JNphi

Abstraction and the Explanatory Gap: Physicalism and Dualism Combined

Michael Remler

Abstract

The explanatory gap between consciousness and science can be understood only in contrast to those things with an adequate scientific explanation. Scientific explanation is built on the measurement of time, distance and mass etc. and relationship concepts such as true, equal to, etc. all of which are derived by the collective abstraction of many subjective experiences. Measurement and mathematics, because they are uniform across the whole of humanity, create an abstract symbiosis of all the separate consciousnesses. Those concepts supported by the symbiosis are of unique authority from which accepted explanations can be derived. Although they retain not one iota of the subjective experiences from which they were abstracted, they retain the authenticity of the subjective experience. The explanatory gap regarding consciousness derives from the fact that there is no analogous abstraction of the totality of mental processes of the human mind, in particular consciousness. Turing's defined 'thinking' as that portion of human mental function that can be represented by a computational algorithm and that abstraction validated by the imitation game. To define the limits of symbiotic abstraction of human mentation, and document that via the imitation game, it is needed to further develop that approach to model the totality of externally observable individual behavior. Consciousness and its associated features like free will, qualia, etc. are intrinsically not observable and cannot be abstracted directly into the symbiosis. Non-observables entities such as black holes, quarks, etc. are identified, validated and explained as the most parsimonious understanding consistent with the structure of ideas anchored in the observables. Consciousness is irreducibly idiosyncratic and non-material and therefore irretrievably not possible to directly abstract. Therefore, a Materialist explanation within the current understanding is impossible. Dualism is the most parsimonious theory but fails for lack of a plausible interface with current physics. Entanglement and the mechanism of quantum collapse are established phenomena within physics not by any known material mechanism. Further understanding of these phenomena may provide the conceptual basis for an abstract non-material dualist model of consciousnesses with no explanatory gap. A model which both physicalist and dualist.

Key Words: explanatory gap, abstraction, consciousness, Turing, physicalism, dualism

DOI: 10.5281/zenodo.15259742

Corresponding author: Michael Remler
Address: Dept. of Veterans Affairs, 150 Muir Rd., Martinez, CA 94553, USA
e-mail ✉ mike@remler.com

Introduction

The explanatory gap (Levine, 1983) is a term developed and utilized to describe the problems of relating individual subjectivity (Nagel, 1974) and collective objectivity, most especially science. To engage the issue two antecedent problems must be addressed. 1) What is an 'explanation'? 2) What are the foundations and validity of the conceptual tools of Science with which it can seek to 'explain' subjectivity?

Popper defined three forms of reality, three worlds (Popper, 1992). World 1 is physical reality. However, he asserted as equally real, ideas. Ideas exist in two separate venues, two separate realities. First in the minds of individuals, a part of each person's separate subjectivity, and are unavailable to any other person, World 2. The ideas of any one individual are completely autonomous of any other subjectivity. Ideas also exist in symbolic representation in the common culture of communication, available to all participants in that common culture, World 3. Popper focused on asserting the independent existence of each world as manifest in their capacity be the cause of an effect. He explicitly avoids the issue of physicalism versus dualism as the mechanism of action of world 2 in world 1 but insists that the origin of consciousness, world 2 is a product of the creative action of world 1. He asserts that world 3 arises from world 2 before it becomes existentially independent, but avoids any detail regarding the mechanisms of its creation.

Explanation is an intra world 3 construction. An explanation is constructed of objects and relationships. The explanatory gap exists because the objects of world 2, subjectivity such a consciousness, qualia, affect, etc., even though the word exist in world 3, are incompletely, ambiguously or even incorrectly formed.

A related but distinct issue avoided here is the distinction between the objects of world 3 as representations of the object in world 2 derived from perceptions of world 1 or related in some way directly to the object in world 1. The doctrine of Scientific Realism is that the ideas of world 2 about world are latent in physical reality, world 1 (Larenz 2024). Alternatively, one can emphasize all experience with reality begins as the subjective experience of individuals (Dowling, 2021). Berkeley said that all knowledge begins with subjective perception, *esse est percipi*, 'to be is to be perceived' (Berkeley, 1948). Human beings even know they exist because they perceive their existence, their subjective thoughts (Descartes, 1988). Existence, consciousness and perception are so intimately entwined as to be essentially inseparable. Either way, the linkage between physical reality world 1, human perceptions and thoughts about that reality world 2 is broadly accepted. Therefore, the linkage of world 3 to world 1 is effective.

The essence of the explanatory gap is exposed by close examination of the origins of world 3 from world 2 and the mechanism of its

independence from world 2. A green light, as a physical reality (world 1), can be perceived by humans (world 2) and then utilized as a traffic signal (world 3) and therefore function reliably to control traffic on the street. But no one has any idea what the subjective perception of green, the idea of green, the qualia of green (world 2), is to any other person, but themselves. The green of world 3 did not arise *ab initio* directly from world 1 but rather via the ill-defined nature of world 2.

