Received: 10 February 2019 | Accepted: 11 February 2019
DOI: 10.1002/asmb.2442

SPECIAL ISSUE PAPER

WILEY

Is reliability a new science? A paper from the panel session held at the 10th International Conference on Mathematical Methods in Reliability

Nozer D. Singpurwalla[1] | Vitali Volovoi[2] | Mark Brown[3] | Erol A. Peköz[4] | Sheldon M. Ross[5] | William Q. Meeker[6]

[1]Department of Systems Engineering and Engineering Management, City University of Hong Kong, Hong Kong
[2]Alpharetta, Georgia, USA
[3]Columbia University, New York, New York
[4]Boston University, Boston, Massachusetts
[5]University of Southern California, Los Angeles, California
[6]Iowa State University, Ames, Iowa

Funding information
City University of Hong Kong, Grant/Award Number: 9380068 and 9042083

Abstract

The 10th International Conference on Mathematical Methods in Reliability, MMR 2017, held in Grenoble, France during July 3-7, entailed a panel discussion entitled "Is reliability a new science?" with Mark Brown, Regina Liu, William Meeker, Sheldon Ross, and Nozer Singpurwalla as panelists. Bill Meeker also doubled as a chair and as a moderator of the panel. The panel discussion was spawned by the recent appearance of a book by Professor Paolo Rocchi, Docent Emeritus of IBM, titled "Reliability is a new science: Gnedenko was right" published by Springer in 2017. The panel discussion was well attended and enthusiastically received and could serve as a forerunner to other such panel discussions at future MMR conferences. This paper presents some elements from the lively debate generated by this discussion.

KEYWORDS
chance, hypotheses, probability, propensity, quantifiability, reproducibility, significance test

1 | CAN RELIABILITY BE CONSIDERED A NEW SCIENCE?

Nozer D. SINGPURWALLA, City University of Hong Kong, Hong Kong

1.1 | Introduction: the philosophy of science

The term "science," as a seeker of truth, carries a century old aura of legitimacy and respectability. However, every endeavor of research cannot be called a science. Fields such as biology, chemistry, physics, and their spinoffs constitute the "hard sciences," whereas fields such as economics, psychology, and sociology are known as the "soft sciences." This distinction has to do with two broad philosophical issues, articulated by the Austrian-British philosopher Sir Karl Popper, as *objectives* and *method*.[1] The objective of the natural sciences is to devise and refine approximate descriptions or models of physical universe by

1. asking a question;
2. formulating a hypothesis;
3. testing the hypothesis, and then either rejecting it or provisionally accepting it until new evidence forces its modification or its rejection.

Per the Popperian view, science grows by framing hypotheses and subjecting them to increasing severity. Progress is achieved by the fact that each successive hypothesis has to pass the same test as its predecessor, and at least one of those that its predessor has failed. This view is in contrast with the older view wherein science was about framing laws derived by induction from a multitude of particular and observational facts. To Popper, generalizations come first and the observations used to test the generalizations come next. From Popper's viewpoint, this then is the *philosophy* of *science*.

1.2 | The method of natural sciences

The method of natural sciences has the following characteristics:

1. Clearly defined *terminology*. We all know what constitutes a cell, but what does carelessness or sexism means?
2. *Quantifiability* (or *Measurability*). One can measure density, velocity, or toxicity. Thus, how can one measure happiness, inspiration, and generosity?
3. *Controllability*. A scientifically rigorous study maintains direct control over as many of the factors that influence the outcome. This is to ensure reproducibility, be it in Alaska or Zanzibar.
4. *Reproducibility*. A rigorous science is able to reproduce the same result over and over again. The failure to reproduce enables one to disprove a finding, ie, proof by contradiction.
5. *Predictability*. A rigorous science is able to make testable predictions. For example, Murray Gelman's prediction of the "quark" particle before its existence was confirmed.

1.3 | Is mathematics a science?

Were science be viewed broadly as "systematic and formulated knowledge," then mathematics would qualify as a science, though not as a hard (or natural) science.

All the same, since mathematics provides a language used by natural scientists to describe and discuss the universe, there is a *link* between mathematics and the natural sciences. However, the link is not enough because the ultimate arbiter of correctness in the natural sciences is empirical evidence, whereas the ultimate arbiter in mathematics is proof and logic. Indeed, any hypothesis which cannot be falsified by empirical data is not scientific, and that any scientific theory is a hypothesis which is only provisionally acceptable at any given time, that is waiting for new empirical evidence to falsify it. By contrast, the conclusion of a theorem, obtained by a deductive approach, is always and forever true, whenever its conditions (or hypotheses) are satisfied.

