Article

How Self-Reference Builds the World (Part 2)

Cosmin Vișan[*]

Abstract

If you were to build a world from Nothing, how would you do it? By investigating the nature of self-reference, we will show how this can be achieved, how starting from Nothing, Everything can be obtained. Various implications of the definition of self-reference will be investigated, showing how it can account for various aspects of the phenomenology of consciousness, thus showing how starting from only 1 principle, a world of infinite complexity can be obtained. Parallels with set theory will be made along the way.

Part 2 of this two-part article includes: 7. Phenomenological Consequences of Self-Reference; 8. The Evolution of Self-Reference; 9. The Difficulties of the Formless Realm; & 10. Is Idealism Really the Truth?

Keywords: World, self-reference, consciousness, qualia, formless, form, set theory.

7. Phenomenological Consequences of Self-Reference

I hope that by now the reader got a feeling of how self-reference constructs consciousness through the process of looking-back-at-itself. But there are some points that need to be made explicit about the consequences of this process for phenomenology. It has to do with the process of including in the present object all the previous objects. This gives rise to some interesting phenomenological manifestations that are good to be pointed out in order to see how the definition of self-reference happens everywhere in consciousness. Because it becomes difficult to keep analyzing objects step-by-step, I will take in this section directly higher-level qualia and show that they also follow the same properties of self-reference that all qualia follow. Let's investigate the visual domain and see the example in Figure 5.

As self-reference keeps bringing into existence meanings relative to previously existing meanings and contexts, eventually it gets to generate the levels that we see in Figure 5. At the base of the visual domain, we see the black-and-white qualia, which has the meaning of "being visual", and this by necessity looks like black and white. This is where the visual domain starts. Then as self-reference continues to look-back-at-itself and to define itself relative to previous manifestations of itself, it continues to give birth to various other levels, like shades-of-gray which include the black-and-white object, then colors which include the shades-of-gray object, then on top of colors it creates shapes, then on top of those shapes it adds various meanings and

* Correspondence: Cosmin Visan, Independent Researcher, Master in Physics at The University of Manchester.
  Email: visancosmin17@yahoo.com

thus brings into existence the qualia of visual objects, and eventually it goes to create the full visual scene which includes various visual objects.
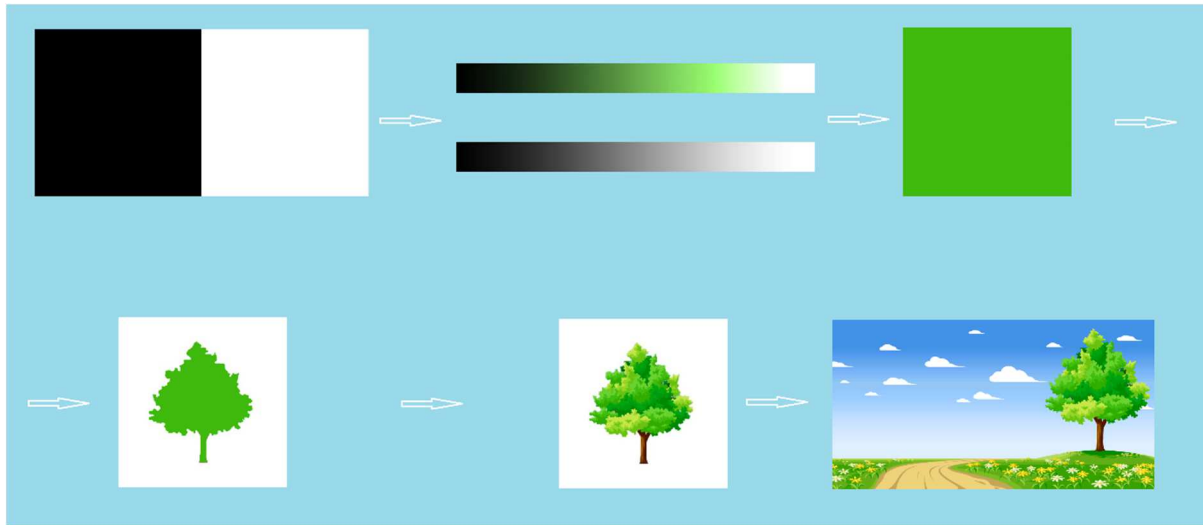
Figure 1. Levels of Self-Reference in the visual domain

Now there are various considerations that this more complex and more concrete example exposes. Firstly, it just goes to confirm the functioning of self-reference that we discussed previously. While the initial levels of Self, Vividness, Diversity, Memory and Time might seem obscure, especially since they cannot be easily experienced directly, but only exemplified in higher-level qualia, now we have a clear picture of what is going on. Colors, shapes, objects, are not anymore obscure, but are the most trivial experiences that we have in our human consciousness. Another interesting aspect is that now we have access directly to the black-and-white level both in itself and as experienced in shades-of-gray, and similarly for shades-of-gray level both in itself and as experienced as part of colors.

In principle, our consciousness could have turned out to not contain as possible experiences the black-and-white and the shades-of-gray objects/qualia; our consciousness could have turned out to only allow us to experience colors directly. But the fact that a certain hue would have varied from dark to bright, would have suggested to us that maybe it is not because the hue itself changes brightness, but that the hue remains the same and what changes is some mysterious object inside the hue that we might have postulated to be called "shades-of-gray". In the same way that in our current human consciousness we cannot experience directly the objects Vividness or Diversity, but we can deduce their existence from the manifestation of everyday qualia, as shown in Figures 2 and 4, similarly if we wouldn't have experienced directly the object "shades-of-gray" we could have deduced its existence from the way the object Color varies its brightness.

Another important manifestation of self-reference is exposed by the current example. Previously, we might have got the simplistic view that what the main engine of self-reference is, is its construction of levels based on von Neumann inclusion and transcendence of "I am"s. And then the fact that we dropped the "I am"s notation and we just employed the higher-levels objects

themselves as notation, like writing "Memory = "I am <Diversity & Vividness & The Self>"" seemed just a more convenient notation and nothing more, we see now that there is more than this.

As we noted previously, we not only replaced the "I am"'s notation by the objects notation out of pure convenience, but we did this to express the fact that the actual process that is taking place is the creation of meaning out of previously existing meanings and contexts. But at that point, because the process was linear, it seemed to make no difference. But now we see that the process is not linear anymore. For example, from the shades-of-gray, not only 1 color is created, but many different colors can appear. Similarly, from the shape (first image on the second row), not only a tree could have emerged, but also a leaf. Another well-known example is the duck-rabbit case, where from the same shape, either a duck or a rabbit can emerge.

