Without Cause
Mark Feeley
June 17, 2013

Abstract
Physicists increasingly accept that information is more fundamental than material things, but if material things are not fundamental, then neither are material causes: we will live in a world without cause. We thus examine the steps and missteps by which information came to be seen as more fundamental, examine the flaws and risks of a purely informational view, and consider a possible approach to restoring a belief in material things and material causes.

“Some mon just deal wit’ information. An’ some mon, ‘im deal wit’ de concept of truth. An’ den some mon deal wit’ magic.”
- Nernenny, Rastafarian “Bush Doctor”

Introduction
It will come as a surprise to most of the general public, and even to most beginning students of physics, that a great many theoretical physicists believe in magic and not physical law. Guided by the dogma of quantum theory, many (and perhaps most) physicists accept that in the so-called quantum world, events can happen with no natural cause at all: a particle decays into other particles, particles are detected here versus there, or a spin is resolved as up or down. In the orthodox quantum view, the outcomes in these examples are said to be defined at the instant of measurement as the result of some undefined sort of stochastic process. Unfortunately, though the term “stochastic process” has a pleasantly scientific tone, if there is no natural cause for such events, then we can safely replace this term with “supernatural cause” or “magic” without any change in meaning. Even if most physicists do not admit to a general belief in magic, they must admit that there is a general loss of faith in natural causes which has caused the search for physical law to be almost abandoned. No one now seeks to understand why a particle decays at a given time or a spin resolves as up or down; very few even believe there is a why. Expressed in John Wheeler’s terms, physicists no longer believe in “It”.

The modern idea is that we live in an informational world, not a physical one, and that the fundamental laws that we should seek are the laws of information, not physical laws. This is misguided and it is dangerous. Information theory necessarily augments physical theory when knowledge is limited, but cannot replace it. The choice facing physics is not one of information theory versus physical theory, it is information theory plus physical theory versus information theory plus magic. That physicists do not believe in a physis, a physical world governed by physical law, can only be seen as a crisis for physics. However, with some intellectual discipline and some retraining, we might escape this crisis.

1 email: markfeeley@sympatico.ca
2 Nernenny, quoted in [1], p. 1,
Into the crisis: steps and missteps
To understand the way out of the crisis, we must first understand the way in. Ernst Mach laid the groundwork for the rise of the informational view by arguing that physical science could not aspire to be a true description of reality, but should instead be the best summary of the available facts about reality:

“The goal which it (physical science) sets for itself is the simplest and most economical abstract expression of facts.”

By introducing the critical distinction between “facts about reality” and “reality itself”, Mach allows the facts and the reality to diverge, and thus admits into physical theory two elements that are now essential: that the facts may be observer dependent and that the facts may be limited. Of course Einstein used the observer dependence of facts with stunning success in his developments of both relativity theories. The second element, the limitation of information, generally necessitates the use of probabilistic or statistical treatments, and is a key feature of both statistical mechanics and quantum mechanics. Claude Shannon created modern information theory in 1948 as a development of standard probability theory, and a few years later, Edwin Jaynes pioneered the application of information theory in physics. Jaynes redeveloped much of statistical mechanics in terms of Shannon’s information theory, and gave us a new understanding of entropy as an informational or epistemic concept rather than a thermodynamic one. However, the uses of information theory in quantum and statistical mechanics can be sharply contrasted. Statistical mechanics acknowledges the limitation of knowledge without denying the existence of an underlying reality; the physical view of reality is still unquestionably more fundamental than the informational view. The physical model actually informs the statistical model. On the other hand, depending on the interpretational flavour, quantum theory is either indecisive about the nature and existence of reality or denies reality outright. It is only with quantum theory that an informational view begins to be considered as somehow more fundamental than a physical view. However, the confusion and doubt about reality in quantum theory, and thus the support for the primacy of an informational view, stems from a misunderstanding: the creators of quantum theory simply did not understand or apply probability theory correctly.

The way out: relearning probability and quantum theory
Those of us who struggled with discipline in our early school years may recall “writing lines” as punishment for our misdemeanours: 500 lines of “I will not chew gum in class...” and so on. This type of remediation would be of great benefit to physicists today. Understanding of physics would be vastly improved if in the first class of every course in quantum theory – beginning and advanced, undergraduate and graduate – students were handed a stack of foolscap paper and assigned to repeatedly write a version of the extraordinarily lucid and pointed line given to us by de Finetti:

“Probability is not real.
Probability is not real.
Probability is not real...”