The creation of the linkage across the gap between worlds 2 and 3 may be described in four components. First is the creation by the physical reality (world 1) of the subjective perception in individuals (world 2). Secondly, individuals with the subjective perception (world 2) create a public objective symbol, word, idea (World 3) to represent subjective experience. Thirdly, the aggregation of multiple individual subjective experiences, world 2, via the mechanism of language to create a shared concept in world 3. Finally, to create the linkage, the education of subsequent individuals regarding the public objective idea structure to build a correlation to their individual subjectivity, reinsertion into world 2. The purpose here is to contrast a variety of phenomena in the physical world 1 which are accepted as explained in science, world 3, as compared with features of subjectivity, world 2, which are perceived to be beyond such explanation. It is the third step which creates the broad and intimate relationship between worlds 2 and 3 such that the distinctions can be subtle and elusive.

The critical feature of the explanatory gap derives from which world 3 symbolic conceptual structures relate to world 1 in such a way that they can be shared by all members of the collective. In contrast, world 2 subjective phenomenon, are private and unavailable to be shared, and fundamentally idiosyncratic to each individual. This gap between the intrinsically subjective private versus the public objective is both fundamentally irreducible and central to the issues of explanation even while in common use they are invisible.

Therefore, before we can examine effectively the explanatory gap, we must understand the basis of explanation where there is no gap and explanation is accepted as effective and complete. Science, a portion of world 3, the standard of effective and complete explanatory power begins with Newton and the relationships of the laws of mechanics and gravitation. The units of the terms of mechanics, 'time', 'mass' and 'distance', and the equations governing their relationships, are at root the abstraction into public objective quantitative measurable form, world 3, of subjective experience, world 2. The explanatory power of science rests on the preservation of the authenticity of the subjective experience, world 2 into the abstracted quantitative scientific units, world 3.

Abstraction from Subjectivity

The essential first step is for individuals to share their separate *perception*. Perception here is used to mean any subjective mental state, anything an individual perceives. However, central to this discussion is the distinction between those perceptions derived from external physical reality, world 1, and those unlinked to external reality, originating in world 2. Critically important, it is of course impossible to actually '*share'* the perception, the subjective experience, itself. Next what is done is to create a *'symbol'*, a manifestation in public, to representation of the perception, classically a sound, a word (Locke, 1975). Equally fundamental is the *'delineation'* of different symbols, words, for different perceptions. The meaning of the symbol is then repetitively *'correlated'* over the breadth of the collective to reduce and eliminate any public ambiguity and confusion. That process of strengthening the collective meaning of the symbolic representation of the perception is possible only because not one iota of the individual idiosyncratic subjective perception itself is involved.

Perception, sharing, symbol, delineation, and correlation are presented in conceptual order, but there is no necessity to believe that there was any specific historical separable sequence. Aspects of this process existed animals (Suzuki, 2016). In the millennia of human cultural evolution, various portions of the process have no doubt coexisted and overlapped. The net result is two-fold: 1. A public symbol is created and accepted with decreasing ambiguity across the collective. 2. An unambiguous bond is created within each individual between their subjective perceptions, and the correlated objective symbols. However, there is no reason to believe there is any uniformity amongst the subjective perceptions which are bonded to the same public symbol. This apparent anomaly exists and the symbol works *because* no aspect of anyone's subjectivity itself is present in the symbol. The continuing autonomy of the subjective qualia from the public symbol is made most unambiguous by the phenomenon of synesthesia, where some individuals perceive an input associated with one modality such a sound with a qualia associated with a completely different modality, such as color (Banissy, 2014).

A subset of symbols to describe the relationship amongst other symbols; words like cause, true, logical, effect, etc. They are developed to correlate with the subjectivity that the relationship *'makes sense'*. Of course, as with all perceptions, makes sense can be very different things in each individual. But worse, any statement about it is intrinsically tautological. The only thing that can be said about making sense is that it makes sense.

The validity of the symbols derives from its power to coordinate human behavior, but it has no independent meaning or authority. An enormous variety of the subjective life of individuals can be represented usefully in the symbolic structure of language. The

richness of that representation is the essential step to the coordination of large numbers of people and a functional society. Nevertheless, the unknowable and irreducible heterogeneity of subjective experience, correlated to the same delineated symbol results in an irreducible ambiguity of meaning of the symbols.

