*Journal & Proceedings of the Royal Society of New South Wales*, vol. 155, part 2, 2022, pp. 182–188. ISSN 0035-9173/22/020182-07

# *Quis custodiet ipsos custodes?* From quantum gravity to second-order cybernetics

**Dean Rickles**

University of Sydney

Email: dean.rickles@sydney.edu.au

## Abstract

This brief note will discuss an interesting byway that emerged from my historical and philosophical research on quantum gravity (for which I was awarded the 2021 History and Philosophy of Science Medal of the Royal Society of NSW). The byway in question concerns the seemingly essential role played by observers (or subjectivity) in our models of reality. That is, the deepest probing of reality in fundamental physics (which quantum gravity research amounts to), suggests a necessary role for humanity, or something like it. Science is truly a human story on this account.

> We have found a strange footprint on the shores of the unknown. We have devised profound theories, one after another, to account for its origins. At last, we have succeeded in reconstructing the creature that made the footprint. And lo! It is our own. — Sir Arthur Eddington, *Space, Time, and Gravitation*, Cambridge University Press, 1920, p. 200)

Readers with a fondness for the classics will immediately recognise the above phrase in my title from Juvenal's *Satires* (Satire VI, lines 347–348): "Who watches the watchers?" The implication being, of course, that somebody *should* watch those watchers! This is, most will agree, general good-sense, unfortunately not heeded nearly enough in our political hierarchies and, according to some scientists, not heeded enough there either.[1] Within the former context, the concern is over tyrannical rulership, and the latter concerns, as we shall see, another sort of tyranny that we might label "the tyranny of objectivity." The English astrophysicist Sir Arthur Eddington (1882–1944) used this latin phrase as a rather neat summation of his idiosyncratic scientific epistemology, interpreting it as: "Who Observes the Observers?" That is: since all empirical science starts from our immediate experience (how can it be otherwise?), and spreads out by inference to everything else, no matter how remote (see Eddington, 1928, p. 281), then how are we to account for this elementary fact within science and our scientific theories? We must seemingly pretend that we do not exist and that the universe would go on being its same old self in our absence.

1 Juvenal's discussion was itself a (*very* politically incorrect, by present standards) skit on why men should not get married to women, given the nature of women (which he proceeds to outline in a way that would make even Jordan Peterson blush), especially when there are far more pleasant alternatives available, such as committing suicide. The line about who watches the watchers refers to the clever expedient of having a eunuch watch one's wife. And you thought the MGTOW movement was a recent thing? It was Edmund Burke who in 1756 applied the phrase to government, and the idea that because of the 2nd-order watcher problem we should do away with any and all forms of government in favour of a "natural society against politicians".

Eddington spent his final years working on what he believed was a *truly fundamental* theory, in which the human contributions were fully worked out, so that the invariant structure of the universe (purged of human epistemology) could be laid bare.[2] Eddington lectured on this material at Dublin, producing a pamphlet *The Combination of Relativity Theory and Quantum Theory* (1943) which contains the basic ideas of his *Fundamental Theory*. At the same time, in 1943, and in *his* lectures to the Dublin Institute for Advanced Study (converted into the book *What is Life?* a year later), the physicist Erwin Schrödinger pointed to Albrecht Dürer's painting *All-Saints*, in which we find two circles of believers gathered in prayer around the Trinity elevated in the skies, a circle of the blessed ones above, and a circle of earthbound humans. Within the earthbound humans we find kings and emperors and popes, and, he notes, we also find a tiny portrait of the Dürer himself, figuring as "a humble side-figure that might as well be missing" (1944, p. 137). Similarly, in science we find our models "colourless and soundless and unpalpable".[3] Schrödinger would later refer to this curious omission of the source of knowledge as an "exclusion principle" of sorts radically distorting this knowledge of the world if not taken account of.