---

4 Original: “Probability does not exist”, B. de Finetti, [6]
In the second class, students could discuss at length: if probability is not real, then what is it exactly? The familiar game of heads and tails, coin tossing, tells us all that we will need. Students can be asked to consider a coin tossing experiment \( C \) in which they are given a coin tossing machine of such precision that a coin initially placed in the machine with a given face up will land upon a table with the same face up with 100% certainty. The table is glass so that the coin may be read by an observer above or below the table. Before any toss, information \( I \) will be given, stipulating whether the coin is placed in the machine with heads up/down/unknown and whether the observer plays the game by reading the coin from above/below/unknown. They will be asked to understand the meaning of the term “the probability of heads or tails given this information”, denoted \( P(H|I) \) and \( P(T|I) \).

First, they must decide what heads and tails actually are. Of course, coins have symmetries and physical features, namely two faces embossed with pictures, and we can arbitrarily label each face as heads or tails. However, by “the probability of heads” we mean something like “the probability that the face arbitrarily labelled heads is visible to an observer at the end of the experiment”. In this context, heads is the name for a state, not a face. But a state of what? The first guess may be that heads and tails are final states of the coin, but brighter students will object that they cannot simply be states of the coin, since the reading also depends on the position of the observer. Furthermore, heads and tails have no meaning at all until the experiment is complete, so coins cannot be said to “have” heads or tails states. They will conclude, hopefully, that heads and tails are outcomes of the experiment. They are not properties or states of the coin; they are states of the outcome. The difference is profound. As we allow no other outcomes, they will agree that the set \{heads, tails\} is the entire outcome space of the experiment.

The students will then be asked to explain how probability is determined and what variables affect it. They will consider the probabilities that would seem reasonable to them depending on whether they are told the coin is initially oriented with heads up/down/unknown, supposing in all three cases that they are told the position of the observer (above). Knowing the precision of the machine, they will decide: if the coin is known to be initially oriented with heads up then we should reasonably assign a probability of 1 to the heads outcome, \( P(H|I) = 1 \), if known to be initially oriented with heads down then we would reasonably assign \( P(H|I) = 0 \), and if the initial orientation is unknown then we would assign \( P(H|I) = 0.5 \). From this they will easily conclude the key features of probabilities: probabilities are assigned by us, and probabilities are related to our information about the conditions of the tossing experiment and not directly to the coin, the tossing machine, or any other physical thing.

So, we can explain, to quantify the state of our knowledge using the methods of probability theory, we first assume an outcome space, and then we assign a probability to each outcome in that outcome space. In probability theory, this probability distribution over the outcome space entirely quantifies our knowledge. Probability is then used to make predictions and we can describe the methods of probability theory, statistics, or information theory as various types of epistemological calculus.

At this point, students should understand “Probability is not real”. Probability is not itself physical and thus does not exist in space or time. Probability is not a property of, or directly associated with, either physical things or physical systems, or even states of physical things or systems. There is no such thing
as a “physical probability”. Probability is an *epistemological* measure assigned *by us* to outcomes of experiments, and is used to quantify our knowledge in an epistemological calculus.

The third class could entail some more advanced lines:

> “Whenever I see probability in an expression, I will interpret the expression as epistemological, not ontological.

> Whenever I see probability in an expression, I will interpret the expression as epistemological, not ontological…”

In plain English: “If a theory or an expression has a probability in it, then it’s not about something physical, it’s about outcomes we think might occur”.

After this rote training and discussion, it might then be safe to introduce quantum theory. We will attempt to present the theory without the history, the mystery, the philosophies, or the interpretations. We will hope that students have not been tainted by too much prior exposure to the theory – in popular science books, FQXi contests, and so on. We will give them only equations and some guidance regarding the notation, and ask them to deduce what they can of the meaning of the equations and the theory. We need only tell them that the theory features expectation values, and even the least astute will recognize that they are dealing with a probabilistic theory: an epistemological theory, not an ontological one. They will expect a theory of experiments and outcomes.