The ideas of world 3 are said to *represent* the ideas of world 2 (Lycan, 2023). However, that representation may be said to be generally of two types; One, ideas derived reasonably directly from external reality, world 1, and those generated internally in the mind, world 2. It is that distinction which determines the explanatory power of world 3 ideas regarding world 1 and the failure of explanatory power regarding world 2.

'The aggregation and correlation of multiple individual subjective experiences, world 2, via the mechanism of language to create a shared concept in world 3' demands close examination. Separate subjectivities, world 2, with absolutely no knowledge of other separate perceptions, correlate an element of physical reality, world 1, with a symbol, world 3. It is obviously impossible to produce a comparable correlation of their own subjective perceptions since those subjectivities do not exist in world 1, only world 2.

The tiniest fraction of the symbols, world 3, correlated with elements of physical reality, world 1, correlated closely via subjective states world 2, can be rendered into a measurement, an unambiguous relationship of world 1 to world 3 fundamentally bypassing world 2. Time, distance, mass etc. are uniquely important amongst the variety of human symbols, elements of world 3, not merely because they represent aspects of world 1, but because they can be 'objectively measured', invariant to the idiosyncratic, private nature of any one consciousness of world 2, unknowable to the other elements of world 2. Thus, the measured scientific units are completely different entities in world 3 from any other symbols representing aspects of world 1 as they are uncontaminated by subjective perceptions, world 2. Scientific time is external to human consciousness and subjectivity. The idea of time exists in two completely separate and *autonomous* conceptual structures, the objective and the subjective, or as Wiener has called them, Newtonian and Bergsonian time (Weiner, 2019). However, Bergsonian subjective time is existentially, historically and conceptually first (Bergson, 1910). But as discussed below, Newtonian objective time is essential for explanation.

In addition to those perceptions which arrive *ab initio* from subjectivity and are rendered independent of subjectivity by measurement, the perception that the ideas fit together, make sense is also refined into mathematical and logical ideas independent of subjectivity. And critically, those mathematical ideas were linked by repetitive correlation to the subjective sense that something *makes sense.* Mathematics' autonomy from subjectivity derives from its absolute lack of ambiguity. The combination of measurement and mathematics

creates a structure of ideas, totally autonomous from the subjectivities from which they were derived and miraculously accurate in relation to world 1 (Wigner, 1960).

Scientific Explanation - Subjective Collective Symbiosis

An 'explanation' is a structure of ideas represented in a public symbols, world 3. Its importance derives from when transmitted as a communication it evokes the associated ideas in the subjectivity of many individuals, world 2. More specifically, it includes the relationship amongst the ideas. Thus, it requires the transfer of both perceptual and relationship ideas from the objective venue to the subjective. And in that transfer the objective abstraction assumes the authority that belonged originally to the subjective experience and thought processes of the separate individuals (Givental, 2024). The authority of direct experience, which existed initially only in idiosyncratic individual subjectivity and was transferred by the process of correlation via world 1 now reasserted by objective ideas, world 3, into individual subjectivity, world 2. This creates within world 2 a new entity, not each separate consciousness, and not any aggregation of still separate consciousnesses, but rather a Symbiosis of the separate consciousness structured on the belief, internal to each that certain elements of their unique subjectivity are substantively similar in the other subjectivities.

Fundamental to this discussion, is that what has been transferred from subjectivity to objectivity and then back is not the essence of the original subjectivity. That was lost. There is no essence of time transferred from the subjectivity of individuals to the clock and therefore what is delivered in explanation is validated by the status of the objective ideas not any essence of the original. We may say that what was transferred was 'explanatory power' but that is misleading. The power to understand time, distance, force remains in the individual. The measurements and the equations that use them have no comprehension of the essence and so therefore they can hardly explain the essence. The Symbiosis does uniquely have the authority to commandeer in each individual their actual experience with the subjectivity and utilize that commandeered understanding via rational thought such that the individual can feel they understand something separate and distinct from their direct experience. That is explanation.

Science is that portion of the Symbiosis built on measurement and mathematics. What gives the objective idea structure built on measurement and mathematics the authority to command the ideas of subjectivity, specifically that the relationships described make sense? Paramount of course is the enormous success of science and technology in effecting the world (Wigner, 1960). Understanding is at its essence the sensation that is set of '*ideas*' '*fit together*' to '*make sense*'. The success of technology validates that understanding.

Fundamental is that all three components of explanation are not primary experiences. Rather, they are the internal subjective construction in each individual human consciousness by education of the abstractions that were originally derived from many long-gone people's subjectivity, but now with no elements of such subjectivity itself. It is individual development and education that create a cognitive link between the authenticity and existential reality of primary experience, world 2, in each individual and the objective idea structure, world 3.