Thus, mathematics is not in the same league as a natural science, and it ought to be classified with the arts and humanities, though some claim that mathematics cannot be classified as an art either because it is not accessible enough to be labeled as art. All this goes to suggest that probability theory as a branch of mathematics cannot be classified as a science either, but it can be justifiably claimed that, like mathematics, probability is a language of science. Quantum theory is an example, at least to Einstein, whereas Bohr and Heisenberg would be inclined to label probability as a natural science. Anscombe and Aumann[2] claim that depending on how one interprets probability, as a plausibility (or reasonableness of belief or of expectation) or as a random or chance (physical) phenomenon, it can belong to logic or to physics, respectively.

1.4 | What about statistics: is it a science?

The term "statistics" appeared for the first time in Sir John Sinclair's 21 volume treatise between 1791 and 1799, entitled *Statistical Account of Scotland*.[3] Sinclair's work was to assess the political strength of a country via its inhabitant's happiness. However, Sinclair did not like the term *political arithmetic*, used by William Petty in 1690[4] in the context of reasoning by figures upon things relating to government, and so he annexed the German word statistics, which pertained to a kind of political inquiry. Thus, as it now stands, *statistics is an artificial word*, which stands for anything dealing with data.

Inferential statistics has as its genesis in the work of Sir Ronald Fisher (who was tutored by Keynes and Russell), and whose signal contributions were significance testing and the notion of a null hypothesis, randomization and the design of experiments, and the theory of estimation for implementing the principle of induction as a philosophy of science. Fisher's work was preceded by that of Bayes and Laplace whose main goal was to advocate probabilistic induction as a philosophical principle for assessing cause and effect. Fisher's work was followed by that of Neyman and Pearson who improvised on Fisher's significance test by introducing the notion of an alternative hypothesis, errors of commission and omission, and confidence limits. This in turn was followed by Wald's work on sequential analysis, losses, and utilities. Thus, Fisher remains firmly as the true creator of inferential statistics, and whose motivation was to facilitate the methodology of science to achieve the objective of scientific inquiry. All the above, and despite the fact that Fisher's significance tests played a key role in asserting the presence of the Higgs Boson, statistics is an enabler and a technology of science, like engineering. Per Popper's criteria, it does not qualify as a science, per se.

1.5 | Can reliability be labeled a science?

For reliability to be labeled a science, it should necessarily satisfy the requirement of an unambiguous and clearly accepted terminology. This is not the case because

1. to many, practically every reader of this article, reliability is a probability, and probability is not uniquely defined nor can it be directly measured.
2. to the likes of Popper, de Finetti, and Lindley, reliability is a *propensity* or a *chance* (see the work of Singpurwalla[5]).
3. to the mathematician Koopman,[6] reliability would be an *intuitive* concept prior to objective experience.
4. to Kolmogorov, Bartlett, and de Groot, reliability would be an *undefined primitive*.

This means that reliability fails the terminology and quantifiability aspect of the scientific method.

However, reliability also fails the other attributes of the hard sciences because the subject of reliability entails a craftful combination of engineering (not a science), mathematics (the language of science), and statistics (the technology of science). However, reliability, useful and valuable as it is, leans on the application of science, the philosophy of science, the language of science, and the technology of science to facilitate decision making in the face of uncertainty. In conclusion, despite the great and wonderful Academician Boris Gnedenko's wishful thinking, reliability is an essential technology; it is not a basic nor a natural science!

2 | IS RELIABILITY A NEW SCIENCE?

Vitali Volovoi, Alpharetta, Georgia, USA

2.1 | Introduction

Despite the recent populist backlash against science, the material and moral benefits associated with science—and the prestige—are indisputable. The mere use of the word "science" does not guarantee the legitimacy of that use (eg, Christian Science and Scientology*). Following the panel discussion at the MMR 2017 conference, the merits of reliability's claim on being a new science are discussed.

2.2 | Scientific methods

Since the time of Aristotle, a distinction has been made between two types of knowledge: intuitive and informal vs more deliberate, "demonstrative" (ie, scientific) knowledge. In Russian and other Slavic languages, the word for science, *nauka,* literally means "learned." In some languages there is a related, albeit more general, distinction between things perceived via intimate awareness and things learned: *eg, kennen* vs *wissen* in German and *connaître* vs *savoir* in French.[8] Blaise Pascal in *Pensées* further expanded this distinction by juxtaposing "intuitive temperament" (*esprit de finesse*) with "geometrical temperament" (*esprit géométrique*); the former case deals with ideas and perceptions that are not amenable to exact definition, whereas in the latter case, the mind works with exact definitions and abstractions of science and mathematics.

Originally, the "demonstrative" nature of scientific knowledge ought to have been derived using formal deduction as in Euclidean geometry: from the first principles (basic, foundational propositions that are deemed self-evident and absolute). As a result of the scientific revolution (and with notable contributions from Francis Bacon), another, empirical path to scientific knowledge has been developed. This path relies on observation and experiments, and entails a gradual evolution of partial truths that are repeatedly refined and clarified in the light of new evidence. Deduction is also utilized, but the first principles are not carved in stone and are subject to refinements as well.

