality estimates (provided they are contained within the overall system of constraints created by accounting balances) are preferable to no estimates at all. This

defense was unacceptable to the methodologist statistician. The differences between the two camps may seem purely sociological, explained only by their belonging to distinct socioadministrative networks: national accountants are closer than statisticians to economic-policy decision making and implementation. But each group had its own approach to the orchestration of reality. Statisticians force themselves, sometimes ascetically, to be nothing other than methodologists specializing in good metrology; at the opposite end stands the systemic reality of the accountant, which only makes sense within the framework of policy guidance and monitoring.

Proof in Use

We can distinguish these two initial forms of realism from a third form, which is implicit or explicit for the user outside the two universes where these two concepts of reality reign: the official statistical departments, and enterprises producing their accounts in a continuous flow. Typically, this user is a researcher, or a social player in the administrative, political, or economic sphere (notably in other enterprises). For users in this third group, "reality" is nothing more than the database to which they have access. Normally, such users do not want to (or cannot) know what happened before the data entered the base. They want to be able to trust the "source" (here the database) as blindly as possible to make their arguments—backed by that source—as convincing as possible.

We are confronted here with a "metadata paradox."³ From a *normative* standpoint, users must be given a maximum of detailed information on the data-production process. It is also true that, from a *descriptive* standpoint (i.e., without passing judgment), many users do not welcome an abundance of metadata: "ideal" information is that which seems self-sufficient, without footnotes to interfere with the message. This unfamiliar field—generally

mentioned (if at all) with a touch of irony—would deserve to be studied in terms of *cognitive economics*—that is, in terms of the yield (cost-effectiveness) of a statistical argument. These issues cannot be dealt with in purely normative terms such as: *we must* supply metadata (which is, of course, true). A sociology of the social uses of the statistical argument remains to be developed. It would be especially helpful in exploring “quality” issues, often discussed in strictly normative fashion.

The examination of uses—the argumentative contexts in which statistical data are introduced—reveals a realism of the third kind, based on the consistency and plausibility of the results obtained. This is especially visible in the development of econometric methods. They provide many tools or arguments for an *internal validation* of a data set, without the need (at least in “normal” conditions) to examine the prior stages (the recording and coding of these “data”). Alas, contrary to what the etymology of the unfortunate term “data” suggests, very few “data” are actually “given”: they come with a high price tag, in both financial and cognitive terms. Coding always involves sacrificing something with a view to the subsequent use of a standardized variable, that is, an *investment in form* (Thévenot, 1984). This is comparable to the industrial investment needed to produce the standardized and mutually validating parts of a machine.⁴

It is striking that the statistical institutes of certain countries (such as the Netherlands) place such emphasis on the concept of integration or, in other words, the achievement of consistency among the statistical data produced by the institute’s various departments. This goal is promoted even if it entails corrections and adjustments of the “raw data”—of what may appear to be the “reality in the field.” This substitution of a validation internal to the statistical system for a more external validation—with respect to a putative field—is relevant to our exploration of the different forms of realism. It converges toward the realism desired by users: the internal consistency of their data set. Indeed, it is precisely in such terms that the most staunchly “integrationist” statisticians

justify their extensive efforts to whittle down any anomalies: "our users would not tolerate our giving them inconsistent data." This point of view actually resembles that of national accountants, who—at least for macroeconomic variables—adopt the same *use-oriented* approach in their macroeconometric models, and whose *book-balancing equations* represent a major constraint.

The Three Realisms: Summary and Comparison

To conclude this review of the three ways of being "realistic," we can try to compare the reality tests that characterize them. The first is that of the pure statistician, trained in a probabilistic culture; this test is a remote descendant of the theory of measurement errors in eighteenth-century astronomy. The latter science regarded observations, however numerous, as independent of one another. The object's reality and substance are proved by the normal distribution of error-ridden observations. A confidence interval can be presented in probabilistic terms. This metrology was imported into the social sciences through the sampling method. The astronomers' basic hypotheses have been transposed to this new universe: statistical units are "homogeneous" (but the definition of this term is ambiguous, notably in the case of enterprises); the distributions of the variables studied do not diverge too significantly from the normal curve; and the law of large numbers can be applied. The central notion of this transfer is that—just as with the distributions of astronomical observations—the computed moments (averages, variances, correlations) have a substance that *reflects an underlying macrosocial reality, revealed by those computations*. Therein lies the essence of metrological realism.

Accounting realism is altogether different. It is internal to the enterprise. Accounting is already an aggregation, in monetary terms, of heterogeneous elements. Some of these are measured with certainty (cash positions, at least when the monetary unit is

reliable and stable); others are estimated with uncertainty and imply a degree of subjective probability. The choices of these values are guided by the two potentially contradictory requirements for prudence and for communication with other players. The main underpinnings of this approach are the "interdefinition" of the variables and their recording in balanced tables. Its overall realism is more pragmatic—in the sense in which we describe someone as "realistic"—than metrological. In any event, the two orders of realism are closely intertwined, and this combination forms the core of a practice of construction and use of quantitative data that differs from the statistician's approach.