Three features emerge which are not evident in the derivation of the Symbiosis. The first, is that explanation is created in the Symbiosis, the collective objectivity, and delivered to each separate individual subjectivity. The other people's subjectivities are no longer involved. Secondly, as Newton immediately demonstrated, the method, iconically the *Universal* Law of Gravitation, could be applied, and carry that mantle of truth, to questions outside the range of ordinary human experience, the movement of the planets. Now that is extended to quarks. The third, as stated, explicitly in Newton's, three laws of motion (Smith, 2008), force was the *cause* of the consequences in the motion of objects. Where Kepler merely *described* the movement of the planets, Newton *explained* the movement of the planets was *caused* by of gravitational force. However, when force is created by what is subjectively an act of will, it may be said to be a manifestation of human consciousness. When force is created by the law of gravitation between two presumptively unconscious pieces of mass, then it is not a cause, but simply a correlation (Pearl, 2018).

Inner Limits and Outer Expanse of Explanation

The subset of world 3 whose links are built on measurement and mathematics world 3, have a unique power derived from the fact that is substantively identical in each of the individuals, world 2, and they correlate with the external reality, world 1. This is despite the fact that the essence of subjectivity, the qualia, remains idiosyncratic in each individual. It creates special portion of world 2, a Symbiosis amongst the many separate individual's world 2, linked by their commonality in worlds 1 and 3, because the power of their separate link to the collective ideas is strengthened by the knowledge that the link in everyone else is identical. The public component of the Symbiosis is possible if and only if that which is to be abstracted can be rendered into a public structure that can be shared without ambiguity. The Symbiosis creates the unique and essential components of the explanatory power.

Abstraction makes possible the Symbiosis which makes possible Explanation. Anything and everything linked fundamentally to subjectivity such as religious faith, the beauty of Michelangelo's David etc. cannot be explained.

Time is a very ephemeral reality as best. Even before the explanatory gap, regarding the nature of consciousness arises, the same fundamental issue may be seen in the subjectivity of time. In any one individual, regarding any one particular aspect of their experience, their subjective perception may or may not correlate well with the fully refined abstraction (Wiener, 2019). They accept both 'that is how they feel', world 2, and the validity of chronological scientific time, as a matter of physical reality, world 3. As a result, reality as understood by each individual, is bifurcated into those portions which fit into the abstract structure, those that do not. We accept that bifurcation because we must. *Those elements of experience that do not fit into the Symbiosis, the abstract structure, cannot be explained.*

In contrast, things that are beyond experience, but do fit into the abstract structure are considered to be explained. Once the Symbiosis is established, the relationship of individual cognition to the collective is bifurcated into two parts: the part where the Symbiosis applies contrasted to those issues for which the Symbiosis cannot be created. Where the Symbiosis is created, the power and authority between each individual and the collective shifts totally to the collective. So great is the shift in authority that it is the first article of faith in scientific realism that the mathematical structure developed to explain the physical world was in fact latent, preformed, in physical reality before it was understood (Larenz, 2024; Wigner, 1980; Spinoza, 2002).

Education by the collective of each and subsequent individuals entrenches the dominant of the objective perspective but only for those matters for which the Symbiosis was established, or those matters where the lack of ambiguity can produce uniformity amongst the many separate individuals (Givental, 2024). Since objects appear to create virtually all the cause evident in the world, the question arose is there anything else, such as human will, that can function as an autonomous cause (Laplace, 1995).

Most subtly and most revolutionary, the Symbiosis co-opted to itself, the most profound and elusive features of subjectivity, existence. You cannot separate the subjective perception of time, from the subjective perception of existence. You cannot separate that perception from consciousness. For each individual, the perception of their own existence is consciousness and subjective perceptions. To make it most explicit, Descartes said *Cogito Ergo Sum* (Descartes, 1988) meaning there was no distinction between his subjective perception and fact of the existence of his mind, his consciousness as a separate reality, existence. He expected his readers, everyone else, to accept that as fact. But after Newton, after the rise of science, its autonomous existence as a set of ideas, the products of Descartes' subjectivity, dualism, the primacy of subjective reality, existence, no longer carry the argument. It is this shift of the arbiter of reality from individual subjectivity to collective ideas which creates the 'Explanatory Gap'. Therefore, when we establish the authority of the Symbiosis, we take

with it the right to define the reality of existence. Science, not subjectivity, became the arbiter of existence (Givental, 2024).

Now, the existence of human consciousness, world 2, must be discovered by science, the public ideas, world 3, as if it were a part of world 1. It has discovered black holes, the process of evolution, and quarks, all entities accepted and explained by science without any possibility of direct perception. The acceptance is true of concepts which are either beyond perception and/or contrary to apparent perception was revolutionary (Kuhn, 1962). The prerequisite for revolution is the existence of a crisis because existing explanatory structures cannot explain or contradict an observed phenomenon. The public observation of human beings includes that they claim consciousness and free will, but there is no explanation within world 3.