The philosopher Edmund Husserl labeled this abstractive methodology "the Galilean style," which, as Steven Weinberg puts it, involves making "abstract mathematical models of the universe to which at least the physicists give a higher degree of reality than they accord the ordinary world of sensation" (1976) — Husserl called the latter world "the life-world." This style is often aligned with the so-called Copernican Principle, stating that we, as observers, do not occupy a privileged position in the cosmos. Yet we forget that the Galilean style is, when viewed more closely, a slap in the face of the Copernican Principle, since it assumes the world we observe is how it is, rather than involving features of the observers themselves. The Galilean style abstracts out too much, and ends up leaving the observer so privileged that we assume we can ignore it, treating it as part of the background. A more thoroughgoing scientific Copernicanism would involve including the standpoint of knowledge-creation or discovery within its models, which would also thereby tame some of the more hubristic elements of modern science.[4]

2 This book was published in its unfinished state as *Fundamental Theory* by Cambridge University Press in 1944. I have discussed his viewpoint in many places, including Rickles (2017) and (2020).

3 Lest the objection that science's ethos involves replacing observers with instruments, with none of the flaws of observers, Schrödinger writes (in his later Tarner Lectures, *Mind and Matter*): "the observer is never entirely replaced by instruments; for if he were, he could obviously obtain no knowledge whatsoever" (1955, p. 162). I would also add that those instruments do not come out of nowhere: they are infused with theoretical assumptions that must be in place for us to believe that they reveal what we take them to reveal: they are themselves 'embodied theories.'

4 This more integrated approach aligns somewhat with Iain McGilchrist's ideas (e.g. in his book *The Master and His Emissary*: Yale University Press, 2012) of the brain's two hemispheres (which also stand symbolically for the world at large, as a kind of East and West), which should be synchronised in order to avoid the excesses of either being dominant (and so balancing subject and object; reason and feeling, and so on).

We find the foregoing sentiment expressed in the cyberneticist Heinz von Foerster's marvellous phrase[5]: "Objectivity is the delusion that observations could be made without an observer."[6] The so-called scientific method makes essential reference to observation, and yet we forget this when it comes to our theories. Both Eddington and von Foerster appear to have shared the same view about the foundations of the scientific enterprise: you must take due account of the observer if you want to properly understand the business of science *and* its produce. The cybernetic approach to systems (including science), specifically when viewed from a meta- or 2nd-order perspective, takes the role of the observer-participator (i.e a bounded, purposive system) seriously in the models we make of the world. It is curious, then, that our models of ordinary systems (such as social systems) appear to now be converging on the same conclusion as our models for the most non-ordinary systems (those of quantum gravity).[7] I will try to explain a little of the overlap here, though to go into detail would require a book.[8]

While von Foerster thought in terms of computational processes (inputs, outputs, and feedback), Eddington, true to his times (and with the likes of Bertrand Russell nearby) thought in terms of *sense data*. But the same problem faces both accounts: what is the nature of scientific epistemology given that it must come from such finite beings, as we are? This is essentially epistemology viewed as a sampling problem. We are dropped into a potentially infinite desert, with no innate map and compass telling us how the world around us really is, and must orient ourselves somehow from our personal data stream of (what we suppose is) a mere fraction of all possible grains of sand. Science, like everything else beyond our personal screen, is a grand exercise in inference. Moreover, even as we develop our scientific theories to help us navigate in this desert, we never get beyond these representations (or maps).

Interestingly, this basic idea (going back at least as far as Hermann von Helmholtz), of the brain as an inference-engine, making best-fit guesses from a data stream, is, like cybernetics, making a strong comeback, with Karl Friston's notorious 'free-energy' approach, with the associated predictive processing/Bayesian brain elements. These approaches share the same tendency to focus the attention of where knowledge is being created (the mind or brain). It considers the possibility that much that we take to be objective about the world out there, is a result of mechanisms that allow us to navigate the world, and might have more to do with Us than It. However, ultimately, it

5 See Ernst von Glaserfeld (1996, p. 279).

6 See Lynn Segal's book (Segal, 2001) on von Foerster for an excellent account of those aspects that relate to this paper.

7 This is not an historical statement, but one that describes the present conditions in quantum gravity research. I am currently engaged, with a PhD student, on a study of Stephen Wolfram's new theory of quantum gravity (involving the space of all possible computational rules) which, again, must, if it is to recover the universe we experience, involve the introduction of computationally-bounded observers that sample this space (see, e.g. https://writings.stephenwolfram.com/2021/04/why-does-the-universe-exist-some-perspectives-from-our-physics-project/ for Wolfram's viewpoint).