Since they expect experiments, we first develop a classical theory of experiments, once again referring to a coin tossing experiment $C$. As heads and tails are outcomes rather than mathematical entities, we must first choose a mathematical representation for these states. We will make use of vector algebra, and so choose to represent outcomes as “directions” in the outcome space. Thus, we define $|H\rangle$ as a unit vector in the “heads direction” ($H$ or $\vec{H}$ would be more obvious vector notations, but the bra-ket notation will be useful later). We then want to assign two different quantities to an outcome state: a value of some kind (usually numeric) and a probability. The values we assign for head and tails are denoted $c_H$ and $c_T$. The choice is entirely arbitrary, but $c_H = 1$ and $c_T = -1$ would be a typical choice.

To represent this value assignment mathematically, we define an operator $\hat{C}$ satisfying

$$\hat{C}|H\rangle = c_H|H\rangle, \quad \hat{C}|T\rangle = c_T|T\rangle.$$  

Thus, $\hat{C}$ simply represents the process of assigning a numeric value $c_H$ to the abstract outcome state $|H\rangle$. The key benefit of this operator representation is that it clearly distinguishes between the value we have assigned to an outcome state and the outcome state itself. We call $|H\rangle$ and $|T\rangle$ the eigenvectors of $\hat{C}$ and call $c_H$ and $c_T$ the eigenvalues of $\hat{C}$. We require that our operator $\hat{C}$ generates values for all possible outcomes of the experiment, and formally this requires that the eigenvectors of $\hat{C}$ span the outcome space. We will further specify that the outcomes are defined by orthonormal vectors, so that:

$$\langle H|H\rangle = \langle T|T\rangle = 1, \quad \langle H|T\rangle = \langle T|H\rangle = 0.$$  

If $|H\rangle$ and $|T\rangle$ are defined as orthonormal vectors spanning the outcome space, then any vector $|\Psi\rangle$ in the outcome space of the experiment can be written

$$|\Psi\rangle = h|H\rangle + \epsilon|T\rangle.$$
We next assign probabilities $P(H|I)$ and $P(T|I)$ to the outcome states, and according to probability theory, the expectation value for a coin toss $\langle C \rangle$ is

$$\langle C \rangle = c_H P(H|I) + c_T P(T|I).$$

This is all we need for many purposes, but to capture the full representation of the coin toss experiment, and recognizing that the process of assigning values was arbitrary, we may wish to express $\langle C \rangle$ in terms of an outcome vector $|\psi\rangle$ and the operator $\hat{C}$ explicitly. To do this, we define quantities $\phi_H$ and $\phi_T$, called probability amplitudes, which are complex roots of the probabilities, satisfying $\phi_H^* \phi_H = P(H|I)$ and $\phi_T^* \phi_T = P(T|I)$. We will ignore the detail of why we choose complex versus real roots in this essay, and indeed the phases will rarely matter. We then further define our outcome vector to be

$$|\psi\rangle = \phi_H |H\rangle + \phi_T |T\rangle,$$

so that

$$\hat{C}|\psi\rangle = c_H \phi_H |H\rangle + c_T \phi_T |T\rangle.$$

Standard vector algebra defines the scalar product of vectors (with complex components) as

$$\langle \Psi_2 | \Psi_1 \rangle = h_2^* h_1 + t_2^* t_1,$$

and we can use this scalar product to link the operator and outcome vector representations to probability theory and give us $\langle C \rangle$ in terms of $\hat{C}$ and $\Psi$,

$$\langle C \rangle = \langle \Psi | \hat{C} | \Psi \rangle = \langle \Psi | \hat{C} | \Psi \rangle = \phi_H^* (c_H \phi_H) + \phi_T^* (c_T \phi_T) = c_H P(H|I) + c_T P(T|I).$$

We now have a complete representation of our coin toss experiment, which allows us to capture outcome states ($|H\rangle$), the probabilities ($P(H|I)$) and probability amplitudes ($\phi_H$) assigned to those states, the values ($c_H$) assigned to those states, and the procedure ($\hat{C}$) which assigns those values. This theory of experiments is simply standard probability theory combined with an operator representation of experiments and some vector algebra. We can easily generalize to more outcomes or to continuous outcomes. Although it should go without saying, coin tossing is perfectly classical.