*Scientology is, effectively, science squared, combining for extra gravitas both Latin and Greek origins (the sin that originally hampered the legitimacy of the word "scientist"). The word "scientist" eventually won out over a simultaneously suggested alternative, "nature poker."[7]

This second path grew in stature over the years, to the point that mathematics, which used to be a paragon of scientific knowledge, was either excluded from the branches of science, or admitted only with the qualifier "formal" (as opposed to the "natural" sciences that have risen to the top in their importance).

The ability to predict has become the key distinguishing feature of true science. This has led to the concept of falsifiability[1]; if a theory predicts Event A, and Event A fails to occur, the theory is deemed falsified (and in need of modification). However, the occurrence of Event A does not verify the theory but merely provisionally spares it from falsification until the next test. A theory that can explain any event (eg, Freudian psychoanalysis[†]) is not falsifiable, has zero predictive power, and is therefore not scientific. From this viewpoint, mathematics itself (as opposed to its application within a different "true" science, say physics) does not deal with predictions of events in the real world. Deductions within mathematics's own domain are either true or false, and thus do not fit the scientific method's paradigm of perpetual improvement. However, mathematics is not a real science.

Coming to the defense of poor mathematics is not that hard, however, the world of mathematics is so vast and complex that not all conjectures can be immediately verified (and hence can be falsifiable in the interim). For example, the four-color theorem[‡] was formulated in 1852 by Francis Guthrie, whereas computer-assisted proof was obtained only in 1977 (by ruling out 1,936 patterns of maps) by Kenneth Appel and Wolfgang Haken (see the work of Wilson[10]). The last Pierre Fermat theorem was conjectured in 1637, but not proved by Andrew Wiles until 1994. Thus, in full compliance with the scientific method, the four-color and last Fermat conjectures were considered to be provisionally true but falsifiable for 125 and 357 years, respectively.

For a reader with a particularly long-term view who considers a century as a mere blink of an eye, there is Kurt Gödel's incompleteness theorem and its conclusions about the inherent limitations of formal axiomatic systems. This theorem proves the existence of true but unprovable statements. If Gödel's theorem feels too esoteric for reliability, consider the lack of analytical solutions for non-Markovian system reliability models or maximum likelihood estimates given the non-convex nature of the underlying optimization problem. In all these scenarios, perfect solutions are unattainable, and the gradual evolution of approximation techniques is par for the course.

At the same time, the ancient viewpoint of the superiority of deduction is not completely dead either. In fact, the book that prompted the panel discussion at MMR 2017 calls for the development of a deductive foundation to reliability, thus upgrading it to a "real" science.[11] On balance, it seems fair to conclude that both deduction and empirical iteration are equally valid scientific methods, often used in combination. In any case, since reliability ultimately concerns itself with behavior in the real world, and as such can be viewed as applied mathematics, objections related to any lack of experimental connection are moot. However, this suggests another question, discussed next.

2.3 | Science vs technology

As with the distinction between intuitive and demonstrative knowledge, the distinction between theoretic knowledge (*epistêmê*) and pragmatic "craft" (*technê*) can be traced to Aristotle. The two distinctions can be considered different facets of the same belief in the innate superiority of "pure" knowledge unencumbered by any connection to reality. This bias is deeply entrenched in classical civilization and is credited by Drucker[12] for the decline of the British Empire; it can also be detected via the further stratification within the sciences, with physics crowned as the most "fundamental." As Ernest Rutherford put it (in accordance to Lev Artsimovich, an academician of Soviet Academy of Science), scientists are "divided into two categories—physicists and stamp collectors."[13]

On the surface, this might disqualify reliability from consideration as a science, since it mostly pertains to man-made systems. One can argue that, as artificial systems grow in complexity, they resemble natural systems more, especially at the aggregate, network level.[14] Engineering designs can be considered as special "memes," which are the informational version of genes.[15,16] As a result, the similarities between the failures of engineering and of biological systems are certainly worth exploring. Analogies in analyzing engineering maintenance vs medical records (eg, the effectiveness of cancer-screening procedures) are obvious, but it is also important to understand some of the differences. For example, in contrast to engineering systems, biological systems have local continuous processes related to rejuvenation and recycling, eg, autophagy. Rocchi[11] points out a study that explores the difference in the trends for human mortality rates and hazard

[†]Karl Popper personally knew Alfred Adler, a pupil of Sigmund Freud, and shared both Jewish roots and German-language culture with him as well as with Karl Marx, Freud, and Albert Einstein. By highlighting the inverse relationship between the explanatory and predictive powers of a theory, Popper contrasted the scientific value of Einstein's work with the lack of such value in the "theories" promoted by the other three.[9]

[‡]The theorem states that any map on a plane can be colored with four colors, so that any two countries sharing a border may always be colored differently.

rates for engineering systems.[17] The analysis is based on representing a human being as a redundant system with initial defects. This is an interesting idea, but representing a human body as a nonrepairable system flattened to a single-level hierarchy appears to be too simplistic.