So now we properly understand why the main engine of self-reference is not the mindless concatenation of "I am"'s strings, but is the creation of meanings relative to contexts. Once self-reference gets past its first few levels, past "I am" and "I am "I am"", the "I am"'s themselves stop to become relevant, and the new meanings that are brought into existence are the ones that start to take central stage, and then based on them the further evolution of self-reference takes place. That's why I used throughout the paper the expression "similar to Von Neumann" and not "identical" or "equivalent", because the only similarity is the inclusion in the current level of the previous levels, but from the current level multiple levels can then be emerged. Each one of them is similar to Von Neumann, but they together don't reflect the Von Neumann construction anymore.

If we would want to maintain a as close as possible identity, that would be something like, after numbers 0 and 1, there would be numbers 2 and 2', and then after number 2 would be numbers 3 and 3' and 3" and 3''' and so on. I'm not a mathematician, so I don't know if such a construction makes sense in mathematics, but as far as consciousness is concerned, this is how self-reference generates the qualia of consciousness. Actually, if such a manifestation doesn't exist in mathematics, then a mathematician reading this paper might have at this point some interesting inspirations that he can use to develop new theories in mathematics and why not maybe even bring a full mathematical theory of consciousness if that is even possible.

Things become even more complicated (and beautiful). As we all know, an object is usually made out of multiple colors. Trying to make a connection to mathematics, this would be similar to something like after 4, 4' and 4", a single 5 would come. So instead of an axis of numbers, we would have something like a tree, but not just any tree, but a tree in which the branches after they have split from the larger trunk upstream, they unify back downstream. So, the potential mathematics of consciousness would either be more complicated than might appear at the first sight, or because of such weird parallel numbers appearing, it might stumble upon some Godel-like inconsistencies and no mathematics of consciousness would be possible. And it will get even more complicated as we will develop the theory in the paper.

For example, we would see cases in which after 4, 4' and 4", 5 would come, and at the same time after 4 and 4', 5' would come. So, there would be something like superpositions of

superpositions and so on. This, of course, would ultimately be because of the contradictory properties of self-reference of being no-thing and every-thing both at the same time and other equivalent contradictions. But such contradictions are no problem, because they are in the formless realm, the form realm not only not being affected by them, but being actually created by them. We can see a representation of such cases in Figure 6.
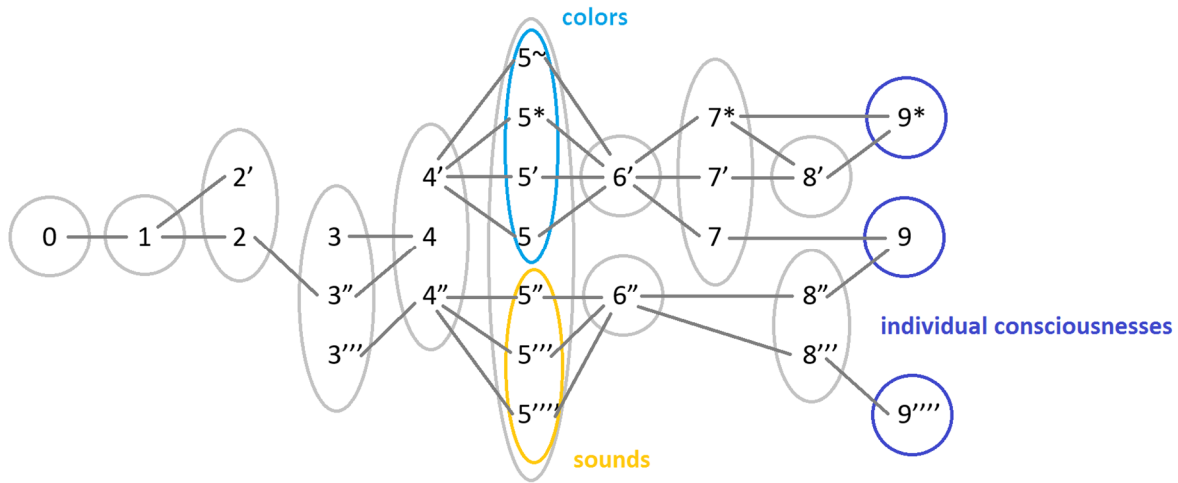
Figure 2. The intricacies of Self-Reference's levels

What we see in Figure 6 is a representation of various cases that self-reference can give birth to. They don't necessarily represent real levels, but various intricacies that can be generated by self-reference. For example, even though 4' and 4" might be on the same level, they can in turn generate levels with quite different qualities, like 4' generating 5~, 5*, 5' and 5 which can be colors qualia, while 4" generating 5", 5''' and 5'''' which can be sounds qualia. Once the colors qualia and sounds qualia are brought into existence, they become independent and they give birth to utterly different qualia domains, like the visual, that might include levels 6', 7*, 7', 7, 8' which can be shapes, visual objects, etc., and the auditory, that might include 6", 8", 8''' which might be language, music, etc.

Another interesting case might happen on the last level, that represents the individual consciousnesses. For example, you can have case 9* which emerges on top of 7* and 8' and which might represent a consciousness that only sees. Also, you can have case 9'''' which emerges on top of 8''' which might represent a consciousness that only hears. But interestingly, you can have case 9 that emerges on top of the visual quale 7 and the auditory quale 8" and thus it represents a consciousness that both sees and hears. This is what is called in neuroscience the binding problem or in panpsychism the combination problem: how do different qualia unify into 1 single consciousness? And we can see that this can happen by self-reference looking-back at 2 different manifestations of itself and bringing them together under a singular looking-back and thus creating a consciousness that is able to experience multiple qualia domains.

Besides the unifications that happen inside a single individual consciousness, these workings of self-reference are also responsible for telepathies: two independent consciousness can be unified temporarily under one single consciousness, have one single experience, and then split back and each of them remembers that one single experience. Thus, telepathy is not some signal transmission-reception, but is a unification of individual consciousnesses into a single consciousness that has a certain experience, and that experience is remembered by the individual consciousnesses once they split back. Also, because each individual consciousness is made up of various distinct levels, the telepathy/unification can happen between sublevels. Thus, me dreaming being on a trip to the mountains and my partner dreaming being on a trip to the seaside, is still a telepathy, is a unification of the level "trip". But since each one of us is made up of different sub-levels, my level "trip" is in turn unified with level "mountains", and her level "trip" is unified with level "seaside", the telepathy will only be partial. Thus, a not-perfect telepathy is not a reason for rejecting the phenomenon, but is actually a door towards the intricacies of self-reference.