Now, finally, we can begin quantum theory. Students can be grandly told their first “fundamental postulate of quantum mechanics”:

\textit{With any observable $A$, we associate an operator $\hat{A}$ which acts on $\Psi$, and the only results of a measurement of $A$ will be one of the eigenvalues $a_i$ of $\hat{A}$, satisfying $\hat{A} |\psi_i\rangle = a_i |\psi_i\rangle$.}

Having just seen this in the context of coin tossing, this will seem blasé. They will not view this as having the exalted status of a postulate or even in any way quantum, just the standard structure of a theory of experiments – classical, quantum, or otherwise. They will instantly recognize that an observable is a name for a type of experiment (coin tossing), not a property of some physical thing, that a measurement is an instance of that experiment producing a single outcome (a toss), that the eigenvectors $|\psi_i\rangle$ are the outcomes (heads or tails), and that the set of eigenvalues $\{\psi_i\}$ is the outcome space of the experiment ($\{|heads, tails\}$), and the eigenvalues $a_i$ are values $(1, -1)$ which we have chosen to assign to outcomes.

Given the “fundamental postulate” for expectation values

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle = \sum (\phi_i | a_i | \phi_i | \psi_i \rangle = \sum a_i \phi_i^* \phi_i = \sum a_i P(i),$$
they will recognize \( a_i \) as the value assigned to the \( i \)th outcome of experiment \( A \), and \( P(i) \) as the probability of the \( i \)th outcome. They will see a straightforward probability-based theory of experiments.

The meaning of the wavefunction \( \psi \) would be quite mundane and just as it was for coin tossing: it is a vector in the outcome space of an experiment. The wavefunction captures both the outcome states (defined as part of the definition of the experiment) and the probabilities which we have assigned to those states. Not wishing to write more lines, they will not accept that the wavefunction is in any way physical, or even directly associated with anything physical. Probability is not real.

If you were to tell them, with fanfare befitting such a great mystery, that the function \( \psi \) “collapses” upon measurement, instantaneously and everywhere, they would be astonished only at your theatrics, as it is quite obvious that the probability of an outcome becomes 1 when that outcome is known.

We can then ask them to attempt to determine, from equations alone, what sort of physical thing the theory might describe. Of course, they will accept that there is some physical thing, as they have no particular reason to suspect otherwise. We can tell them that we seem to get the best predictions in many experiments if we assume that the function \( \psi \) evolves according to wave-like equations such as

\[
i \partial_t \psi = \left( -\frac{i}{2m} \nabla^2 + V \right) \psi \quad \text{or} \quad (\partial_t^2 - \nabla^2) \psi = -m^2 \psi.
\]

But why should we choose to assign a spatially and temporally varying probability to an outcome? Since any information is given at the start of the experiment and does not change, we must have reason to believe that some feature of the physical thing in our experiment evolves in a wave-like fashion. For example, suppose that our experiment (our observable) is named “water heights on vertical sticks in the Bay of Fundy”, our outcomes are heights \( h \), our outcome space is the continuous domain \([h_{lo}, h_{hi}]\), and our information \( I \) is that the Bay of Fundy is famous for its tides, that tides have exhibited periodic behaviour with period \( \approx 12 \) to 24 hrs in all previously known cases, and we are given \( h(t_o) \). Given \( I \), we have sound reasons to believe that we should assign a time-varying probability to any outcome \( h \). If \( h \approx h_{hi} \) now, we should assign a high probability that \( h < h_{hi} \) in about 6 hrs. With more information, such as a physical theory of tides and positions of the sun and moon, we could refine our probability model further. The equations of quantum theory do not cast doubt on the existence of an underlying physical world. On the contrary, the evidence offered by the equations positively suggests a physical world with wave-like features. Students would be keen to understand this physical world.