Given the universal nature of the hierarchical composition of complex systems (whether natural or man-made)[18] developing dynamic multistate system reliability models that more realistically reflect at least top-level hierarchy of a human body can be a worthwhile undertaking. It would also be interesting to see whether some of the coarse-grained observations from complexity science can be reproduced using such models. For example, one curious fact is that all mammals have roughly 1.5 billion heartbeats during a lifetime, with humans fully following that rule right up until the 19th century, and then pulling ahead to 2.5 billion with the advent of modern medicine. (Remarkably, the improvement of the average life span left the upper bound virtually unchanged).[14]

Regardless of the success of applying some of the reliability theory to natural systems, the distinction between *epistêmê* and *technê* has never been absolute, given the role of technologies in scientific discoveries (such as improvements in lens technologies that led to telescopes and the heliocentric view in astronomy). Since the late 19th century, a conscious effort has been made to reverse the direction of influence and use "pure" science to drive technologies.[8] As a result, fields like materials science (that evolved from metallurgy) and computer science have emerged. The common theme appears to be the bestowal of the title of science upon a field with high status (as perceived by society at the moment). The latest example supporting this empirical rule is data science, a field that is close to reliability (as revisited in the conclusions to this paper). In the context of data science, the distinction between the "passive" and "active" approaches to the natural world is reduced to a single step in classifying analytics (from "predictive" to "prescriptive").

2.4 | Falsifiability of probabilistic statements

Predictions in reliability are probabilistic in nature, which makes the question of falsifiability tricky for two reasons. First, there is no consensus regarding the interpretation of the meaning of "probability." Second, there is the difficulty of interpreting the truth of a probabilistic prediction, especially as it pertains to a single event (the so-called decidability problem). Fortunately, for reliability, quantum mechanics also relies on probabilistic statements, which prompted Karl Popper to treat both issues at length.[1]

Similar to Andrey Kolmogorov, Popper interprets the probability by introducing an axiomatic definition of probability that can be manipulated independently of the assigned meaning. The decidability issue is resolved in a statistical sense based on a reliance on methodological rules for empirical evidence (eg, adhering to significance levels that are considered sufficient in the imperfect real world, not unlike the precision of deterministic measurements). An additional restriction is made regarding the salience of reproducibility: the methodological rules should forbid the predictable and reproducible occurrence of systematic deviations.

This last point is nicely illustrated by the recent developments in economics (a.k.a. "the dismal science"); while the presence of deviations from the "rational choice" model (which assumes that individuals always make rational decisions) has been long recognized, the lack of actionable alternatives made the assumption of rational choice the only game in town. This changed when Daniel Kahneman and his co-workers[19] revisited Pascal's dichotomy by separating "slow" (deliberate), rational thinking from intuitive "fast" thinking. Importantly, they demonstrated that intuitive thinking is prone to systematic biases.[§]

Popper's use of a "reference class" for a single event provides a prescient view on treating both frequentist and Bayesian views of probability in a unified utilitarian fashion. For example, a reference class can correspond to a particular model (or to a set of models used in a particular prediction). While each hurricane is unique, one can scientifically compare the performances of European and American hurricane prediction models for multiple hurricanes, regardless of the specific nature of the models. Similarly, the popularity of the website FiveThirtyEight is based on the aggregate performance of the models that Nate Silver and his colleagues employ.

Moreover, the probabilistic model itself does not need to be fully scientific to be scientifically tested. Tetlock and Gardner[21] have run a large-scale series of experiments where they compared "reference classes" corresponding to specific individuals who made multiple predictions and tried to understand the traits of so-called "superforecasters" who consistently outperform their peers. In this case, a model is the combined ability of a person (or a group of people) to make predictions. Digging a bit deeper, one can compare the performance of two models within the same person—this being

[§]See also the work of Ariely,[20] as well as the work of Richard Thaler, who recently was awarded a Nobel prize in Economics for his work on the subject.

one of the main messages of the Kahneman's book[19]: when making singular decisions in life, it pays to understand the aggregate performance of slow (deliberate) vs fast (intuitive) thinking, even for unique events where the impact is immeasurable, since the counterfactual conditional (ie, reproducing the same decision point using a different mental model) is impossible.

Furthermore, the benefits of overruling one's intuition for singular events are more dramatic *because* the impact is hard to measure: presumably, intuition evolved based on analogous experiences in the past, so the rarer the event is, the less relevant are the prior experiences coded in fast thinking. (However, as Richard Thaler quipped, the theory proclaiming that people learn from their mistakes, and thus approach rational behavior is good for "buying milk."[22])

In summary, what is good enough for quantum mechanics is good enough for reliability,[¶] but with one important caveat: the time scale of the statistically significant testability of a model matters. The slower the feedback is, the harder it is to gauge the predictive power of a theory.