And maybe the greatest revealing of Figure 6 is that we are all connected. Even though it might appear on the surface that we are individual consciousnesses, appearance highly accentuated by the fact that we appear to have separated biological bodies, deep down in our consciousnesses we are connected. Some might be connected by multiple levels, some might be connected by fewer levels, and similarly members of the same species might be interconnected more than between individuals from other species, but ultimately there is a connection between all the consciousnesses in the world, even if between some consciousnesses that connection might be only at the level of the Self. Some might experience colors, other might experience sonars, other might experience infrared, but all of us, humans, animals, plants, etc. experience one and the same Self. At the base of our consciousness, we are all one and the same Self. And ultimately, we are all one and the same self-reference. So even though death might destroy parts of self-reference, self-reference itself is indestructible.

More so, given that we are all one and the same Self, we all experience that Self at all times. So even from the point of view of an individual consciousness, the Self of each individual consciousness is indestructible. So not only that self-reference endures forever, but given that the Self is eternal and is part of all individual consciousnesses, in a way each individual consciousness endures forever as well. Sure, it might lose parts of itself at the moment of death, but it cannot lose the Self. So, whatever might happen to us at the moment of death, we will still continue to have experiences, at least the experience of the Self, which then will emerge new levels and our individual consciousness will evolve again and we will start a new life somewhere else in the webs of self-reference.

Actually, depending on the exact configurations of the webs of self-reference, we might not even lose much at the moment of death. Given that our sub-levels are also shared by other individual consciousnesses, and thus are sustained in existence by those other individual consciousnesses, our death might only destroy some surface levels of our individuality, and we might end up into another life pretty much intact, thus explaining cases of reincarnation in which those new individual consciousnesses retain memories from their past individualities. All these

Scientific GOD Journal |December 2023 | Volume 14 | Issue 4 | pp. 263-281
Visan, C., *How Self-Reference Builds the World (Part 2)*

268

considerations can be put into 1 short sentence: I am God. Self-reference is basically God. And each one of us are various instantiations of God. Through individual consciousnesses God knows about itself. And the destruction of individual instantiations is not capable of destroying God. God is eternal, being forever maintained into existence by the eternal Self.

Let's take another example to illustrate the inclusion of the lower levels into the higher levels. As we saw in the previous sections, each object from the previous levels is included in the newly emerged level. As a consequence, it goes on to share its quality towards the full quality of the newly emerged level. And it continues to do so for all the levels that emerge on top of it. As we also mentioned previously, we feel alive either if we see something, or we hear something, or we experience some emotion, and this is because in all these objects, the object "I am" is included. Also, the fact that we see motion happening in the visual domain or we listen to music or we hear someone talking, is all because in these objects, the object Time is included. So once an object is included in the immediately above level it doesn't stop its existence, but continues to manifest itself all throughout the higher levels. Let's take another series of levels and specify exactly the manifestations of the levels involved in them. We will take the example of language.

Shapes: quality of "visual objects": entities with spatially defined boundaries.

Letters: inherits the quality of the Shapes, thus becoming themselves visual objects, and emerges on top of it its own quality of "unities of language".

Words: inherits the quality of the Shapes, being themselves visual objects, inherits the quality of the Letters, being themselves unities of language (just more complex than letters), and emerges on top of them all its own quality of "carriers of linguistic meaning".

Sentences: inherits the quality of the Shapes, being themselves visual objects, inherits the quality of the Letters, being themselves unities of languages (just more complex than both letters and words), inherits the quality of the Words, being themselves carriers of linguistic meaning, and emerges on top of them all its own quality of "carriers of ideas".

Thus, we start to understand that one technique of how to probe deep into consciousness, is to start with everyday experiences and try to find as many qualities in them as possible, and then try to see how we might arrange them in a holarchy of levels. Actually, the way I got to the ideas that I'm presenting in this paper is precisely this way. All that I had was normal everyday consciousness, and I started to do introspection for many years. And only after a great deal of introspection, did I get to the idea of self-reference. So, in practice, the order of gaining knowledge was opposite from the order that I chose to present in this paper. But in order to make an exposition starting from first principles, I decided to present consciousness starting from self-reference, and only then obtain our every-day familiar consciousness. By doing this, we can see how starting from first principles, we beautifully obtain everyday consciousness without having to resort to additional ontological entities. All that we have is self-reference and this is everything that we need to explain consciousness. No need to postulate some additional Self, some additional Time, some additional qualia; all of them are self-reference. And all of them follow the definition of self-reference, all of them manifesting emergent structuring of being a

level of consciousness that includes in itself lower levels and is itself further included in higher levels.

Moving further with the phenomenology, we notice a further aspect, that will push the manifestation of self-reference even further from the von Neumann construction of numbers. We saw how the ramification of levels, like multiple colors emerging on top of shades-of-gray being similar with numbers 3, 3', 3", etc. following number 2, is a departure from the construction of numbers in set theory. But now we face another departure, which either will make the potential mathematical formalization of the theory even richer or even "more" impossible. And that is the top-down influence in levels. Not only that the top meaning is obtained as relative to the context of the previous meanings from the lower levels, but the top meaning itself is able to modify the lower levels that it contains. This would be similar to something like, once number 5 is obtained after numbers 0,1,2',3 and 4", number 5 decides to change the order and make itself come after 0,1',2,3" and 4'''. Let's look at Figure 7 and understand this manifestation.
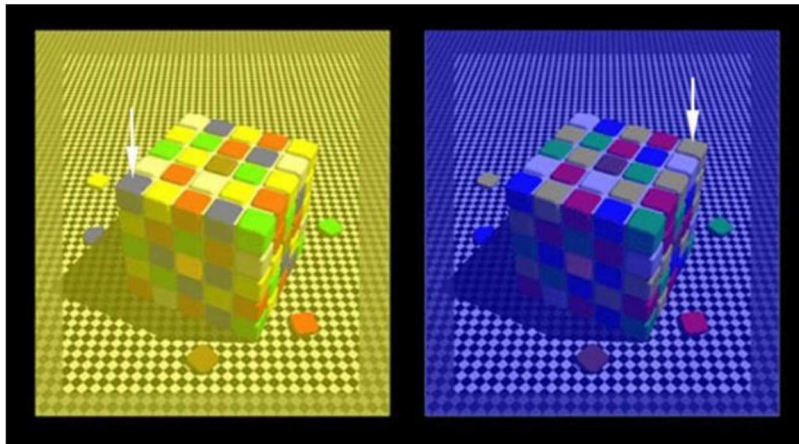
Figure 3. Top-down influence in levels