Unfortunately, students are not taught quantum theory this way. Instruction often begins with tales of the supposed failings of “classical physics”, and proceeds to such incomprehensible pronouncements as “we may associate a wave function with every particle, and the wave function is a complex probability amplitude whose squared modulus represents the probability of finding the particle at a point in space-time”, “the wavefunction captures all that we may know about the system”, or “the wavefunction is a probability distribution over the state space of the system”. These ideas are nonsense, as they all clearly associate the probability with the thing – the particle or the system – not the outcome of an experiment given some information. The faulty notion of physical probabilities is intrinsically assumed. The idea of physical probabilities pervades beyond wavefunctions, and we might hear “the electron has a probability of \( \frac{1}{2} \) of being found in a state of spin up or down in any direction”. Just as a coin does not “have” a heads state, so an electron need not “have” a spin state. An electron has features which can be
manipulated in experiments with outcome states labelled $|\uparrow\rangle$ and $|\downarrow\rangle$, we can say no more. We are told the Uncertainty Principle relates two properties of a particle, but it cannot: it relates only the outcomes of two different experiments. These ideas are confused and wrong.

Students taught “our way”, with a clear idea of what a probability is, would reject all of these assertions out of hand. They would know that probabilities are not real. Our students would demand to know the information which was given in order to assign a probability to an outcome of a given experiment. They would not doubt the existence of a physical thing upon which they experiment, and they would demand to know of any physical models which they might use to better estimate probabilities. They would recognize that they have limited information, but would not doubt the existence of causes. In short, they would not be deceived.

"It from Bit": here be dragons
Unfortunately, most people have been deceived. The flawed concept of physical probabilities is almost inextricably tied into the foundations of quantum theory and leads directly to most of the confusion in physics. It leads to all of the confusion about epistemology and ontology, about information and reality, and it leads, inevitably, to “It from Bit”. Now, John Wheeler has made very many important contributions to physics, but “It from Bit” is simply not his finest hour. In fairness, the essential idea behind “It from Bit” is not even his, as in 480 BC the Milesian Greek philosopher Anaxagoras taught that all things were created by the mind. Wheeler’s principal innovation over Anaxagoras was to assert that the information received by the mind is digital, which indeed may not be trivial as a form of limitation of information, but the central idea remains the same. However, “It from Bit” effectively captures the zeitgeist of post-quantum physics, and as a quote, has a mystical, Zen-like quality which gives it great power. Unfortunately, it is both wrong and unhelpful.

To see why Wheeler’s “It from Bit” is wrong, we can examine his own explanation:

> Otherwise put, every "it" — every particle, every field of force, even the space-time continuum itself — derives its function, its meaning, its very existence entirely — even if in some contexts indirectly — from the apparatus-elicted answers to yes-or-no questions, binary choices, bits. "It from Bit" symbolizes the idea that every item of the physical world has at bottom — a very deep bottom, in most instances — an immaterial source and explanation; that which we call reality arises in the last analysis from the posing of yes-no questions and the registering of equipment-evoked responses; in short, that all things physical are information-theoretic in origin and that this is a participatory universe.\(^5\)

This passage is not terribly lucid, but we can attempt to parse it. Wheeler defines “Bit” as a set of apparatus-elicted answers to yes or no questions. To make any logical sense of the explanation, we must ask two questions: “Is the apparatus It or Bit?” and “Questions about what?” The first question actually has no satisfactory answer. If the apparatus is “Bit” then we unfortunately are not left with any “It”, all is “Bit”. Such a position may well be valid, but anyone seriously holding this opinion is best

---

\(^5\) J. Wheeler, in [3]
advised to abandon physics and pursue philosophy or psychology. On the other hand, if the apparatus is “It”, then following Wheeler, this apparatus owes existence to yes or no questions asked by another apparatus, and that one by yet another, in infinite regress. This is nothing more than a fancy proposal for implementing the famous “turtles all the way down”, and is really not very useful. Sadly, turtles that ask yes or no questions are no more believable as a basis for reality than regular turtles. As to the second question, unless the apparatus concocts answers of its own volition, then we must presume that the apparatus gives its yes or no answers based on what it can determine of some domain external to itself. Even from Wheeler’s own definition of “Bit”, it is thus apparent that information is information about something, not information about information or information in the abstract. “Bit” is about “It”. As should have been quite obvious at the outset (except apparently, to theoretical physicists), the information is derived from the something, not the something from the information. Despite the imaginative sophistry, “It from Bit” lacks any logical consistency and really does not pass muster.