2.5 | Conclusions: the social side of reliability

The importance of reliability's role in society is likely to be the determining factor as to whether reliability can be considered a science, as otherwise it fits the bill. However, it faces headwinds, some of which affect all sciences, as well as others that are specific to reliability. The former put the entire scientific ecosystem under stress and can be traced to two mutually reinforcing sources:

- The "democratization" of information exchange due to the Internet revolution that has, along with its multiple benefits, severely diluted the quality of content and reduced the incentives for future creation of quality content. The end result is a less coherent community and hence more transactional (ie, more short-term) interactions.[#]
- A more general increased belief in the invisible hand leading to the reduced role of government, especially in the United States and the United Kingdom. Using Thaler's terms, science is not the same thing as buying milk: its benefits are too long term to be appreciated by the markets. The negative impact of short-termism on research has been recently documented, especially in publicly traded companies that are motivated by quarterly reports.[24]

One can argue that both trends are related to the type of market failure called "the tragedy of commons," but regardless of the label, the phenomena are real and significant. Importantly, both trends are particularly pertinent to reliability and, especially, safety (safety events are rare, so that the time scale of feedback is even greater than for other scientific endeavors). In the United States, the reliability community appears to be shrinking in size (for example, at RAMS, the main reliability conference, the number of participants has roughly halved in the past twenty years). A similar trend can be observed with Google searches for the word "reliability": the frequency has halved in the past 13 years. As Professor Mark Brown, a member of the panel discussion at MMR 2017, recalled, one of the previous panels was entitled "Is reliability dead?"—and the issue has not exactly gone away.

At the same time, data science and the technologies related to the Internet of Things (IoT) create both threat and the opportunity for reliability.[25] Data science deliberately emphasizes correlation over causation, therefore minimizing the importance of deduction. This might seem to make data science less scientific and more like superefficient stamp collecting. However, at the moment, the upside of the scalability of domain-blind (black-box) data-driven solutions outweighs the lack of causal understanding. When managing failures within the IoT, knowledge about causal relationships of engineered systems and the scarcity of failure data make purely data-driven solutions objectively inferior. This provides an opening for reliability that traditionally combines causal and data-driven models, but taking advantage of this opening requires a concerted effort on behalf of the reliability community.

Scientific progress is viable only in the presence of a thriving community that enables the free exchange of information. This has been true at least since the founding of the Royal Society of London for Improving Natural Knowledge in 1660 (of which Sir William Petty, the father of social statistics, was one of the 35 founding members). Reproducibility implies transferability (portability) of knowledge, so communication is critical. Members of the scientific community understood that they were in the same boat, in stark contrast to their contemporaries engaged in alchemy who hoped to literally strike

[¶] Actually, reliability can claim some advantage: it is relatively clear why the described phenomena must be treated probabilistically. In fact, if anything, there are too many sources of nondeterminism (including variability in manufacturing processes and usage profiles). In contrast, there is no consensus as to whence probabilistic effects appear in quantum mechanics.[23]

[#] An analogy can be made with game theory: the "prisoner's dilemma" results in inferior outcomes for all participants, whereas cooperative strategies appear only in the iterated versions of the game, ie, when long-term effects are taken into account.

gold in solitude. In the area where reliability and the IoT intersect, the combination of the fragmentation of the reliability community combined with the IoT gold rush seems to lead to behavior more reminiscent of alchemy than science. This can be risky in the long run. The role of community is particularly important for reliability and safety, as it can effectively broaden "reference classes," thus improving the pace of feedback (and hence the pace of progress).

In summary, adopting the red list for the classification of endangered species, the category "vulnerable" seems to be an appropriate status for the field of reliability today: there is still time to react, but we are not yet out of the woods.

3 | REFLECTIONS ON NEW DIRECTIONS FOR AN ANCIENT SCIENCE: RELIABILITY AND BLOCKCHAIN

Mark BROWN, Columbia University, USA
Erol A. PEKÖZ, Boston University, USA
Sheldon M. ROSS, University of Southern California, USA

Reliability is not a new science; it is as ancient as life itself. The field of reliability, however, must evolve to face the current challenges of the day. For the field to flourish, we should not only write for one another using our specialized vocabulary. Rather than looking inside, we should look outside trying to help develop methodology essential to other areas and disciplines. This entails recognizing application areas where we have something useful to contribute. One of the most promising new areas today where reliability is a linchpin is the blockchain technology that is the foundation for many new rapidly growing businesses. This technology holds the promise for new ways of facilitating cooperation and trust without requiring a centralized authority. Many new businesses that have the potential of disrupting established industries rely on the reliability of the blockchain. Though the blockchain technology has many uses, its application to digital currencies is not without controversy since some people believe it has the potential to disrupt government agencies and the banking sector. We next outline a few of the technical problems in this setting where expertise from the reliability community could create great synergy and value for both the blockchain and reliability communities.