8 Some of the philosophical background can be found in Atmanspacher and Rickles (2022) — though there we speak in terms of action-perception cycles, information, and meaning, rather than implicitly cybernetics terms.

breaks down the Us/It distinction, since we become part of the system to be modelled (that is, we model the modeller and use this to understand the models [i.e. theories] that the modeller generates).

Eddington argued, given such ideas, that the theories that result from our observations take on a Procrustean flavour. The physicist, on this view,

> Might be likened to a scientific Procrustes, whose anthropological studies of the stature of travellers reveal the dimensions of the bed in which he has compelled them to sleep. (1936, p. 328)

As Procrustes would make sure any travellers visiting his establishment would fit his 'one size fits all' bed (by stretching them out or chopping their legs off), so the scientist achieves the comprehensibility of the universe by selecting out those parts that *make* it comprehensible. This view Eddington labels "selective subjectivism." We find cyberneticists, such as von Foerster, using "radical constructivism" to describe much the same thing. Likewise, Wheeler's phrase "It from Bit" (on which more below), and his later followers' "participatory realism." Henri Poincaré, in *The Value of Science*, had earlier presented a similar view of the nature of laws and order as part of his thesis of conventionalism. Thus, he writes: "Does the harmony which human intelligence thinks it discovers in nature exist apart from such intelligence? Assuredly no. A reality completely independent of the spirit that conceives it, sees or feels it, is an impossibility".

It is important to note that both Eddington and Wheeler were led to their views by a deep knowledge of the gaps in our foundations of physics, namely the theories of relativity and quantum theory. Because Wheeler, and many other quantum gravity researchers, are on a quest to figure out the deepest structures of our world, they tend to view our current frameworks as *provisional*. It is fairly easy to see how Eddington was led down this path, through his study of Einstein's theories of relativity. For Eddington, these revealed in stark detail the essential role of the observer (the reference frame in this context) in properly interpreting the results and making sure they are physical, rather than mere artefacts of representation. A huge controversy in the early part of classical and quantum gravity research was precisely over whether gravitational radiation (since detected, with a pair of black holes, with Nobel prizes duly awarded) is representation or reality. That is, the result needed to be disentangled from the coordinates an observer uses to describe a gravitating system.

Wheeler's view came more from the side of quantum mechanics, though mixed up with the idea of background independence that we find in the general theory of relativity — i.e. the notion that there are no absolute structures in general relativity, so that all physical quantities must be relational in form. This is sometimes expressed in terms of there being no spacetime existing as a fixed stage on which the actors (matter and energy) play out some performance. Rather, spacetime (the stage) is one of the actors itself. Wheeler essentially extended this to include observers too: they are part of the relational structure, and should not be viewed as standing apart from the stage production, as audience members. But Wheeler also viewed Niels Bohr as the sage of quantum mechanics, and in particular took his view "No elementary phenomenon is a phenomenon until it is an observed phenomenon" as the basic principle of

quantum mechanics. In other words, there is nothing *manifested* in the world (indeed, no world at all) until an observation is made (by a reflecting subject) that *objectivises* it.[9] Hence, there is a kind of relativity or duality (between subject and object) in quantum mechanics too.[10] But for Bohr this was no problem since physics was not concerned with probing deeper than the manifest reality (since such a thing cannot be probed in this way). As he claimed: "In our description of nature the purpose is not to disclose the real essence of the phenomena but only to track down, so far as it is possible, relations between the manifold of our experience" (1934, p. 18).