Now, some ideas turn out to be illogical or wrong but nevertheless useful, and Newton’s theory of gravity as instantaneous force-at-a-distance is one such example. Instantaneous force-at-a-distance is now thought to be wrong, but the idea remains computationally useful in many domains and was also very useful as a stepping stone in the search for better physical law. Wheeler’s idea has no such merits. “It from Bit” suggests that we should consider information theory, not physical theory, as fundamental. Whenever we have limited information, some form of probabilistic or informational theory actually must be used. However, these methods must only be used to augment a physical theory. Whatever the many uses and merits of information theory, it is fundamentally empiricist, and does not require or seek causes, mechanisms, explanations, or physical laws. Thus, the view that physics is information theory implicitly suggests that we can abandon the necessary search for those mechanisms, causes and laws.

Jaynes, on the other hand, was always particularly careful to distinguish between reality and what we know about reality. He even gives a name to the mistaken assumption that what we know about reality is reality: he calls this the Mind Projection Fallacy. The principal danger of the Mind Projection Fallacy is the denial of causes: “I do not know the cause” therefore “there is no cause”. Unfortunately, by claiming that reality “owes its very existence to” what we know about reality, Wheeler became the poster boy for the Mind Projection Fallacy. Jaynes describes the danger of a loss of a faith in physical causes very well:

*In current quantum theory, probabilities express our own ignorance due to our failure to search for the real causes of physical phenomena – and worse, our failure even to think seriously about the problem. This ignorance may be unavoidable in practice, but in our present state of knowledge we do not know whether it is unavoidable in principle, the “central dogma” simply asserts this, and draws the conclusion that belief in causes, and searching for them, is philosophically naïve. If everybody accepted this and abided by it, no further advance in understanding of physical law would ever be made; indeed, no such advance has been made since the 1927 Solvay congress in which this mentality*
became codified into physics. But it seems to us that this attitude places a premium on stupidity; to lack the ingenuity to think of a rational physical explanation is to support the supernatural view.⁶

Of course Jaynes, of all people, is not arguing that we should not use statistical methods, only that statistical methods should augment but never replace the search for physical theory. Recent work by a number of physicists on reconstruction of quantum theory with a new a set of informational axioms is valuable and will provide clarity to be sure, but will not and cannot produce a physical theory. Indeed, informed as it is by the “It from Bit” philosophy, such work does not even strive to do so. A physical theory underlying quantum theory is also needed, and it is most definitely not naïve to pursue it.

Conclusion

“What is the relationship between epistemology and ontology, mind and matter, information and reality, or Bit and It?” However we phrase it, the question is a very old one – older than Wheeler and Jaynes, older than Mach, older than Anaxagoras, and possibly older even than the cave painters of Lascaux. Just as the ancient painters used pigments to create representations of the reality they saw, we use mathematics to create representations of the reality we see. To be sure though, there is a reality, “It”, and the information creates a representation of that reality. Our information is very likely constrained to be digital as Wheeler suggests, thus “Bit”, and may be limited in other ways, but “It” does not derive from “Bit”. “Bit” manifestly derives from “It”.

Since information about reality is necessarily limited, physicists can and must make use of information theory to understand physics, but it is wrong and dangerous to assume that information is fundamental. Information theory is a tool which allows us to quantify and best use our knowledge about the physical world in a concise mathematical form. Physicists should use this tool, but the task of physicists is nothing other than to discover physical theory, physical law and physical causes. The task cannot be avoided or shirked from, and cannot be wished away with information theory, stochastic processes, or mystical incantations. Quantum theory, the source of the problem, will have to be rethought and relearned. Since information about reality and reality itself are different things, they must be differentiated in any theory. Quantum theory provides no such distinction, and thus, despite its predictive value, must be wrong. Quantum theory must be reworked or replaced with a theory which provides the same results, but which offers a clear epistemological/ontological (or Bit/It) boundary, or it is unlikely that progress can be made.

Probability is not real, but causes are real. We must not believe in magic. We can be optimistic that a physical theory underlying quantum theory can be found – that “It” can be restored to primacy. Indeed, it is Wheeler himself who best inspires us to continue the search:

“Behind it all is surely an idea so simple, so beautiful, that when we grasp it – in a decade, a century, or a millennium – we will all say to each other, how could it have been otherwise?”

— John Archibald Wheeler⁷

---

⁶ E. Jaynes, in [2], p.1013
⁷ J. Wheeler, in [4], p.304
References