The two squares indicated by arrows are gray in isolation, but when put in the context of the cube they become blue and yellow. We can consider in this case for gray to be some number, say 2, then on top of 2, 3 and 3' can emerge which would correspond to colors blue and yellow, and then on top of numbers 3 and 3', number 4 would follow which would correspond to the image of the cube. Notice that in both cases is the same cube, there are not two different cubes. And the cube selects on the left the blue square and on the right the yellow square. So once number 4 appears, it selects on the left to follow after 2 and 3, and on the right to follow after 2 and 3'. But this selection happens only after the cube comes into existence, not before. So, the order of the sub-levels is changed once a higher level comes into existence. We are dealing with a top-down influence in levels.

This again, is not something that happens in the von Neumann construction of numbers. So, if a mathematical theory of consciousness is possible, it needs to be an extension of set theory. Now, why do I say there is the same cube and not two? This is similar to how the brightness of a color varies in Figure 5. Though the color is the one that seems to vary in brightness, the actual variation comes from the sub-level of shades-of-gray. The color is the same, like for example

color green. Is just green. The fact that it varies from light green to dark green is because of the level of shades-of-gray that varies inside the holarchy of the color. If you want, is something like platonic ideas. The idea of "green" is one. And then this idea gets into various combinations with other ideas, and the full context of those combinations is what determines the final experience. Similarly, the platonic idea of "cube" is one. And then it goes into various combinations with other ideas, like the ideas of "colors", and the full context of these combinations is what determines the final experience.

It is interesting to note here that the platonic ideas cannot exist on their own. You cannot have some abstract "free-floating" cube. In order to manifest itself it needs to appear on top of some pre-established structure. And this is ultimately because of self-reference itself. Self-reference itself in order to exist must do so based on some pre-established structure. And on that primordial level, the only pre-established structure is itself. So "I am" comes into existence on top of itself. The form "I am" sustains its existence on top of the formless "I am". And then this necessity for formless/abstract/platonic ideas to come into existence on top of pre-established structure is maintained as a necessary principle of existence, such that in order to experience the abstract ideas of "tables" and "chairs" and "rocks" and "trees", all these platonic ideas must be brought into existence on top of colors or on top of sounds or whatever other forms of manifestation of consciousness might be experienceable by other beings.

Universals manifest themselves as particulars by emerging on previously existing structures. That's why you can see a chair that is brown or a chair that is white or a chair that has 4 legs or a chair that has 3 legs. Its particular instantiation is dependent on what it previously exists in consciousness. But "chairness" is the same in all cases. "Chairness" is a formless universal, that can only manifest on the form realm if it emerges on top of particular combinations of some other pre-existing universals already instantiated as particulars. "Itselfness" is the only universal that emerges as particular on top of itself. As such, it is unique. While you can have a chair that is brown and a chair that is white, there is only 1 unique "I am". And the beauty of it all is that we are all that 1 unique "I am".

To acquire an even better intuition for top-down influence in levels, an additional example involving Time is suitable to be presented at this point. We saw the levels that Time includes, we also mentioned how Time itself is included in higher levels, like watching objects moving in the visual domain or listening to music in the auditory domain, but a residue of false physical intuition might reside in the materialistically inclined reader that maybe the motion in the visual domain is actually because objects really move outside consciousness. To dispel this last false intuition, let's look at Figure 8.
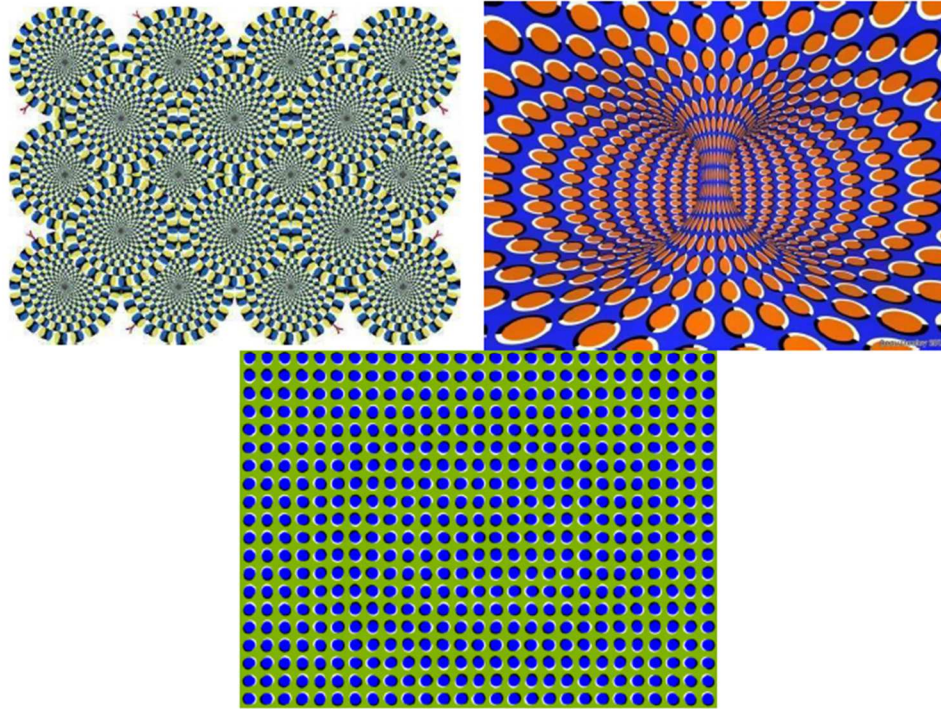
Figure 4. The level of Time manifesting in the visual domain

We notice something peculiar about these so called "motion illusions". The reason why they work is because of the alternating black-and-white qualia in each of the three images. But how could color qualia have anything to do with time? Aren't they separate ontological categories? And here is the beauty of the theory developed throughout this paper: they are not separate ontological categories. They are all qualia. And as qualia, they are all structured on the emergent holarchy of meanings that arises as self-reference looks-back-at-itself. Both time and colors are certain meanings/forms that self-reference identifies itself with. And as such, they all follow the same rules of self-reference: inclusion and transcendence of levels together with the later top-down influence in levels.