Bitcoin[26] is a digital currency that relies on the blockchain technology and maintains the integrity of a ledger of transactions using the so-called proof of work consensus protocol. At the heart of the security of this protocol is a pseudorandom number generator developed by the National Security Agency that takes any alphanumeric string as its "seed" and outputs a number between zero and one that appears uniformly distributed. The random number generator is considered noninvertible, meaning that it is impossible to know which seed will give a particular output without simply trying it. Many agents, called "miners," compete to earn a reward by being the first to discover a number that, when appended to the current blockchain and the current batch of pending transactions and used as a "seed" in the random number generator, will yield a random number that is below some very low predetermined threshold. (Actually, only the output value from the previous block is appended, in lieu of the entire blockchain, and essentially uniquely specifies the entire previous blockchain's contents. That is why it is called a "chain.") When this discovery is made, the current batch of pending transactions, called a "block," is considered sealed and the miners move on to trying to seal the next block. As such, the number of attempts until a given miner finds a correct number follows a geometric distribution, and thus the time required is roughly an exponential distribution with rate proportional to the computational power of the miner. Miners who are the first to seal a block receive a reward after a "cooling off period" provided their sealed block is part of the longest chain known so far, measured in terms of the number of sealed blocks. This incentivizes miners to only work on the longest known chain so far, and to publicize discovery of sealed blocks. It also incentivizes other miners to immediately accept a sealed block and continue working on the next block. Once a miner publicizes that the current block has been sealed, it is very easy for everyone else to immediately check this by using the pseudorandom number generator. Anyone is allowed to be a miner; there are no restrictions on who may do this.

Since the number required to seal a block depends on the set of all previous transactions, if a dishonest miner ever alters a previous transaction, it will be obvious to everyone that the subsequent numbers are not the correct ones necessary to seal the subsequent blocks. A standard model is that each miner discovers the correct numbers to seal the block according to an independent Poisson process with rate proportional to the computational power of the miner, and elementary properties of the exponential distribution imply that the chance a given miner earns, the next reward is also proportional to her rate. Because storage of the chain of transactions is decentralized and held separately with each miner, the agreed-upon rule is that the longest chain of transactions is the authoritative one. A problem occurs when a dishonest miner tries to alter a previously sealed transaction (say, N blocks in the past) and then reseal all subsequent blocks until they have the longest version of the chain, which would then be considered as authoritative. This is called the "double spend" problem because the dishonest miner could spend the same money twice. Since the honest miners continue to extend the original version of the chain, the dishonest miner will only be able to catch up if the number of events in her Poisson process ever gets at least N events larger than the number of the events in the superposition of the honest miners' process. Calculating this probability under different assumptions is crucial to evaluating security of different mining protocols and lies squarely in the wheelhouse of the reliability theorist.

As a more concrete example of the meaning of "double spend," suppose you have only 23 bitcoin in account 123 and you buy a car for 2 bitcoin from a dealer who has account 456. The car dealer of course waits for the transaction to appear as sealed into the blockchain (usually people wait until several blocks have passed) before he lets you drive away with the car. When you drive home, you can go to the previous block and insert the transaction "account 123 pays 23 bitcoin to account 789" where 789 is a second account that you control. Suppose you have fast computational power and you can seal several blocks quickly to catch up and get ahead of the miners on the main chain, but you do not include the transaction to the dealer. Once you then publicize your chain to everyone, since it is the longest they will accept it and continue building on your chain. The transaction to the dealer will still be in the pending pool but everyone will notice that it is an invalid transaction because there is no 23 bitcoin in account 123 anymore. Subsequently, no one will ever include it in the blockchain. This means you have the car and you get to keep the 23 bitcoin, and you can spend it again. That is why it is called "double spend."

Many solutions have been proposed to the "double spend" problem described in the previous paragraph, and the necessity of reliability calculations in evaluating potential solutions is essential. To reduce the chance a dishonest miner can catch up, some researchers have suggested changing the protocol so that the times that a miner seals the next blocks follows a more general renewal process rather than the Poisson process.[27] If the interevent time has a smaller coefficient of variation than the exponential distribution, the chance of catching up can decrease and thus the reliability will increase. In this case, calculation of the reliability of the protocol involves determining the chance that one renewal process with a larger mean interrenewal time ever exceeds a second renewal process with a smaller mean by more than N events. There is, however, economic tradeoffs. Miners receive a reward for each block they seal and they have the incentive to pool their resources if being a small miner does not pay. This could eventually destabilize the reliability of the block chain since larger miners will have more than a proportional increase in rewards. Computing these probabilities lies within the field of the reliability theorist and is essential for understanding the tradeoff between reliability and economic stability.