The Explanatory Gap is when the Symbiosis is asked to explain individual people's consciousnesses which cannot be perceived by any other consciousness. No correlation is possible and therefore from which there is no abstracted representation. It is on this basis that some claim that since consciousness is private and cannot be subjected to correlation, it is an illusion (Dennett, 1992). The abstraction methodology and metaphysics of science imputes the existence of black holes and quarks for which no perception is possible from the observable behavior of matter. That imputation depends on the *rule governed, reproducible, regularities* in the *observed behavior*, in world 1. It is agreed that there can be no world 3 explanation of world 2 because there can be no process of correlation between private idiosyncratic subjectivities, world 2. The problem must be reformulated. The challenge is to find those features in human behavior which can be extracted and correlated to become effective scientific, measurable concepts in world 3 and then use them to impute the existence of world 2 mentation.

Abstraction of Individual Thinking

An explanatory structure possessed of enough power to be accepted as an explanation of any part of human mentation must possess enough rule governed content, such that the abstraction of the world 2 individual into an objective world 3 concept can be effectively shared with the collective. The purpose of the above discussion is to make clear that explanation exists in two phases. The first phase is to change an idiosyncratic personal subjective perception, via a process of abstraction and correlation into an objective public idea. The second phase is to utilize the abstracted ideas, preferably in quantitative form via mathematics and apply it to issues unrelated to the perceptions from which it was abstracted. Consciousness is, *ab initio*, absolutely no more or less subjective than time, etc. Therefore, the essential problem is how to abstract consciousness into an objective public idea.

The first step to the abstraction of consciousness was famously taken by Alan Turing (Turing, 1950) in asking the question 'Can Machines Think' and answering it with the public abstraction of a Computational Algorithm Machine (CAM). Beyond that, he created a concept for comparison and correlation of the performance of the CAM called the 'Imitation Game' (IG). Computational algorithm can be compared and correlated with the mental functions of a human being by asking them to simulate human interactions to the point of being indistinguishable from the human being (Jones 2025).

Turing sharply defined the limits of his goals, well short of the abstraction of the totality of human consciousness. He states, explicitly,

> "I do not wish to give the impression that I think there is no mystery about consciousness. There is, for instance, something of a paradox connected with any attempt to localize it. But I do not think these mysteries necessarily need to be solved before we can answer the question with which we are concerned in this paper (Turing, 1950)."

Turing asks if machines can 'think', not can they perform all the capacities of the human mind, let alone be conscious. Indeed, from the beginning, he makes it clear that he wishes to abstract the essence of that component of human intelligence that matches symbol manipulation (He calls it 'thinking'.) by a CAM and demonstrate the correctness of the abstraction by the IG, in comparison to actual people. He specifically avoids any effort to define thinking or intelligence by any other means. As the discover himself of the relevant theorems, Turing knew perfectly well that machines can perform any well-defined symbol manipulation. The purpose of the imitation game is to demonstrate that symbol manipulation corresponds closely to what is understood as human 'thinking'. Therefore, a CAM that can pass the TT in the IG, match that portion of the generality of human intelligence in the imitation game may be said to effectively abstract 'thinking' from the complexity of human behavior as a computational algorithm. Thinking defined this way is an unambiguous public abstraction of a major portion of human mentation.

However, even the generality of human of symbol manipulation is but the merest fragment of the specificity of each separate idiosyncratic individual human mind. A crude analogy might be to say a CAM that meets the TT is to a specific human thinking, as the phrase 'it's a thing' is to the precise quantitative measurement of mass on a scale. Consider an Individualizable CAM as one that having interacted with a specific individual for a period of time and afterwards cannot be distinguished from that individual using simple communication by an imitation test. The general individualized CAM, trainable to mimic a specific individual, maybe thought of analogously to a clock, a measuring rod, or scale, etc., and after training to mimic a specific Individual be thought of as analogous to a specific measurement of 5.3

seconds, 2 meters, etc. In the modified Specific Individual imitation game, where the machine is designed to imitate one specific person, the interrogator is asked to determine which is the real specific person and which is the machine designed to abstract the total intelligence of that specific person. Thus, although neither Turing nor current machine implementations of human intelligence have gone there yet, in the spirit of this discussion, the individual and idiosyncratic nature of one person's 'thinking' may be compared and correlated with the thinking of other people.

Abstraction of Intelligence Beyond Thinking

It has been strongly argued, based on Gödel's incompleteness theorem and related ideas, that there are severe limitations on any CAM (Penrose, 1989). The belief that machines cannot be creative in the way that real human beings are, arose at the very beginning of modern computational methodologies (Turing, 1950, *Lady Lovelace's Objection*). Contemporary methods of machine learning including iterative feedback can in the most minimal sense say something that has not been said before. The creative objection is already shown to be, at least in the limit, false.