Wheeler was drawn into *quantum gravity* by the ineluctable logic of following the theory of gravitational collapse into its extreme limit. He couldn't resist probing deeper beyond the manifest, to try and figure out what made it go. Pre-geometry, which is really just pre-physics, is the natural outcome of this route. Why? Because a gravitational collapse of the universe is capable of eliminating everything in its big crunch: all of space, time, matter, *and* law. If so, what sense are we to make of space, time, matter and law, making up our cosmos? Wheeler believed that we must seek a deeper theory of "cosmogenesis" beyond the black hole and beyond the big bang. These cannot be fundamental elements of reality. In understanding the deeper underpinnings of physics, in such a world, John Wheeler notes that:

> In brief, we confront two imperatives and one great issue. First, the gates of time tell us that physics must be built from a foundation that has no physics; or still more briefly: "Must Build." Second, elementary quantum acts of observer-participatorship: "Do Build." Finally, how are billions upon billions of these elementary building acts organized — if they are to make up the grand structure that we call "reality"; or, in brief: "How Build?" No more attractive clue offers itself for attacking this great issue than the way information is processed to make "meaning." On what else can a comprehensible universe be built but on the demand for comprehensibility? (Wheeler, 1980, pp. 6–7)

Comprehensibility by whom? Us, of course — hence, 2nd-order cybernetics. Where there is information, there is meaning. Where there is meaning, there are creatures like us required to interpret it. Since Wheeler's viewpoint (known colloquially as "It From Bit") puts information and meaning at the very deepest layer of reality, we must also place there the *interpreters* that provide meaning to the bit strings. Interestingly, this view of information as the correct fundamental ontology of the world is becoming near-orthodoxy — developments in CRISPR, AI, quantum information theory, and so on, are propelling this view to the centre even more.

9 We should not underestimate Bohr's deeper views, especially on account of his decision to embed the *Taijitu* (Yin-Yang) symbol at the centre of his coat of arms. Bohr strongly believed that without a subject there was no object, and vice versa. It was this aspect of Bohr's approach to quantum mechanics that so disturbed Einstein, *not* the probabilistic nature of the theory.

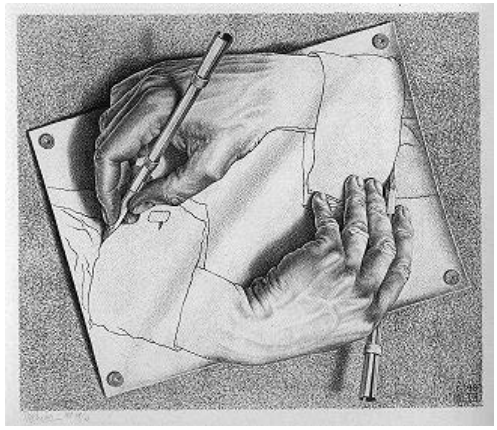
10 We can find several others espousing variations on a theme of this basic idea of reintroducing the subject or observer, e.g. David Bohm and more recently the cosmologist James Hartle, with his scientific model of an observer known as an IGUS: an "information gathering and utilising system." But it remains the exception rather than the rule.

But, to repeat, it is not correct to say that the world is *made of information*. The world outside of such creatures as Us does not contain *any* information as such since it involves a relation between It and Bit and Us.[11] The science of information developed by Claude Shannon and Warren Weaver (1949), especially as presented today, might have you believe that things are otherwise, but we must remember that Shannon was concerned with transmission efficiency and the elimination of redundancy to achieve this. Yet even in Shannon's works, for example on the identification of redundant elements from messages to code more tightly, the criterion for removing redundancy must be referred to a criterion of synonymy that ends with Us once again. Without Us, there is no way to distinguish one bit string from another; nor is there a way of seeing how one and the same bit string can represent an infinitude of possibilities.

As in cybernetics, this way of thinking often leads to loops and cycles. Roger Penrose puts the circuit (which Wheeler calls the "meaning circuit") linking subject (epistemology) to object (ontology) as follows:

> We have a closed circle of consistency here: the laws of physics produce complex systems, and these complex systems lead to consciousness, which then produces mathematics, which can then encode in a succinct and inspiring way the very underlying laws of physics that gave rise to it. (Penrose, 2004)[12]

If we take this participatory view of reality seriously we face a kind of construction by us that in turn constructs us, with each pulling the other into being much as Baron von Munchausen pulls himself and his horse out of a swamp by his own hair.[13] As if mirroring the Eddington quote with which we began, let us close with T. S. Eliot: "The end of all our exploring will be to arrive where we started."



*Drawing Hands* — M.C. Escher

11 However, in the aforementioned book (Atmanspacher and Rickles, 2022) we make a case for a kind of primordial or latent meaning that sits underneath both It and Bit, and subject and object.