Therefore, there is no problem for colors to influence time. Time is a certain level of self-reference that lies below the level of colors. Therefore, it can receive top-down influence in levels from the level of colors. In this particular case, if the black and white are disposed in a certain manner, this will create a top-down influence upon the level of time and thus create motion in those particular images. I think these are the clearest examples that time indeed is nothing more than just another quale in consciousness, that also follows the general rules of self-reference structuring consciousness on a holarchy of levels.

8. The Evolution of Self-Reference

But why that particular arrangement of black and white and not others? Because clearly if you randomly put black and white on a piece of paper you don't always get motion. The reason is that this has to do with the evolutionary contingent appearance of meanings in the history of self-reference. To understand what this means, let's look at Figure 9.
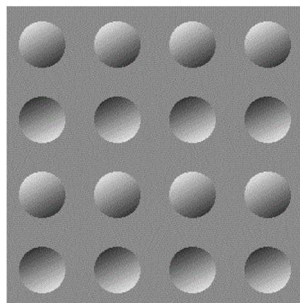
Figure 5. Evolutionary contingency of meanings

The reason why on the first and third rows those particular arrangements of shades-of-gray are experienced as bumps and on the second and fourth rows are experienced as dimples is because of evolutionary contingency: the Sun just happened to shine from above, and any wall of rocks from the environment upon which the Sun shined would look like the first and third rows if it contained bumps and it looked like the second and forth rows if it contained dimples. And this particular emergence of bumps and dimples qualia on top of shades-of-gray qualia passed on across generations and now we can have the same qualia just by looking at a picture on the computer screen. And this is an interesting fact: the history of self-reference is encoded in the qualia that we possess. Even more interesting, the meanings might be buried so deep that we can never uncover them.

For example, in the future, some human descendents might end up living underground for many generations such that the knowledge of the Sun could be completely lost. Yet, they would still experience these qualia from Figure 9. But they would not be able to make sense of them anymore. Similarly, it is to be expected that we have in our qualia catalogue countless qualia of which origins are completely lost to us. Don't forget that the story about Big Bang and evolution of galaxies, is just a story that we tell ourselves now from the limited observations that we can make about the universe. But that is a story told from the third person.

The true history of self-reference took place from the first person. And this didn't involve Big Bangs and galaxies. This involved creatures living in evolutionary contexts, and the meanings that they acquired were meanings that were relevant for their survival and reproduction in those particular contexts. And those contexts might not be relevant today anymore. But we might very well still possess their qualia, unable to understand their origins anymore. For example, our ancestors might have lived in "100-dimensional realities" in which they fought for survival. But those "100-dimensional realities" are completely lost to us, yet we might still have relics from those times in our qualia, but forever beyond our grasps to understand their origins.

We thus see that the complexities of self-reference explode beyond our wildest imaginations. Initially we started innocently, just concatenating "I am"s strings, but now we ended up with whole realities that self-reference explores within itself, with whole worlds that are nothing but self-reference interacting with itself. Thus, though initially the evolution of self-reference seemed to be linear and deterministic, we now see that it is highly non-linear and based on contingent contexts in which individual consciousnesses compete for survival through their employment of their free will. The individual consciousnesses are of course nothing more than self-reference itself taking on various roles.

And the breaking of symmetry happened because of free will. While initially self-reference evolved linearly and deterministically, the fact that it kept creating more and more objects within itself, those objects could then be included and transcended in different individual consciousnesses. And at that point, a new game starts in which each individual consciousness has its own goals and pursue them at the detriment or benefit of others. And that is by employing free will. At that point, the next levels/qualia that self-reference gets to experience will be born out of whatever new contexts the free will of each consciousness puts that consciousness into. And that's where the linearity and determinism of levels end and the era of evolutionary contingency starts and the science that can be done shifts from purely theoretical to empirical, the new levels not being able to be predicted like we did in the previous sections solely by analyzing strings of "I am"s, but they have to be uncovered by investigating the evolutionary history of self-reference.

And this raises one interesting and deep problem related to evolution: Why do we eat? Our current materialistic science tells us a story about needing energy and needing various organic compounds to make the physical body maintain itself. But given that "physical world" doesn't exist, that consciousness is all there is, why do we actually eat? Couldn't self-reference create dream worlds in which it can freely float on the fields and forever lives in a state of nirvana? Maybe it could also create such worlds or maybe it cannot. Maybe it will do that eventually, but it is not there yet. So, what is then stopping it from creating a paradise in which no animal and plant needs to be killed anymore? The answer is very revealing about what is actually going on with this world. Remember that self-reference is all there is. If there is anything that can stop it from achieving nirvana is itself. As self-reference evolves more meanings within itself, at some point it might as well develop multiple meanings at the same time.

For example, let's say that it is still at the level of shades-of-gray and it has some chair "in front of it". It might happen that it can go on and see that chair both red and green. But since as forms, red and green are subjected to the principle of non-contradiction, they cannot exist at the same time in the same consciousness. So, the solution is for 2 different consciousnesses to be brought into existence, one to see the chair red, the other to see the chair green. Since they don't know one about the other, there will be no contradiction, each one of them thinking it is the only one in existence and its perspective on the world is the absolute truth. Both of them will be self-reference itself. But as we saw, even though at the highest level they appear to be individual consciousnesses, deep down in their emergent structure they are connected by shared levels. Given this connection, and given the top-down influence in levels that we previously discussed,

whatever a consciousness does in its own world, will have repercussions deep down in its structure. But since that deep structure is shared with other consciousnesses, whatever a particular consciousnesses changes in its higher level experience, will affect what other consciousnesses experience in their own higher levels worlds. Exactly how the consciousnesses are connected and thus how they ultimately interact is a matter of evolutionary contingency and can only be revealed by empirical science.

And here lies the answer to why do we eat. Even though self-reference might indeed strive for nirvana and for life without food, its strivings towards that state are slowed down by its other instantiations that want to pursue other purposes. In one instantiation self-reference might bring into existence meanings that tend towards nirvana, but in other instantiations it will also bring other meanings that will pull towards other purposes. And through top-down influence in levels, this will disturb the instantiation of self-reference that tends towards nirvana. And this process of disruption of ultimate goals is what eating is. In order to reach nirvana, self-reference must fight against its own instantiations that don't strive towards nirvana. Therefore, it has to eliminate them. It has to purge itself of lower meanings in order to achieve its highest possible meaning.