Neural networks and other parallel processing implementations and quantum methodologies offer the potential of massively increased speed and efficiency of calculation to produce qualitatively different results that cannot be envisioned by current methods (Penrose, 1994). Words like realistic, aesthetics, virtuosity, genius, and others reflect the performance of tasks which may be within the range of an algorithm but with an implementation of a totally different order. These performance characteristics constitute the extended range of human intelligence, without including the issue of subjective consciousness, which are beyond the range of thinking as defined by even the best contemporary CAM. Following Penrose, let us allow that these machines, can reproduce expanded portion of human mental capacity, call it the 'total intelligence'. And it can project a hologram of a person which looks, sounds and moves like the original, and can be trained to mimic essentially all individual human behavior externally manifest. We may assume that this advanced machine with its total holographic simulation of an individual could pass the Total Turing Test (TTT), a side-by-side conversation with the hologram simulation and the real individual it is programmed to mimic. There is no substantive difference between the hologram suggested here and the more popular philosophical construct of a zombie, an atom by atom physical reconstruction. In both cases the hologram/zombie can appear to be human by external examinations but with no reason to believe from the methods used for construction that it has consciousness etc.

It is often said that there are mental capacities of real human beings, which no machine will ever be able to perform such as express feelings, humor, resilience etc. But what if the machine is programmed to discuss 'its deepest thoughts and feelings', 'it's conflicts with its mother', and 'its sadness when she died'. Even though you know, it didn't have a mother, it would pass the TTT as well as you or your best friend. Is there still an explanatory gap? Perhaps if one accepts that externally verifiable appearance is the only reality. The fact that we know the hologram/zombie is lying changes nothing as we know human beings can also bluff and lie. Clearly there is no explanatory gap regarding the 'expression of emotion', but equally clearly the behavior and expression of subjective affect is not the 'reality of subjective emotion' any more than the hologram/zombie that mimics every nuance of a person's behavior is the reality of the person. There would be appear to be no explanatory gap between scientific understanding and the totality of human 'intelligence' even while the reality of subjectivity and consciousness remains untouched and completely unexplained.

This exposes the fundamental flaw in the standard symbiosis of subjective and objective manifestations of the same part of reality that is papered over by scientific realism (Larenz, 2024). The truth value of subjective reality which cannot be effectively abstracted from the behavior into a Symbiosis cannot be affirmed or contradicted by the absence of that reality, world 2, in the public objective intellectual construct, world 3.

The current capacity of machines implementing AI Large Language Models to simulate human conversation, and it's continuing rapid advance, suggest that the problem of total simulation is already solved, or almost. Indeed, it is claimed by some, that's such AI robots possess all the mental capacities to be regarded as human, and therefore entitled to human rights on a par with physical human beings (De Graaf, 2021). But in fact, they are not designed to have, nor do they even have, any of the mechanisms and capacities of human intelligence (Bishop, 2008). They are intended to be pure façade, a simulation, or as Pearl says, mere "curve fitting" (Pearl, 2018).

One may legitimately ask, is the façade, the maximum extent of the world 3 scientific Symbiosis the ultimate reality? Is the external behavior of a human being, a complete and accurate reflection of the interior mental activity of any actual person? After all, the façade, the external behavior of a human being, is all that any of us know about anyone else except ourselves. It is at best problematic to grant to a robot or a hologram the same rights that belong to a human being without any confidence or reason believe that the interior corresponds in any way to those features of human being, which are the basis of empathy, consciousness and free will, and human rights. Is there Truth, such as the existence of consciousness, beyond all Knowledge, that which can accepted by the maximum imaginable Science?

The critical point, is that from the contemporary scientific perspective, the Symbiosis, those parts of world 3 which cannot be rooted in and abstracted from world 1, does not exist. That is the argument against Cartesian duality, the view that consciousness is a different stuff than ordinary matter. Machines may *claim* consciousness and Free Will. Individuals may assert the uniqueness of their consciousness, and the exercise of free will but if they cannot be abstracted into the Symbiosis from observable behavior that assertion will be outside the mechanisms of contemporary science, and outside the metaphysics of scientific realism. World 3 reality is the Reality (Givental, 2024).

Free Will and Consciousness?

There are two broad understandings of a speculated gap between 'total intelligence', and the full capacities of human mind. 1. Conceivable capacities of human consciousness, externally observable and but presumably not mediated through any currently understood physical properties and therefore impossible to simulate. Consider mental telepathy. 2. Properties of the human mind available to introspection which seem unique in being problematic for any machine, 'free will', which is defined as an apparently unpredictable expression in public behavior, and 'consciousness', and sense of subjectivity itself.