The reality judgments of users represent yet another category. Here the technical and social divisions of labor, between the production and use of statistics, have engendered their social and technical effects. The data set is a black box whose input and output sides can be clearly distinguished, provided that the input side is perceived as meeting "quality standards." Today these are increasingly explicit and guaranteed, whereas they used to be more implicit. The user's trust in the data-production phase is a precondition for the social efficiency of the statistical argument. The reality test is provided by the consistency of the results and constructions issued from the data set.

Let us now stand back from these three forms of realism, which we have separated for analytical purposes. We can observe that, if "ultimate reality" is never accessible directly but only through different perception systems, then the three realisms come together in a single test—that of the consistency between the various perceptions. However, the three approaches contrast with a fourth, which, in its concern to reconstruct the chain of coding and measurement *conventions*, effectively challenges the reality of the objects. This attitude—which we may describe as *nominalist* or *constructivist*—generally does not result from a theoretical philosophical choice but arises in situations marked by controversy, crisis, innovation, and changes in the economic, social and

administrative contexts: this is indeed the situation in the 1990s and 2000s.⁵

From Measurement Conventions to the Languages of Reality

Even more than social statistics, business statistics is a field in which the measurement conventions for most of the objects studied are continually being debated and called into question. This trend has been accelerated and made more visible by the requirements of European harmonization. Witness the list of topics to which INSEE has devoted its annual one-day seminar on business statistics between 1995 and 1999: enterprise groups, networks, accounting standards, frontiers between goods and services, and restructuring. This list alone enumerates fields where the “reality” of the objects dealt with by statisticians are constantly eluding their grasp and changing appearance. This makes the job of the official statistician in charge of business statistics uncomfortable but stimulating. The statistical activity for each of these topics is determined and constrained, at an early stage, by a multitude of negotiations—at the micro- or macro-level—among a multitude of macro- or micro-players: the government, trade organizations, the European Commission, competition-protection agencies, labor unions, and enterprises, from the largest to the smallest. The seminar topics seem to have been chosen to cover a succession of fields where all these players exercise their negotiating skills.

The existence of enterprise groups—whose boundaries are often problematic—raises the issue of which statistical unit is relevant for business statistics, which is an inexhaustible subject of debate among European statisticians. What better illustration of the problematic nature of the “reality” sought by business statisticians than the recurrence of controversies over the *statistical unit*, the most basic and central link in the system of business statistics? A host of related questions ensue: What is an activity? A product?

A "relevant market"? A "dominant position"? As we know, the answers to these questions are interlinked, in keeping with the famous "chicken and egg" principle. Scientific and legal issues are enmeshed, and some of these have major implications. For example, the December 25, 1999, issue of *Le Monde* carried four pages of dry but fascinating "legal notices" describing a "Ruling by the Competition Council of July 20, 1999, on practices observed in the sector of thermal applications of energy," consecutive to a dispute between the national power utility (Électricité de France) and the national library (Bibliothèque Nationale de France). The concept of "relevant market," which is central to this type of dispute, thus made its way into Santa Claus's stocking.⁶

In dealing with these issues, business statisticians are faced with a dilemma that puts their professional identity at stake. On the one hand they must assert a public-interest objective guided by science alone, above individual interests and their contingent controversies; on the other hand, they cannot ignore that their output will serve as arguments in such controversies. "Reality" will, in that case, be no more than a rhetorical instrument brandished by business lawyers. Some statisticians go to great lengths to claim they can single-handedly define a theory of economic reality and then observe that so-defined "reality" in compliance with that theory. The chief exponents of this position are the Dutch business statisticians. Their "statistical units" are, in principle, entirely distinct from "legal units." They are the object of painstaking, quasi-monographic definitions and observations. The French, in contrast, try not to distance themselves too much from corporate bookkeeping methods and practices, although it is harder to synthesize this material in a consistent, comprehensive, and theoretically elegant manner.

The theme of accounting standards and their harmonization at the European and global levels that was mentioned earlier lies at the heart of the questions concerning the social, negotiated, and conventional character of business statistics. The accountant's work carries within it a language of reality that defies the statisti-

cian's language; the accountant constructs, from the ground up, a universe to serve as a reference for action. This undertaking is more or less constricted by rules and standards that vary from one country to another, or rather between a few broad geographic areas, since the leading Anglo-Saxon corporations and their auditing and consulting firms are imposing de facto world standards, which the European Commission's Fourth and Seventh Directives briefly tried to anticipate.⁷

Enterprise networks are new production and distribution structures that also pose a problem for business statistics and make it difficult to use a confident realistic language. Because of their lack of a central entity and their high mobility, their boundaries are hard to draw—in contrast to enterprise groups, which generally possess a decision-making center whose sphere of influence one can attempt to define. The frontiers between the goods-producing and service-producing sectors of the economy are equally blurred by the practice of subcontracting and outsourcing entire segments of corporate production activities. Last, the permanent restructuring of large corporations and even small and medium-size enterprises makes it particularly difficult to monitor the producing sectors over time and therefore to build time series—the daily fare of econometric research.