Note that this process happens through top-down influence in levels, not through biological eating. What biological eating is is a mechanism developed by self-reference in order to enhance the elimination of lower meanings of itself. Fundamentally, eating happens through top-down influence in levels. But since that process in its pure form was cumbersome for self-reference to be used to eliminate its lower meanings, it created highly sophisticated mechanisms that took the shapes of plants and animals, and through them self-reference maximized the efficiency of eliminating lower meanings of itself and even more than this, it created a competing environment in which the strive for survival encouraged the emergence of higher meanings through which eventually nirvana could be reached.

Note that this process might be blind, maybe the end goal is not nirvana, maybe there is no end goal. But the explanation for eating remains the same: eating is instantiations of self-reference competing with one another for dominance. This competition happens through top-down influence in levels directly inside self-reference. Biological eating is a highly sophisticated second-order phenomena that is an appearance of the true eating which happens not by physically killing and physically eating, but by top-down influence in levels directly in self-reference.

A disturbing moral problem arises here: As the highest instantiations of self-reference that we are, should we kill the lower instantiations in order to reach nirvana? Fortunately, given the unimaginable complexities of the entanglements between the instantiations of self-reference involved, most likely than not just killing other forms of self-reference will be a self-defeating purpose. Just killing all the animals and plants on the planet will just leave us humans with nothing to eat and then we in turn will die as well, and self-reference will have to make another "Big Bang" to start from scratch. Most likely, the process is so fine-tuned, that the only possibility is to let go of any plan of domination and let "nature" do its course. If we still insist that nevertheless we want to accelerate the road to nirvana, the way forward will not be through

blind and mindless killing, but through highly sophisticated understanding of the interactions that happen inside self-reference.

Only disentangling the knots through which individual consciousnesses are connected, we might aspire at eliminating eating, and we might do this even without killing any being. By such understanding of the entanglements of consciousness, we might set all beings free from the sin of eating and bring true paradise into the world, in which consciousnesses, both less evolved like plants and animals, and highly evolved as humans, can live in harmony with no killing for eating ever taking place again. But such a science might take millions and billions of years to develop, and in a self-fulfilling prophecy sense, this might actually be the very road to nirvana.

9. The Difficulties of the Formless Realm

In discussing the top-down influence in levels we saw that there is a pretty intricate way in which levels interact, this making the creation of a formal theory of consciousness challenging. But the reader might still have the impression that the workings of self-reference discussed so far can be formalized into a complete theory. It might appear that all you need is some mathematical theory that talks about levels and the connections between levels, and this appears to be doable. But because this is fundamentally impossible, I will give in this section some more intricate examples to strengthen the intuition of the reader regarding what formless entails and to convince him that no complete formal theory of consciousness is possible.

Before getting to the example, one more aspect of self-reference is worth emphasizing. As we saw, self-reference is an interplay between formless and form: form arises because of the formless looking-back-at-itself. Since a synonymous for "form" is "thing", a synonymous for "formless" is "no-thing", which we can more aesthetically call it Nothing. So self-reference starts from Nothing. But as we saw, by the process of looking-back-at-itself, Nothing is able to bring forms/things into existence. Actually, we saw that all the consciousnesses in the world are self-reference itself. All the qualia that ever existed and will ever exist are brought into existence by self-reference looking-back-at-itself; are self-reference. As such, not only that Nothing brings things into existence, but it brings all the things, or in short, every-thing, which we can more aesthetically call it Everything.

As such, Nothing brings Everything into existence. And even more profound than this, since the things/forms that self-reference brings into existence are not separate from self-reference, but are self-reference itself, Nothing not only brings Everything into existence, but **Nothing = Everything**! It is a very satisfying conclusion, is what we desired all along. We wanted 1 single principle to explain everything. And after reading the analysis done throughout this paper, we finally get to understand that indeed starting from the 1 single principle "let self-reference be the entity with the property of looking-back-at-itself" we indeed are able to explain the entire world. We couldn't have asked for a more powerful principle to explain the world. Does it sound like a nonsense contradiction? Sure, if you read it out-of-context on some fortune teller cookie, that's exactly how it sounds like. But here we arrived at this conclusion after countless pages of serious

analysis, both theoretical and phenomenological, such that we are in the position to actually appreciate this profound truth of reality and to understand that without such ontological contradictions nothing could have ever existed.

We thus get to appreciate that no formal theory is even possible because of the formless part of self-reference which by definition is not formal. But it can be hoped that at least some partial formalization might be achieved, for certain limited cases, along the lines of "if we put consciousness in this context, we expect this to happen". For such undertakings to be successful, a rich enough awareness of phenomenology is desired. Therefore, lets provide yet one more phenomenology that I hope will help towards the creation of partial formalizations or at least of testable experiments. We will see that the Nothing = Everything contradiction doesn't only apply to the totality of qualia that self-reference brings into existence, but since Nothing = Everything is a property of self-reference itself, it actually manifests at all levels of self-reference as well.

Given the formless part, self-reference can maintain contradictions within itself that are alright to exist as long as they are not formalized where the principle of non-contradiction applies. Before a certain meaning is brought into existence to be experienced, whatever happens "behind the scenes" can very well be contradictory. This is not necessarily something unheard of. A similar thing happens in quantum mechanics where before the moment of measurement, you have superposition, which is basically contradictory entities "existing" at the same time. But even in quantum mechanics, even though the contradiction might appear to have been successfully formalized, it is not really the case, because the moment of measurement must by necessity be formless. Quantum mechanics is exactly an example of a partial formal theory. You might be able to come up with some formalization for some part of the phenomenon, but in the end, there must be a place in the phenomenon where the formalization must make way for the formless part to do its doings. And in quantum mechanics that is the moment of measurement. In a theory of consciousness, such places where the formless will have to be let to its own doings might appear in other places. Where will such places be remains to be seen by whoever might undertake the challenge of developing the partial formal theory of consciousness. In order to facilitate the development of such a theory, let's look at another intricate way in which levels are able to interact because of the formless part of self-reference.

I will talk about a certain type of errors that we make in everyday life, that in general are simply ignored by people, but that actually hide deep mysteries of the workings of the formless realm. That type of error is when for example we write "light tie" instead of "tight lie". It might seem like an innocent occurring, probably just some neurons misfiring. But if we are to take it seriously, we will see the depths of the formless realm. Therefore, let's take it seriously. How could we explain the swapping of the first letters? In order to explain it, we need to be aware of what we are dealing with. Namely, we are dealing with meanings/qualia. As such, any explanation would involve the workings of meanings. And as we saw throughout this paper, meaning is created by self-reference looking-back-at-itself and constructing a layered holarchy of levels. Therefore, we might regard the present meanings as being on the level of "group of words" that is a level that is on top of the level of "words" that itself is a level that is on top of the level of "letters".