Another phenomenon where the reliability theorist can play a role involves the analysis of the so-called selfish mining attack,[28] which is a possible attack on the reliability of the block chain. Here, an agent keeps sealed blocks secret from the rest of the community so that other people will waste their resources mining a shorter chain, which will not end up being part of the authoritative block chain. The selfish miner keeps her mined blocks secret until the group of honest miners is almost about to catch up, and then she releases the blocks so that everyone else's work has been wasted. This waste of resources slows down the mining process and essentially lowers the computation power of the honest miners, and therefore allows the dishonest miner to have an unfair advantage. The theory goes that honest miners will realize this and eventually switch over to being selfish miners, and this destabilizes integrity of the block chain. Computing the chance that various selfish mining strategies succeed, under different assumptions on the distribution of time it takes to mine a block, is of paramount importance and involves calculations within the purview of the reliability theorist.

It may seem at first glance that the selfish mining attack may harm the attacker. If they are secretly mining a block, someone else may seal the block and the selfish miner may not get credit for it. An important part of the selfish mining strategy is that, if you notice someone else seals the block, you quickly release your block so it appears that both happened at exactly the same time. This then means about half the miners may try to mine on your block and the other half on the other block and you have roughly a 50% chance of having your work count toward a reward. The big value in selfish mining comes if you get more than one block ahead of the honest miners: then you can always just release your chain to the public as soon as the rest are just about to catch up and are one block behind. It can be shown that this strategy hurts the honest miners slightly more than it hurts the selfish miner and can create just barely enough wasted work by honest miners overall that they will want to do selfish mining. It will not profit a selfish miner in the short term, but long term means that there is an incentive for everyone to collaborate with the selfish miner and split the proceeds, which would give the selfish miners a majority of computational power and they thus would be able double spend. People believe that this has not really happened because it would hurt the value of bitcoin, and people who are big players have the incentive not to hurt the value.

Much of the reliability theory of the coming century will naturally involve adversarial systems. This is a natural consequence of the connectivity and decentralization of our society. During the last century when we were not as connected, much of the uncertainty and hazards in systems were due to Mother Nature, aging, and defects. With advances in technology, systems are much more secure from those threats but face new classes of strategic adversarial threats. The burgeoning field of cyber security is a particularly timely example. The field of reliability is well equipped to face these challenges and to show leadership in these new areas, but it must be open-minded into studying new models and developing new theory. This new direction for the field of reliability theory will bear much fruit and be a key into helping human society evolve to better cooperate and flourish peacefully.

4 | IS RELIABILITY A NEW SCIENCE?

William Q. Meeker, Iowa State University, USA

Nozer Singpurwalla argues that reliability is not a science. However, of course, because there are differing definitions of science, one can also argue the opposite. I once worked on a project with a very smart lawyer. He told me that he is always willing to allow the opposing lawyer write the contract if he can write the definitions. If one uses a definition that admits mathematics as a science, then clearly deriving bounds for reliability functions and proving closure results of the reliability functions for different kinds of systems would be contributions to the science of reliability. Under the same definition, it is easy to argue that statistics is also a science.

Although I am defined as a statistician by education and more than 40 years of employment in the Statistics Department at Iowa State University, my views of reliability have been deeply influenced by my experiences and interactions with engineers and statisticians working with engineers in industry. Three years of summer internships at GE Corporate Research and Development (as it was known back then) while in graduate school led to 40-plus years of collaboration with individuals in that organization (initially with Wayne Nelson, but subsequently with Gerry Hahn and Necip Doganaksoy) and numerous visits and consulting engagements with GE engineers. From 1978 to 1995, I spent approximately 10 weeks every summer visiting Bell Laboratories working with engineers and statisticians on applications in telecommunications reliability. Since my book Statistical Methods for Reliability Data with Luis Escobar was published in 1998,[29] I have taught on the order of one hundred reliability short courses, mostly to engineering audiences, and have been involved in many consulting projects, working with engineers and scientists to apply statistical methods in nonstandard reliability applications.

What is indisputable is that reliability is a highly-quantitative engineering discipline that depends heavily on many areas of science including probability and statistics but also relying on materials science, physics, and chemistry. In the most challenging and important applications that I have worked on, companies made sure that they had appropriate experts from all of the needed disciplines working together on the team.

Unlike most survival analysis applications in medical statistics, the analysis of engineering reliability data almost always involves extrapolation. We extrapolate in time when we predict the fraction failing during a three-year warranty period, based on six months of field data. We extrapolate in temperature when we estimate a life distribution at 25 C, based on accelerated tests at 50, 60, and 70 C. We extrapolate from laboratory testing to make predictions about field performance. Some Statistics 101 books state emphatically “You cannot extrapolate.” However, the truth is that reliability engineers extrapolate all of the time. Although it is not always the case, the justification for the needed extrapolation should be and often is detailed knowledge of the failure mechanism and the corresponding chemistry or physics of failure. Thus, reliability, as a discipline, depends on and has benefited from fundamental scientific work that has been done in the areas of physics, chemistry, and materials science.