Not Mediated by Matter: Perhaps the most speculative possible limit on the abstraction of the totality of human mentation by a machine is that of effects of consciousness not mediated by matter employing currently understood mechanisms. Turing himself considers this possibility (Turing, 1950; *The Argument from Extra-Sensory Perception*). We should not assume that all fundamental mechanisms are already known. Modern physics faces vaguely similar issues and explaining quantum collapse and entanglement. Most clearly in an entanglement, an effect on one particle effects another particle, at an arbitrary distance away, by no known physical mediation. Einstein called it "spooky action at the distance" (Einstein, 2004). Consciousness has been called "the ghost in the machine" (Ryle, 1948). Both explanatory gaps will require a new understanding of the physical world to include non-material entities interactions with matter.

Consider if physics to *explain* quantum collapse and entanglement understood the existence of a mechanism without mass, call it a new property or a new thing, which had no mass, no localization and could effect and detect changes in the quantum states of matter. Consciousnesses can be modeled by that kind of source of Entanglement with brains (Remler, 2019). If those same sources could entangle with other pieces of matter, that would be 'extra-sensory perception'.

Subjectivity: The outer boundary 'intelligence' is defined by the modified single person imitation game which is indistinguishable from

the original by all external means. We can ask for high affect riffs on its parents or lovers. If it would make a difference, we might also demand that the hologram be 'dissectible' on command. Absolutely, it would assert that it had Consciousness with Free Will. There is no explanatory gap regarding the 'appearance', 'the simulation', of Consciousness and Free Well. The simulation can include a very convincing discussion of its mother even though we know it does not have a mother. The substantive issue is how to distinguish the real person with subjectivity who has a mother and the simulation of that person solely from the external behavior.

Subjective Consciousness and Free will are by definition, not rule governed. Counterfactual, thinking is thought by some to be intrinsically human, associated with free will, and potentially rich in non-programmable results (Starr, 2022). In general, it is claimed that no rule govern system can reach the open-ended, unpredictable, ability of the human mind (Penrose, 1994). The introduction of a random variable may simulate free will in that the activity of the program will not be predictable, fully rule governed. Such random events may be thought to be the mechanism of creative intelligence in the same way that random mutations, are thought to be creative in the process of evolution. But the unpredictability of the human mind is not random but based upon an understanding of the ideas that make predictability. Free will is the ability to understand what is predictable and then not to be governed by it. This is articulated most clearly by Dostoyevsky;

> Good heavens, gentlemen, what sort of free will is left when we come to tabulation and arithmetic, when it will all be a case of twice two make four? Twice two makes four without my will. As if free will meant that! … Twice two makes four seems to me simply a piece of insolence. Twice two makes four is a pert coxcomb who stands with arms akimbo barring your path and spitting. I admit that twice two makes four is an excellent thing, but if we are to give everything its due, twice two makes five is sometimes a very charming thing too (Dostoyevsky, 1989).

Obviously, to identify an unexpected response based on *understanding* from a random response, the machine must also understand the issues involved. It is not clear how one could distinguish unexpected answers, a non-repeating signal, derived from understanding from unexpected answers, derived from a random variable. The singular advantage a straight forward dualist understanding of subjectivity and free will is that it is the most parsimonious explanation in face of the universal report oft that by all human beings.

Explanatory Gap, one standard philosophical response is to simply deny the reality of those properties, calling them illusions (Churchland, 2013). As things stand, this is a straight up conflict between the truth value of subjectivity, Cogito Ergo Sum (Descartes, 1988), and the truth value of public knowledge (Dennett, 1992),

science. That however would still not prove that they were illusions. We are forced to impute from the behavioral facts the existence of entanglement because the apparently simpler explanation, hidden variable is ruled out by Bell's Inequality. The explanatory gap between public abstract ideas and the subjectivity of consciousness, including with it it's sense of free will, is unbridgeable except by the more parsimonious imputation simply that a non-material consciousness exists (Kuhn, 1962). Entanglement and quantum collapse represent scientific phenomena possessed of the minimal essential characteristics that may one day bridge the explanatory gap, an explanation which is both physicalist and dualist.

References

Bergson H. *Time and Free Will: An Essay on the Immediate Data of Consciousness*. New York, NY: The Macmillan Company; 1910.

Berkeley G. *The Works of George Berkeley, Bishop of Cloyne*. Luce AA, Jessop TE, eds. 9 vols. London, UK: Thomas Nelson and Sons; 1948–1957.

Bishop J. Artificial Intelligence is stupid and causal reasoning won't fix it. arXiv. 2008;07371v1.