More precisely, on the level of letters we have the meanings/forms: "t","i","g","h","t","l","i","e". These meanings are then included and transcended on the level of words which is composed of the meanings "tight" and "lie", which are themselves included and transcended in the highest level of the group of words "tight lie". But sometimes, this highest level can be scrambled and be transformed into "light tie". How is this process done? Is it merely a swapping of meanings from the lower level of "letters"? Is not that trivial. Is actually more subtle, and it has to do with the contradictory nature of self-reference.

In order for the swapping to occur on the level of "letters", those letters need to "be aware" of the higher levels of which they are part. Ultimately, they are part of a group of two words. As such, whatever swapping might take place, it will take place based on the restrictions of the highest level. Thus, even though on the intermediary level of "words", the letters "t","i","g","h","t" went into the word "tight" and the letters "l","i","e" went into word "lie", it is still the case that because of the highest level "tight lie", their position in the structure of the intermediary level is not set in stone. Since only the highest level of self-reference is a form, whatever structure goes into making that form, is itself into the formless realm.

As such, it lives in a state of contradictory superposition. So even though superficially it might appear to us that when we write "tight lie", the letters are set in stone each in their corresponding words, in fact what is happening is that they on themselves are in the formless realm. And I repeat, this is because only the highest level is the level that is the form. Everything else is formless. As such, before the final look-back that is about to create the form happens, all the levels are in superposition in the formless realm, both inter-level and intra-level, and as such they can still be scrambled before the form is experienced, and thus, sometimes it happens that instead of experiencing "tight lie" we end up experiencing "light tie".

The reason why such superpositions exist both inter-level and intra-level is because of the Nothing = Everything contradiction. Before the final look-back that creates the actual form to be experienced, Everything regarding that form exist in the formless state, so in the state of Nothing. So the way the Nothing = Everything property of self-reference applies to this case is for Everything to refer to all the levels that go into the composition of the form, such as group of words, words, letters, shapes, and whatever other elements might appear in more complex cases, and for Nothing to refer to the fact that all these elements exist in a state of formless superposition. This has to happen by necessity, because by necessity self-reference has the property of Nothing = Everything. And this property has to apply at all levels of manifestation of self-reference, not only at the utmost level of the entire world. And the reason is, as we saw all the way in Figure 1, that all levels of self-reference are self-reference itself. Self-reference is not only the "I am" or the entire world of all consciousnesses, but every step along the way is also self-reference itself. As such, each step along the way must have the very same properties of self-reference. Therefore, in any attempt at partial formal theories of consciousness, we have to be careful how we apply the property of Nothing = Everything by properly identifying what the Everything part refers to and what the Nothing part refers to.

Now, exactly how the error happens and what are the full constraints, is a matter of both empirical science for guidance through data collection and then of partial formal theories of

consciousness of the future. What I'm providing here is awareness for the reader regarding the richness of phenomenology that is created by the fact of self-reference having a formless part that is subject to contradictions. So, an innocent swapping of some letters is not just "neurons misfiring", but is the full force of the formless realm of self-reference in action. By investigating such errors, we can begin to develop partial formal theories of consciousness in which to try to account as much as possible for whatever mysteries are happening in the formless realm. Note that this error is not chaotic, but it has structure to it. Contradictions in the formless realm doesn't mean chaos.

They are not chaotic contradictions, but are contradictions leading to structures in the experienced forms. But at the same time, even though this example given here might appear trivial at a first sight, a more mathematically inclined reader might soon discover some unsurmountable difficulties if he tries to come up with some formalization for it. And he might start to get the first glimpses of awe at the magic of the formless realm. Any attempt at formalizing this error example will have first and foremost to take into account the entire theory of self-reference presented in this paper. Only then some progress can be made. Just on its own it will be soon discovered by any mathematician that it is impossible to be captured under any formalism. Is like a fata morgana that it looks like is just within reach, but the closer you get the farther it recedes, until you realize that indeed there is something deep going on here. And what that is, is the formless realm with its contradictory nature.

Some other example of such error, that I will not go into details here, is writing "write" instead of "right". Here we see that the level of letters doesn't even contain the same letters for the two words, yet the error still happens. Needless to say, the formless realm is once again even more mysterious than we just thought just one paragraph above. And these cases will go on forever. Once we thought we got a formal theory of consciousness, one more example is brought forward by someone, that destroys that entire formalization. And the process goes on forever, with more and more intricate cases being continuously discovered. This is what formless mean.

Not being a form, there is nothing to limit it, it can go on forever in its baffling weirdness. And because of its infinite extent, it will also produce infinite forms when it looks-back-at-itself. Forms in themselves will not have contradictions, red is not green and a color is not a sound. But they and their tremendous diversity are possible because of the limitless nature of the formless realm. And I emphasize, its limitless nature is not for some mystic reason, but for the trivial definition that formless is not form. And the reason it doesn't forever remain just in a potential state, is because of self-reference, that by logical necessity looks-back-at-itself and manifests into the realm of form whatever it might happen in the realm of formless.

10. Is Idealism Really the Truth?

Is this really the truth? Self-reference is all there is? How about electrons and space-time and cells and brains? They don't exist? Rather than anything else, this is a question of what existence is. Before wondering if electrons exist, we first need to make clear what "to exist" mean. The clearest definition would be, if we put X into a proposition of form "X exists", we must be able to know all that X is, otherwise how would we even be able to talk about an X? For example, even though we might say that electron exists because we can say that the electron is an elementary particle with electrical charge -1 and spin ½, this cannot fully elucidate what an electron is, since it can still have properties not discovered yet and in principle it might never be possible to know if all properties have been discovered at some point. Thus, "electron" remains a vague notion. Being vague, it cannot be said to exist, because what is that thing said to exist if its character is not elucidated?

On the other hand, this problem doesn't exist for qualia. Red is red. And even though we might not know the entire emergent structure of red, by the very act of experiencing it, we grasp it in its totality. Therefore, red indeed exists. And this can only ever be said about qualia. Only qualia, by the very fact that they are direct experiences, we can know them fully. Therefore, when we say "quale X exists", we fully have the X to which to refer, therefore indeed quale X exists. Therefore, by the very definition of existence, consciousness is all that exists. Therefore, idealism really is the truth. Is really that simple. But what is more subtle is the nature of self-reference. We also talked at the beginning of the paper, but is good to clarify some more, especially now that the reader has a better intuition of what self-reference actually entails.