During my entire career, I have said that I am too young to get deeply involved in philosophical arguments and that I would rather be involved in solving statistical problems. Thus, I will end my discussion here.

ACKNOWLEDGEMENTS

Kevin Wilson (Newcastle University, UK) painstakingly orchestrated the panel discussion and selflessly expended much effort to bring it to fruition. N.D. Singpurwalla is supported by project numbers 9380068 and 9042083 of City University of Hong Kong.

REFERENCES

1. Popper K. *The Logic of Scientific Discovery*. London, UK: Hutchinson and Co; 1959.
2. Anscombe FJ, Aumann RJ. A definition of subjective probability. *Ann Math Stat*. 1963;34:199-205.

3. Sinclair J. *The Statistical Account of Scotland: Drawn up From the Communication of the Ministers of the Different Parishes*. Edinburgh, Scotland: William Creach; 1794.
4. Petty W. *Political Arithmetick*. London, UK: R. Clavel; 1690.
5. Singpurwalla ND. Filtering and tracking the propensity of survival (with discussion). *Stat Sci*. 2016;31(4):521-540.
6. Koopman BO. The axioms and algebra of intuitive probability. *Ann Math*. 1940;41:269-292.
7. Whewell W. On the connexion of the physical sciences. By Mrs. Somerville. *Q Rev*. 1834;LI(CI):54-68.
8. Barzun J. *From Dawn to Decadence: 500 Years of Western Cultural Life : 1500 to the Present*. New York, NY: HarperCollins Publishers; 2000.
9. Popper K. Philosophy of science: a personal report. In: Mace CA, ed. *British Philosophy in the Mid-Century: A Cambridge Symposium*. Crows Nest, Australia: George Allen and Unwin; 1957:182-183.
10. Wilson R. *Four Colors Suffice*. Princeton, NJ: Princeton University Press; 2013.
11. Rocchi P. *Reliability is a New Science: Gnedenko Was Right*. Cham, Switzerland: Springer International Publishing AG; 2017.
12. Drucker PF. Beyond the informational revolution. The Atlantic; 1999.
13. Lang D. The peaceable gypsies. The New Yorker. 1963.
14. West G. *SCALE: The Universal Laws of Growth, Innovation, Sustainability, and the Pace of Life in Organisms, Cities, Economies, and Companies*. New York, NY: Penguin Press; 2017.
15. Dawkins R. *The Selfish Gene*. 2nd ed. Oxford, UK: Oxford University Press; 1989.
16. Frenken K, Nuvolari A. The early development of the steam engine: an evolutionary interpretation using complexity theory. *Ind Corp Chang*. 2004;13(2):419-450.
17. Gavrilov LA, Gavrilova N. The reliability theory of aging and longevity. *J Theor Biol*. 2001;213(4):527-45.
18. Simon HA. Near decomposability and the speed of evolution. *Ind Corp Chang*. 2002;11(3):587-599.
19. Kahneman D. *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux; 2011.
20. Ariely D. *Predictably Irrational: The Hidden Forces That Shape Our Decisions*. New York, NY: HarperCollins Publishers; 2009.
21. Tetlock PE, Gardner D. *Superforecasting: The Art and Science of Prediction*. New York, NY: Crown Publishing Group; 2015.
22. Lowenstein R. Exuberance is rational. New York Times. 2001.
23. Weinberg S. The trouble with quantum mechanics. *N Y Rev Books*. 2017;64:51-83.
24. Jacobs M, Mazzucato M, eds. *Rethinking Capitalism: Economics and Policy for Sustainable and Inclusive Growth*. Chichester, UK: John Wiley and Sons; 2016.
25. Volovoi V. Big data for reliability engineering: threat and opportunity. 2016.
26. Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. 2009. https://bitcoin.org/bitcoin.pdf
27. Bissias G, Levine B. Bobtail: a proof-of-work target that minimizes blockchain mining variance. 2017. https://arxiv.org/pdf/1709.08750.pdf
28. Eyal I, Sirer EG. Majority is not enough: bitcoin mining is vulnerable. *Commun ACM*. 2018;61(7):95-102.
29. Meeker WQ, Escobar L. *Statistical Methods for Reliability Data*. New York, NY: John Wiley and Sons; 1998. *Wiley Series in Probability and Statistics*.

How to cite this article: Singpurwalla ND, Volovoi V, Brown M, Peköz EA, Ross SM, Meeker WQ. Is reliability a new science? A paper from the panel session held at the 10th International Conference on Mathematical Methods in Reliability. *Appl Stochastic Models Bus Ind*. 2019;35:260–269. https://doi.org/10.1002/asmb.2442