12 Penrose writes elsewhere "We are all part of the world, and we are conscious beings, so if the world is itself describable mathematically, then this whole idea of conscious perception must be describable mathematically" (interview of Penrose with Alan Lightman, 1989: https://www.aip.org/history-programs/niels-bohr-library/oral-histories/34322). This kind of 'One World' framework is very much influenced by quantum gravity research whose entire motivation is on the idea that, despite the lack of direct empirical evidence demanding such a theory, since there is one world, there should be a framework that can handle quantum and gravity together. To expand out to consciousness is just another step.

13 The picture *Drawing Hands* of Maurits Cornelis Escher (see figure) provides a perfect visual representation of this co-creation. What's more, it constitutes a kind of new existentialism (quantum existentialism) since, existentialism in its orthodox sense refers to the absence of absolute essences. There is no ready-made world out there. This can be translated into the idea that the division between subject and object is not fundamental, with subjects as much as objects involved in the world's construction.

## Acknowledgements

I am very grateful for funding from: The Foundational Questions Institute and Fetzer-Franklin Fund, a donor advised fund of Silicon Valley Community Foundation [FQXi-RFP-1817]; The John Templeton Foundation [Grant ID 62106]; and the Australian Research Council [Grant DP210100919].

## References

Atmanspacher, H. and D. Rickles (2022) *Dual-Aspect Monism and the Deep Structure of Meaning*. London: Routledge.

Bohr, N. (1934) *Atomic Theory and the Description of Nature*. Cambridge: Cambridge University Press.

Eddington A. S. (1920). *Space, Time and Gravitation: An Outline of the General Relativity Theory*. Cambridge: Cambridge University Press.

Eddington, A. S. (1928). *The Nature of the Physical World*. Cambridge: Cambridge University Press.

Eddington, A. S. (1935). *New Pathways in Science*. Cambridge: Cambridge University Press.

Eddington, A. S. (1939b). *The Philosophy of Physical Science*. Cambridge: Cambridge University Press.

Eddington, A. S. (1943) *The Combination of Relativity Theory and Quantum Theory*. Communications of the Dublin Institute for Advanced Studies. ISSN Series A (Theoretical Physics) 0070-7414.

Eddington, A. S. (1946). *Fundamental Theory*. Cambridge: Cambridge University Press.

McGilchrist, I. (2012) *The Master and His Emissary: The Divided Brain and the Making of the Western World*. Yale University Press.

Husserl, E. (1970) *The Crisis of European Sciences and Transcendental Phenomenology* (Trans. David Carr). Northwestern University Press.

Penrose, R (2004) *The Road to Reality: A Complete Guide to the Laws of the Universe*. London: Vintage.

Poincaré, H (1905/1958) *The Value of Science*. Dover Publications.

Quine, W. V. O. (1951). Two Dogmas of Empiricism. *The Philosophical Review*. 60 (1): 20–43.

Rickles, D. (2017) All Possible Perspectives: A (Partial) Defence of Eddington's Physics. In I. Durham and D. Rickles (eds.), *Information and Interactions: Eddington, Wheeler, and the Limits of Knowledge* (59–72). Springer.

Rickles, D. (2020) *Covered in Deep Mist: The Development of Quantum Gravity* (1916–1956). Oxford University Press.

Schrödinger, E. (1944) *What is Life?* Cambridge: Cambridge University Press.

Schrödinger, E. (1955) *Mind and Matter*. Cambridge: Cambridge University Press.

Seagal, L (2001) *The Dream of Reality: Heinz von Foerster's Constructivism*. Springer.

Shannon, C. and W. Weaver (1949) *The Mathematical Theory of Communication*. University of Illinois Press.

von Glaserfeld, E. (1996) Farewell to Objectivity. *Systems Research* 13(3): 279-286.

Weinberg, S. (1976) The Forces of Nature. *Bulletin of the American Academy of Arts and Sciences* 29(4): 13–29.

Wheeler, J. (1980) Pregeometry: Motivations and Prospects. In. A. Marlow (ed.), *Quantum Theory and Gravitation* (pp. 1–11). Academic Press.