Consciousness indeed exists and is indeed all that exists. What about self-reference? Does it also exist? Strictly speaking, self-reference doesn't exist. But for totally different reasons than why electron doesn't exist. Electron doesn't exist because we cannot make a full concept of it of which to be sure that is complete and no further surprising properties might arise by empirical science. On the other hand, the reason self-reference doesn't exist is because it is not a form. Actually, the very language that we employ to speak about self-reference is faulty, because language is form, and employing form to talk about formless is by its very nature unsuitable. **Correctly, self-reference cannot be spoken of.**

But even saying "self-reference cannot be spoken of" is an utterance about it, so not even such a sentence can be uttered. Even naming it is faulty. Not even saying "self-reference" is correct. Is a very peculiar states of affairs. On the one hand, we cannot speak about it, on the other hand, this "entity" (wrong again, because not being spokeable-about, we cannot call it "entity" either) is responsible for bringing consciousness into existence. Some might wonder, if we cannot speak about it, why are we sure that it is the one that brings consciousness into existence. The reason we can do this is because we observe the phenomenology of qualia (like inclusion and transcendence of levels) and conclude that this is possible only if some entity that we call "self-reference" must "exist". If this peculiar state of affairs doesn't produce awe in the reader, then the reader didn't actually grasp what is being talked about in this paper.

Is this the final theory of reality? Shouldn't we discover it in the year 3000 or year 1 million? Is it really that simple that we can have it today? While clearly the details will take probably forever to uncover, I think the fundamentals are the ones presented in this paper. Why would this be the case? Setting aside the ideas presented in this paper, any other attempt will most likely postulate form entities in the same spirit as "electrons" and "brains". But as we saw, this cannot possibly work, for empirical reasons, i.e. not being ever possible to know all the properties of such empirical entities. And even if a theory might come up with some sophisticated form entities defined apriori as to ensure they are fully defined, this will not satisfy either, because whatever entity someone might invent, that entity is ultimately a thought in consciousnesses.

Say someone might say that "ultimatron" is the ultimate entity able to explain everything. But that "ultimatron" as long as it will be part of some formal system, even leaving aside any Godel-like incompleteness of that system, "ultimatron" will not be able to explain the "I" that is thinking it. It will forever miss the crucial observer that is inventing that very theory. As such, it cannot have the necessary explanatory powers. What is different about self-reference is that it makes room for the observer. By giving up on complete formalism, it leaves the "I" unchained and to its own free will. "Self-reference" is not a theory (in the spirit of "ultimatron") and can never be. "Self-reference" is at most a pointer to some deep aspects of reality that are beyond formalization. Whatever happens in that formless realm might as well be magic, self-reference respects it and allows it to be whatever it desires. The only forms that the theory of self-reference considers are the ones that result from the looking-back-at-itself process.

About those forms, the theory of self-reference can speak about. It can for example tell how they grow in complexity by including and transcending the previous levels. But the "I" that is experiencing those forms is left to its own and is respected for what its mystery forever is. Also, because the forms themselves are a consequence of the ontological mystery of the formless realm/of the observer, they themselves inherit the mystery and don't allow themselves to be known except by putting oneself in the shoes of the observer. Even though the formal part of the theory can be described in terms of inclusion/transcendence, top-down influence in levels, etc., the actual content of the experiences can only be experienced first-hand. If I want to know how it is to ride a roller-coaster, the only way to do this is to actually ride a roller-coaster.

Similarly, from the point of view of my consciousness, all the other consciousnesses are formless. Only 1 form can ever exist at one time, and that is Me. I am all that exists. My eternal present moment right now is all that exists, is all that has form. Everything else is formless. The people that I see around me daily are formless for me. But equally, from their point of view, I am formless for them. Needless to say, this is because of the contradictions of the formless realm, including the Nothing = Everything property. I guess the reader got used by now to the infinite awe-inspiring manifestations of the contradictions of the formless realm. This paper only touched upon a few of them, but in the future of science they will pop-up everywhere.

Therefore, any theory that doesn't take into account these considerations is destined to fail. As such, even if year 3000 or year 1 million comes, if the theories attempted in those years will be strictly formal theories, they will be as incorrect as any such theory that is attempted today. Only by letting the observer be the formless entity that it is, that any theory of reality can have any

chance of being true. And this is exactly what the theory of self-reference developed throughout this paper is doing. It gives credit where credit is due and respects the observer for the formless entity that it is. Knowledge/understanding can only go as far as forms go, since they themselves are form. Beyond that, it is the realm of the formless, that fundamentally will always remain a mystery.

Can empirical science be done under these ontological limitations? This clearly can be done. We can treat the formless realm as a black-box and send towards it various questions and observe the answers that we receive. It might be the case that the questions that we can ask it and the answers that we receive to be infinite. And of course, whatever answers we might receive, they will never form a complete system. With ingenuity we might be able to develop clever patch theories that might apply in certain contexts, but never a full theory will be possible. Another aspect of such an empirical science will be that it will transform from an "objective" cold science to a participatory science, giving again the example of the roller-coaster. In a sense we already do this, though we don't fully appreciate its significance.

For example, we send rockets to the Moon because we have the visual quale of Moon. A blind person on the other hand, cannot even imagine the motivation of someone sending rockets to the Moon. But compared to our primitive 5 sense qualia domains that we have, there might be an infinite number of them, and what the participatory science of the future will do is to open portals towards those qualia domains and thus offer us unimaginable motivations for actions in those worlds and new ways to come up with empirical sciences to patch those realities as well. It is well known of cases throughout history in which people claimed that we reached the end of physics. The truth is, we didn't even start to scratch the surface. We just play around in these limited 5 sense qualia domains and we are so mesmerized by the dream of the "physical world" that we relentlessly research it as if that's where the enlightenment and the absolute truth lies. But we didn't even start. What awaits us is a world beyond our wildest imaginations.

In conclusion, the theory of self-reference developed in this paper, takes into account the unsurmountable difficulty of trying to formalize the observer and leaves it be in its own formless nature. As such, it is indeed a theory that has what is required to be correct. Of course, the details to be filled are probably infinite, but what it does is to give an introductory exposition of what a theory of reality requires. Made aware by these ingredients that a theory of reality requires, the reader can take the next steps of filling out the details and explore the most likely infinite complexities of the formless realm.

(*The End*)