These major issues in business statistics are well known and have been extensively studied. How do they affect the "language of reality" expected of statisticians and economists? This question can be addressed in two ways. Advocates of a defensive realistic approach will try to fill the gaps or glue the pieces back together. Ingenious solutions will be proposed to, for example, track the changes in industrial corporations over a two-year period. In a more constructivist approach, however, one can also be attentive to the way in which the language of reality itself evolves in times of crisis and rapid change. Viewed from this angle, statistical work not only *reflects* reality but, in a certain sense, *establishes* it by providing the players with a language to put reality on stage and act upon it. That is why, beyond the issues of pure statistical measurement that

they raise, the topics of the INSEE seminars listed previously are a rich terrain for observing the current changes *in vivo*.

A significant example of the evolution of languages—naturally linked to that of social relationships—is the rapid emergence of the expression “value creation” (or “economic value-added”) used in the financial world to describe the change in share prices and the gains that shareholders expect from restructuring. This reflects the rise of large pension funds and the use of the stock market to finance investment. In the process, however, the classic language of “value added”—employed in national accounting as well as in a tax context—appears to have been forgotten. One example is the economic press, which makes extensive use of the new vocabulary. In short, a competition appears to have arisen between several languages of reality, each used by different social players.

Amid the uncertainties over the status of the measurements proposed in statistical products, various expressions are often used. Each, in its distinctive way, shows that the metrology of the social sciences is not of the same kind as the metrology of the natural sciences. The terms “index” and “indicator” suggest that the reported measures are like the visible symptoms of a hidden reality that is impossible to reach directly. The expression “latent variable”—used by some—has identical implications. Occasionally, statisticians or academics prudently take refuge in the notion that their quantitative data aim to “encircle reality.” Thus surrounded, as in the game of Go, reality will have no choice but to surrender. More seriously, this apparently military metaphor evokes the notion—which is quite familiar to philosophers—that a reality is known only through external, socially constructed, and historically rooted “points of view.” By multiplying the points of view from different positions, we can always dream of “encircling reality.” But reality will slip away, for new systems—and languages to make them real—are born every day.

Notes

¹The issue of the realism of statistical production is discussed in greater detail in Desrosières (1998).

²On the origins and meanings of the distinction between "objective" and "subjective" probabilities, see Daston (1994).

³In the language of statisticians, "metadata" are the information on the definitions, construction methods, classifications, recording procedures, and processing of disseminated data.

⁴This key concept first appeared in the nineteenth century in rifle manufacturing.

⁵Ian Hacking, in *The Social Construction of What?* (1999), offers a refined analysis of the social uses of constructivist and realist arguments without locking himself into either position.

⁶Santa, in turn, seems to threaten the dominant position of the Magi in Spain (*Le Monde*, December 26, 1999). Is the Christmas cake a relevant market?

⁷European accounting standardization, which regulates the preparation, auditing, and publication of accounts, is governed by the Fourth Directive (1978) for annual accounts and the Seventh Directive (1983) for consolidated accounts of limited-liability corporations.

References

- Daston, L. "How Probabilities Came to be Objective and Subjective." *Historia Mathematica* 21 (1994): 330-44.
- Desrosières, A. *The Politics of Large Numbers: A History of Statistical Reasoning*. Cambridge: Harvard University Press, 1998.
- Hacking, I. *The Taming of Chance*. Cambridge: Cambridge University Press, 1990.
- . *The Social Construction of What?* Cambridge: Harvard University Press, 1999.
- Hopwood, A. G., and P. Miller, eds. *Accounting as Social and Institutional Practice*. Cambridge: Cambridge University Press, 1994.
- Morgenstern, O. *On the Accuracy of Economic Observations*. Princeton: Princeton University Press, 1963.
- Porter, T. M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press, 1995.
- Power, M. "From Common Sense to Expertise: Reflections on the Prehistory of Audit Sampling." *Accounting, Organizations and Society* 17:1 (1992): 37-62.

-
- Stigler, S. *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge: Harvard University Press, 1986.
- Thévenot, L. "Rules and Implements: Investment in Forms." *Social Science Information* 23:1 (1984): 1-45.

Copyright of Social Research is the property of New School for Social Research and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.