Bolhuis J. How Could Language Have Evolved? *PLoS Biol*. 2014;12(8):e1001934.

Banissy M. Synesthesia: an introduction. *Front Psychol*. 2014;5:1414.

Churchland P. *Touching A Nerve*. New York, NY: W.W. Norton; 2013.

De Graaf M. Who Wants to Grant Robots Rights? *Front Robot AI*. 2021;8:781985.

Dennett D. *Content and Consciousness*. London, UK: International Library of Philosophy; 1986.

Dennett D. *Consciousness Explained*. Boston, MA: Back Bay Books; 1992.

Descartes R. *The Philosophical Writings of Descartes*. Vols. 1–3. Cottingham J, Stoothoff R, Murdoch D, Kenny A, trans. Cambridge, UK: Cambridge University Press; 1988.

Downing L. George Berkeley. *The Stanford Encyclopedia of Philosophy*. Fall 2021.

Dumontheil I. Development of abstract thinking during childhood and adolescence: the role of rostrolateral prefrontal cortex. *Dev Cogn Neurosci*. 2014;10:57–76.

Dostoevsky F. *Notes From Underground*. Katz MR, trans and ed. New York, NY: W.W. Norton; 1989.

Einstein A. In: *Born-Einstein Letters, 1916–1955: Friendship, Politics and Physics in Uncertain Times*. London, UK: Macmillan Science; 2004:158.

Givental A. On Quantum Measurement. Private communication; 2024.

Key B, Zalucki O, Brown DJ. A First Principles Approach to Subjective Experience. *Front Syst Neurosci*. 2022.

Kuhn T. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press; 1962.

Hallpike C. *The Foundations of Primitive Thought*. Oxford, UK: Clarendon Press; 1979.

Jones C. Large Language Models Pass the Turing Test. https://arxiv.org/html/2503.23674v1?s=09

Laplace PS. *Philosophical Essay on Probabilities*. Dale AI, ed and trans. Corr. 2nd ed. Vol. 13. New York, NY: Springer; 1995.

Larenz R. Private communication; 2024.

Le Poidevin R. The Experience and Perception of Time. *The Stanford Encyclopedia of Philosophy*. Summer 2019.

Levine J. Materialism and qualia: The explanatory gap. *Pac Philos Q*. 1983;64:354–361.

Locke J. *An Essay Concerning Human Understanding*. Nidditch PH, ed. Oxford, UK: Clarendon Press; 1975.

Lycan W. Representational Theories of Consciousness. *The Stanford Encyclopedia of Philosophy*. Winter 2023.

McTaggart J. The Unreality of Time. *Mind*. 1927;17:457–473.

Nagel T. What Is It Like to Be a Bat? *Philosophical Review*. 1974;83(4):435–450.

Pearl J. To Build Truly Intelligent Machines, Teach Them Cause and Effect. *Quanta Magazine*. May 2018.

Pearl J. *The Book of Why*. New York, NY: Basic Books; 2018.

Penrose R. *The Emperor's New Mind: Concerning Computers, Minds and the Laws of Physics*. Oxford, UK: Oxford University Press; 1989.

Penrose R. *Shadows of the Mind: A Search for the Missing Science of Consciousness*. Oxford, UK: Oxford University Press; 1994.

Popper K. *Objective Knowledge: An Evolutionary Approach*. Rev ed. Oxford, UK: Clarendon Press; 1979.

Remler M. *The Mechanisms, Metaphysics and History of Consciousness in the World*. Sarasota, FL: First Edition Design; 2019.

Ryle G. *The Concept of Mind*. Oxford, UK: Oxford University Press; 1948.

Smith G. Newton's Philosophiae Naturalis Principia Mathematica. *The Stanford Encyclopedia of Philosophy*. Fall 2024.

Spinoza B. *The Complete Works*. Shirley S, trans. Indianapolis, IN: Hackett Publishing; 2002.

Suzuki TN. Semantic communication in birds: evidence from field research over the past two decades. *Ecol Res*. 2016;31:307–319.

Starr W. Counterfactuals. *The Stanford Encyclopedia of Philosophy*. Winter 2022.

Turing A. Computing Machinery and Intelligence. *Mind*. 1950;59(236):433–460.

Wiener N. Newtonian and Bergsonian Time. In: *Cybernetics or Control and Communication in the Animal and the Machine*. Cambridge, MA: MIT Press; 2019.

Wigner E. The Unreasonable Effectiveness of Mathematics in the Natural Sciences. *Commun Pure Appl Math*. 1960;1(1).

Wilczek F. The Origin of Mass. Massachusetts Institute of Technology. Available from: https://physics.mit.edu/wpcontent/uploads/2021/01/physicsatmit_03_wilczek_originofmass.pdf